Wind Energ. Sci., 11, 1205–1225, 2026
https://doi.org/10.5194/wes-11-1205-2026

eawe european academy of wind energy
Open Access WIND ENERGY SCIENCE

Bayesian uncertainty quantification of engineering models for wind farm–atmosphere interaction

Frederik Aerts, Koen Devesse, and Johan Meyers

Department of Mechanical Engineering, KU Leuven, Celestijnenlaan 300, 3001 Leuven, Belgium

Correspondence: Frederik Aerts (frederik.aerts1@kuleuven.be)

Received: 30 September 2025 – Discussion started: 22 October 2025
Revised: 19 December 2025 – Accepted: 28 January 2026 – Published: 14 April 2026

Abstract. Accurate modeling of wind farm–atmosphere interactions is critical for reliable energy yield assessments and flow control strategies. However, formal model comparison methodologies that quantify model form uncertainty by also accounting for parameter uncertainty are still lacking. This study presents an enhanced Bayesian uncertainty quantification framework for the calibration and validation of engineering wind farm flow models. Building on previous work, the framework explicitly incorporates model inadequacy through a parametrized model error distribution, enabling the separation of model and measurement uncertainties. The improved framework is demonstrated using a large-eddy simulation dataset for wind farm blockage and atmospheric gravity waves in conventionally neutral boundary layers. Two models of differing fidelity – a standard Gaussian wake model and an atmospheric perturbation model (APM) – are calibrated and compared. The posterior distribution of the model parameters reveals insights into model behavior and highlights areas for further improvement, for instance, when estimated parameter values are inconsistent across the model chain. In addition, it is shown that not explicitly incorporating model inadequacy results in an overly confident posterior distribution and renders derived stochastic flow models incapable of representing model uncertainty. A comparison of the quantified model uncertainty shows that the APM has significantly lower uncertainty than a standard wake model for this dataset, as the wake model is unable to represent wind farm blockage effects. This demonstrates the utility of the framework for objective model comparison with quantified parameter and model uncertainty given a reference dataset. Both the framework and the parallelized sequential Monte Carlo algorithm for accelerated posterior sampling are made available through the open-source Python package UMBRA.

1 Introduction

Wind farm flow model bias and uncertainty directly impact both the profitability of wind projects through preconstruction energy yield assessments (Lee and Fields, 2021) and the financial targets of wind developers through production forecasts (Ørsted, 2019). Although the estimation bias in annual energy predictions (AEPs) has steadily declined in the last 2 decades, the uncertainty remained of similar magnitude (Lee and Fields, 2021). Historically, wind farm performance is one of the largest contributors to AEP uncertainty, in part due to the uncertainty of the power losses on downstream turbines due to turbine wake effects (Clifton et al., 2016; Lee and Fields, 2021). Although models for these wake losses can reproduce trends in benchmark observations, their precision is highly variable, motivating the use of more precise and delineated observations of wind farms under many different operating and atmospheric conditions instead of those averaged over long periods of time (Moriarty et al., 2014). Moreover, the increasing capacity density and size of wind turbines require new wind farm flow models that consider the atmospheric boundary layer from the surface to the free atmosphere to model effects such as wind farm blockage (Allaerts and Meyers, 2019) and wakes (Bastankhah et al., 2024). To turn new wind farm flow models into reliable and cost-effective tools, objective methods are needed to validate them, quantify their uncertainty, and eventually calibrate them with a wide variety of flow conditions (Rodrigo et al., 2017). The present article studies the

use of Bayesian uncertainty quantification (UQ) as an objective method for model comparison given a reference dataset and calibration with quantified parameter uncertainty.

Many small- to large-scale benchmarking studies have validated the suitability of wind turbine wake models to represent energy losses in downstream turbines with historical power data (see Doekemeijer et al., 2022, for a comprehensive overview). Typically, these studies compare metrics such as farm power, turbine power for a given bin of wind directions, and wake loss while using default wake model parameters (Doekemeijer et al., 2022). This is common practice, as it is the most objective way of quantifying the baseline performance of the models (Rodrigo et al., 2017). However, it is known that the wake model bias is site-specific (Nygaard, 2015) and that the wake recovery differs between offshore and onshore wind farms (Barthelmie et al., 2009; Göçmen et al., 2016) and with atmospheric conditions (Abkar and Porté-Agel, 2015; Niayifar and Porté-Agel, 2016; Klemmer and Howland, 2025). Therefore, site-specific model tuning is crucial for accurate production forecasts. Moreover, in wind farm flow control, the flow model may be tuned to site-specific data in an open- or closed-loop fashion to adequately represent the flow field at any time (Göçmen et al., 2022; Meyers et al., 2022). Therefore, any model validation procedure may benefit from including model calibration so that model performance in practical applications can be compared objectively.

Concerning the quantification of uncertainty, one must distinguish between forward (data-free) and inverse (data-driven) UQ methods (Xiao and Cinnella, 2019). Forward UQ examines the effect of prespecified uncertainties on model inputs on the model outcome and is widely used to quantify the effect of wind resource variability on AEP estimates (Lackner et al., 2007; Kwon, 2010; Clerc et al., 2012). The procedure of Gaumond et al. (2014) to assess the effect of wind direction uncertainty on the predicted power of wake models also adheres to this approach. Despite being rigorous, forward UQ relies on the estimates of the constituent uncertainties, which may be subjective (Nygaard, 2015). Inverse UQ estimates the uncertainty of the model, and possibly its parameters, by comparing it with measured data. The estimated distribution of the discrepancy between model predictions and measurements determines the model uncertainty through its width and the model bias through its mean. With an inverse UQ using operational data from 19 offshore wind farms, Nygaard et al. (2022) showed that the uncertainty on the predicted wake loss relative to the observed wake loss is less than 10 % of the observed loss for the TurboPark model. This is significantly lower than previously estimated (Walker et al., 2015), in part due to thorough data processing and the inclusion of heterogeneous background flow. However, the uncertainty of the model is still overestimated because it is not separated from the experimental uncertainty. With Bayesian inverse UQ, it is possible to separate the measurement and model uncertainty, as demonstrated with operational power data from the Westermost Rough wind farm, while also accounting for the uncertainty of the model parameters (Aerts et al., 2023).

In Bayesian UQ, the model parameters and sources of uncertainties are characterized by probability distributions, which represent our knowledge about them and which can be updated when more data become available. Compared to deterministic model calibration methods (see van Binsbergen et al., 2024, for a comprehensive overview), Bayesian calibration gives not only the "best" parameters, but a joint posterior distribution with information on the parameter uncertainties and their correlations given the dataset (Aerts et al., 2023). The posterior parameter distribution may inform modelers about missing physics, when parameter values are estimated differently throughout the model chain. For example, LoCascio et al. (2023) showed that the posterior mean wake expansion rate and its uncertainty differ for different wake-merging methods, which is a cautionary finding for a modular approach to wake modeling. In addition, Zhang and Zhao (2020) and LoCascio et al. (2023) have proposed using the posterior distribution of the model parameters to obtain stochastic wake models, which can be used in wake steering under uncertainty (Howland, 2021). However, the current approaches to obtain such stochastic models include only the epistemic uncertainty (i.e., due to limited data) on the model parameters and not the model uncertainty due to varying physical phenomena not captured by the (deterministic) model.

In this study, we improve on a previously developed Bayesian UQ framework (Aerts et al., 2023) and demonstrate its use in a controlled environment with large model uncertainty. To this end, we select a large-eddy simulation dataset for blockage due to atmospheric gravity waves (Lanzilao and Meyers, 2024) as reference data and perform an inverse UQ for a standard wake model and a recently developed atmospheric perturbation model (Devesse et al., 2024b). Specifically, we present how the model uncertainty can be incorporated into the Bayesian framework to obtain stochastic models that include model uncertainty. In addition, we illustrate how the adequate inclusion of model uncertainty is crucial to obtain a correct posterior distribution of the empirical parameters, which is used in the current stochastic wake models. Lastly, we show how the Bayesian UQ framework can be used for objective model comparison with quantified model and parameter uncertainty. In contrast to previous studies (Zhang and Zhao, 2020; Aerts et al., 2023; LoCascio et al., 2023), which relied on inherently serial Markov chain Monte Carlo (MCMC) algorithms such as (Adaptive) Metropolis Hastings (Haario et al., 2001) and Hamiltonian Monte Carlo with No-U-Turn tuning (Hoffman and Gelman, 2014) to approximate the posterior distribution, we employ an inherently parallel sampling algorithm that still performs so-called "exact" posterior inference. The parallelized sampler and Bayesian framework are made available in an open-

source Python package coined UMBRA: Uncertainty Modeling toolbox for Bayesian data Re-Analysis.

The improved Bayesian UQ framework is presented in Sect. 2, together with the inherently parallel algorithm to sample the posterior. The setup of the demonstration case is introduced in Sect. 3, by presenting the essential parts of the wind farm flow models and dataset. The results of the inverse UQ analyses are presented in Sect. 4, with emphasis on the framework's adequacy (Sect. 4.1), its application to objective model comparison with UQ (Sect. 4.2), and the generalizability of the findings (Sect. 4.3). A summary and outlook are given in Sect. 5.

2 Bayesian uncertainty quantification

We first discuss how different sources of uncertainty are included in the UQ framework in Sect. 2.1, building on earlier work (Aerts et al., 2023). Section 2.2 demonstrates how this formulation naturally leads to Bayesian updating from a prior to a posterior distribution and introduces the posterior predictive distribution as a key tool for validation in Bayesian UQ. Once the framework is introduced, we demonstrate the consequences of neglecting the model error in Sect. 2.3. The parallelizable algorithm employed for posterior sampling is detailed in Sect. 2.4.

2.1 Sources of uncertainty

Wind farm flow models aim to predict quantities of interest for a given atmospheric state. When focusing on turbine-level power output, the idealized "true process" $\boldsymbol{T}(\boldsymbol{\psi})$ would yield the power $\boldsymbol{\mathcal{P}} \in \mathbb{R}^{N_\mathrm{T}}$ of the N_T turbines, normalized by the output of an undisturbed upstream turbine, given a complete description of the wind farm and atmospheric state in a general vector $\boldsymbol{\psi}$:

$$\boldsymbol{\mathcal{P}} = \boldsymbol{T}(\boldsymbol{\psi}). \tag{1}$$

In practice, the true process is approximated by a model $\boldsymbol{M}(\boldsymbol{\vartheta}_\mathrm{e}, \boldsymbol{\varphi})$, which relies on empirical parameters $\boldsymbol{\vartheta}_\mathrm{e}$ and a partial state description $\boldsymbol{\varphi}$. This approximation introduces model bias and uncertainty, which we would like to represent and quantify (Sect. 2.1.1). Since the true process is only accessible through measurements, measurement uncertainty also arises (Sect. 2.1.2). Moreover, the most representative values of $\boldsymbol{\vartheta}_\mathrm{e}$ for the given conditions are initially unknown. Section 2.1.3 describes how our initial knowledge about the model parameters, bias, and uncertainty can be represented in a prior probability density. In what follows, we write a random variable and its realization with a calligraphic symbol $\boldsymbol{\mathcal{T}}$, its probability distribution as $p(\boldsymbol{\mathcal{T}})$, and the function that uniquely determines it from other random variables as $\boldsymbol{T}(\boldsymbol{\psi})$ so that $p(\boldsymbol{\mathcal{T}}|\boldsymbol{\psi}) = \delta(\boldsymbol{\mathcal{T}} - \boldsymbol{T}(\boldsymbol{\psi}))$, with $\delta(\cdot)$ the Dirac delta distribution.

2.1.1 Model error

Since the wind farm flow model is typically imperfect, we can define an additive model error $\boldsymbol{E}_B(\boldsymbol{\vartheta}_\mathrm{e}, \boldsymbol{\psi})$ that depends on the choice of the model parameters $\boldsymbol{\vartheta}_\mathrm{e}$ and the state of the wind farm and atmosphere $\boldsymbol{\psi}$:

$$\boldsymbol{T}(\boldsymbol{\psi}) \triangleq \boldsymbol{M}(\boldsymbol{\vartheta}_\mathrm{e}, \boldsymbol{\varphi}) + \boldsymbol{E}_B(\boldsymbol{\vartheta}_\mathrm{e}, \boldsymbol{\psi}). \tag{2}$$

Note that the model error is, in fact, deterministic given $\boldsymbol{\vartheta}_\mathrm{e}$ and $\boldsymbol{\psi}$, so that its distribution $p(\boldsymbol{\mathcal{E}}_B|\boldsymbol{\vartheta}_\mathrm{e}, \boldsymbol{\psi}) = \delta(\boldsymbol{\mathcal{E}}_B - \boldsymbol{E}_B(\boldsymbol{\vartheta}_\mathrm{e}, \boldsymbol{\psi}))$. However, since usually only a subset $\boldsymbol{\varphi}$ of the variables $\boldsymbol{\psi} = (\boldsymbol{\varphi}, \boldsymbol{\psi}')$ describing the atmospheric conditions is available or included as input to the model, we are interested in the distribution of the true process $\boldsymbol{\mathcal{T}}$ conditioned on the observed conditions $\boldsymbol{\varphi}$, given by

$$p(\boldsymbol{\mathcal{T}}|\boldsymbol{\varphi}) = \int p(\boldsymbol{\mathcal{T}}|\boldsymbol{\psi}) p(\boldsymbol{\psi}'|\boldsymbol{\varphi}) \mathrm{d}\boldsymbol{\psi}' \tag{3a}$$

$$= \int \delta(\boldsymbol{M}(\boldsymbol{\vartheta}_\mathrm{e}, \boldsymbol{\varphi}) + \boldsymbol{E}_B(\boldsymbol{\vartheta}_\mathrm{e}, \boldsymbol{\psi}) - \boldsymbol{\mathcal{T}}) p(\boldsymbol{\psi}'|\boldsymbol{\varphi}) \mathrm{d}\boldsymbol{\psi}'. \tag{3b}$$

The unmodeled or unobserved variations in atmospheric conditions $\boldsymbol{\psi}'$ thus introduce uncertainty in the conditional process $\boldsymbol{\mathcal{T}}|\boldsymbol{\varphi}$ through their influence on the model error $\boldsymbol{E}_B(\boldsymbol{\vartheta}_\mathrm{e}, \boldsymbol{\psi})$. However, if we allow the model parameters $\boldsymbol{\vartheta}_\mathrm{e}$ to depend on the unobserved conditions, the uncertainty on $\boldsymbol{\mathcal{T}}|\boldsymbol{\varphi}$ may also be reflected in a distribution of the model parameters $\boldsymbol{\vartheta}_\mathrm{e}$. Hence, we have to choose whether to incorporate this uncertainty on the process $\boldsymbol{\mathcal{T}}|\boldsymbol{\varphi}$, which we will further refer to as model uncertainty, within the model parameters $\boldsymbol{\vartheta}_\mathrm{e}$ or in the model error $\boldsymbol{\mathcal{E}}_B$, by fixing the other. Allowing both the model parameters and the model uncertainty to vary with the unknown conditions $\boldsymbol{\psi}'$ causes them to partially contribute in the same way to the total model uncertainty. Such similar contributions to the total model uncertainty are indistinguishable when quantifying the constituent uncertainties from data, rendering the inverse problem underdetermined.

The first option is to incorporate the model uncertainty within the model parameters through a hierarchical stochastic prior (see also Sargsyan et al., 2015; Wu et al., 2018a). To this end, we fix the model error $\boldsymbol{E}_B(\boldsymbol{\vartheta}_\mathrm{e}, \boldsymbol{\psi}) = \boldsymbol{\mu}_B$, so that this relation defines the dependence of the model parameters $\boldsymbol{\vartheta}_\mathrm{e}$ on the wind farm–atmosphere state $\boldsymbol{\psi}$. Propagating the distribution of the unknown conditions $p(\boldsymbol{\psi}'|\boldsymbol{\varphi})$ through this implicit relation yields a distribution of $\boldsymbol{\vartheta}_\mathrm{e}$, which is unknown. If we are only interested in the first- and second-order moments, we can parametrize it as a normal distribution $\vartheta_\mathrm{e} \sim \mathcal{N}(\mu_\vartheta, \sigma_\vartheta^2)$ with mean μ_ϑ and variance σ_ϑ^2 for a scalar parameter $\vartheta_\mathrm{e} \in \mathbb{R}$, which is the maximum entropy distribution in that case (McElreath, 2018). A change in variables then gives

$$p(\boldsymbol{\mathcal{T}}|\boldsymbol{\varphi}) = \int \delta(\boldsymbol{M}(\vartheta_\mathrm{e}, \boldsymbol{\varphi}) + \boldsymbol{\mu}_B - \boldsymbol{\mathcal{T}}) \mathcal{N}\big(\vartheta_\mathrm{e}; \mu_\vartheta, \sigma_\vartheta^2\big) \mathrm{d}\vartheta_\mathrm{e} \tag{4a}$$

$$\approx \mathcal{N}\big(\boldsymbol{\mathcal{T}}; \boldsymbol{M}(\mu_\vartheta, \boldsymbol{\varphi}) + \boldsymbol{\mu}_B, \sigma_\vartheta^2 \boldsymbol{J}_M \boldsymbol{J}_M^\top\big), \tag{4b}$$

where the approximation corresponds to a linearization of the model $\boldsymbol{M}(\vartheta_\mathrm{e}, \boldsymbol{\varphi}) \approx \boldsymbol{M}(\mu_\vartheta, \boldsymbol{\varphi}) + \boldsymbol{J}_M(\vartheta_\mathrm{e} - \mu_\vartheta)$ with the Jacobian $\boldsymbol{J}_M \in \mathbb{R}^{N_\mathrm{T}}$. Hence, the model uncertainty is determined by the model parameter uncertainty σ_ϑ and the model structure through the Jacobian $\boldsymbol{J}_M$. However, the model uncertainty may have a different structure than the model itself. Taking the example of a wake model with an unknown wake expansion rate, the model uncertainty on the upstream turbines due to blockage cannot be captured in this approach, as the sensitivity of the upstream turbine power to the wake expansion rate is zero. Hence, an additional model error term is generally needed. Moreover, a forward UQ is still necessary to translate the uncertainty on the model parameters to uncertainty on the model output, which is typically of interest for many practical applications.

The second approach attributes the model uncertainty only to the model error term. To that end, we fix the model parameter values $\boldsymbol{\vartheta}_\mathrm{e}$ and let the model error $\boldsymbol{E}_B(\boldsymbol{\vartheta}_\mathrm{e}, \boldsymbol{\psi})$ vary with the unknown conditions. Hence, the model uncertainty is represented in the distribution of the model error $p(\boldsymbol{\mathcal{E}}_B|\boldsymbol{\vartheta}_\mathrm{e}, \boldsymbol{\varphi})$, which is no longer a Dirac delta if a part $\boldsymbol{\psi}'$ of the wind farm–atmosphere state $\boldsymbol{\psi} = (\boldsymbol{\varphi}, \boldsymbol{\psi}')$ is unknown. The distribution $p(\boldsymbol{\mathcal{E}}_B|\boldsymbol{\vartheta}_\mathrm{e}, \boldsymbol{\varphi})$ is a priori unknown, but if we are mainly interested in its first- and second-order moments, we can parametrize it as a multivariate normal distribution with expected value $\boldsymbol{\mu}_B$ and covariance matrix $\boldsymbol{\Sigma}_B$:

$$p(\boldsymbol{\mathcal{E}}_B|\boldsymbol{\vartheta}_\mathrm{e}, \boldsymbol{\varphi}) = \int p(\boldsymbol{\mathcal{E}}_B|\boldsymbol{\vartheta}_\mathrm{e}, \boldsymbol{\psi}) p(\boldsymbol{\psi}'|\boldsymbol{\varphi}) \mathrm{d}\boldsymbol{\psi}' \approx \mathcal{N}(\boldsymbol{\mathcal{E}}_B; \boldsymbol{\mu}_B, \boldsymbol{\Sigma}_B). \tag{5}$$

We will further approximate $p(\boldsymbol{\mathcal{E}}_B|\boldsymbol{\vartheta}_\mathrm{e}, \boldsymbol{\varphi}) \approx p(\boldsymbol{\mathcal{E}}_B|\boldsymbol{\varphi})$ (Kennedy and O'Hagan, 2001), but since the model error depends on the choice of model parameters, this requires including sufficient prior knowledge on the mean model error $\boldsymbol{\mu}_B$ (Brynjarsdóttir and O'Hagan, 2014), as will be discussed in Sect. 2.1.3. This second approach yields a model error distribution that is independent of the model structure and is readily interpreted as uncertainty on either the model or the model error due to additivity. To allow a general model error uncertainty parametrization while avoiding the problem becoming underdetermined, we opt for the second approach.

The current parametrization of the model error distribution scales with the number of turbines N_T in the farm since the mean model error $\boldsymbol{\mu}_B \in \mathbb{R}^{N_\mathrm{T}}$ and its covariance matrix $\boldsymbol{\Sigma}_B \in \mathbb{R}^{N_\mathrm{T} \times N_\mathrm{T}}$. To reduce its dimensionality, we follow the same approach as in our previous work (Aerts et al., 2023). The correlations in $\boldsymbol{\Sigma}_B$ are neglected, and the mean and standard deviation of the model error on the power of the turbine i are binned based on the number of upstream turbines $\zeta(i)$ that cause a wake loss on the turbine i greater than 1 % of the free stream wind speed. Hence,

$$\mu_{B,i} = \delta_{\zeta(i)} \tag{6}$$

$$\Sigma_{B,ij} = \delta_{ij} \sigma^2_{B,\zeta(i)}, \tag{7}$$

where δ_{ij} is the Kronecker delta. As such, the model error distribution is parametrized by the parameters $\boldsymbol{\vartheta}_b = \{\delta_\zeta, \sigma_{B,\zeta}\}_{\zeta=0}^{\zeta_\mathrm{max}}$ with $\zeta = 0$ as the index for the upstream turbines, $\zeta = 1$ as the index for turbines with one upstream turbine that wakes them, and so forth.

2.1.2 Measurement error

The true process can only be observed with measurements $\tilde{\boldsymbol{\mathcal{P}}}$, which come with a measurement error $\boldsymbol{\mathcal{E}}_\mathrm{M}$:

$$\tilde{\boldsymbol{\mathcal{P}}} = \boldsymbol{\mathcal{T}} + \boldsymbol{\mathcal{E}}_\mathrm{M}. \tag{8}$$

We will make the simplifying assumption that the measurement error is independent of the unobserved or unmodeled physics $\boldsymbol{\psi}'$ such that $\boldsymbol{\mathcal{E}}_\mathrm{M}|\boldsymbol{\psi} \approx \boldsymbol{\mathcal{E}}_\mathrm{M}|\boldsymbol{\varphi}$ (Aerts et al., 2023). Since most engineering wind farm flow models represent stationary atmospheric flows, their predictions should be compared with time-averaged data. However, observational and high-fidelity simulation data are typically subjected to temporally resolved turbulence. Because only a finite time period is available for averaging due to changing atmospheric conditions or computational constraints, the measurement error consists of both the error of the apparatus and the averaging error. The measurement bias and standard deviation due to the apparatus are typically known a priori, such that the distribution of the apparatus error can be taken as a normal distribution based on these quantities. In what follows, we will presume that the averaging error dominates. This is certainly true for simulation data, as is the case in this article. Due to the central limit theorem and given that the estimator is unbiased, the measurement error then follows a multivariate normal distribution with zero mean and a covariance matrix $\boldsymbol{\Sigma}_T$.

The averaging error covariance matrix $\boldsymbol{\Sigma}_T$ on the mean can either be prespecified or unknown. For independent measurements for the same atmospheric condition, the uncertainty on the average of the measurements can be estimated as the sample variance divided by the number of measurements (Wasserman, 2004). Equivalently, all individual measurements can be used with the sample variance, given that they are independent and represent the same atmospheric state. For a correlated time series, the moving block bootstrap can be employed (Garcia et al., 2005). If no information is available, the averaging error covariance can also be estimated directly from the data based on a parametrization with parameters $\boldsymbol{\vartheta}_t$ (Aerts et al., 2023). However, if the covariance structures between the averaging error and the model error are not sufficiently different, they are indistinguishable. Therefore, it is preferred to use the estimated uncertainty on the mean if available.

In practice, the inflow conditions can also be uncertain due to any kind of measurement error. This inflow uncertainty

may be propagated through the model with a marginalization similar to that in Eq. (4). Note that this procedure is similar to that of Gaumond et al. (2014) to incorporate the effect of wind direction variability on the mean power but also includes the resulting variance of the model output. In that manner, a part of the total variance in the data can be attributed to inflow uncertainty, thereby reducing the observed model uncertainty.

2.1.3 Prior parameter uncertainty

The most representative empirical model parameters of the wind farm flow model $\boldsymbol{\vartheta}_\mathrm{e}$ are a priori uncertain, but the associated model error distribution, parametrized by $\boldsymbol{\vartheta}_b$, and possibly measurement error covariance, parametrized by $\boldsymbol{\vartheta}_t$, are as well. This a priori uncertainty can be quantified or specified in a prior distribution $p(\boldsymbol{\vartheta}_\mathrm{e}, \boldsymbol{\vartheta}_b, \boldsymbol{\vartheta}_t)$, which reflects one's assumptions and state of knowledge before data come along (Trotta, 2008). We will use a weakly informative prior, which is designed to regularize inferences with structural information (Gelman et al., 2017). The provided information is intentionally weaker than any actual prior knowledge available (Gelman et al., 2013), and we choose the shape of the distribution to have the highest entropy given the provided information. Since the joint prior has maximum entropy when the parameters are not correlated, the prior is constructed as the product of the marginal priors. Typically, we know what the range of reasonable or allowable values is for the model parameters. In that case, the proper distribution with maximum entropy is a uniform distribution (Toussaint, 2011). As the exponential distribution has maximum entropy among all non-negative continuous distributions with the same average displacement (McElreath, 2018), the standard deviations of the model error terms are assigned exponential priors with averages of 0.1. This is deemed weakly informative given that the power of an undisturbed turbine is normalized to one.

In Bayesian calibration, particular attention must be given to the choice of the prior distribution for the mean model error or bias $\boldsymbol{\mu}_B$. In the Kennedy and O'Hagan (2001) framework for Bayesian calibration used previously (Aerts et al., 2023), model inadequacy is a priori considered independent of the model output. To make the model parameters identifiable, we constrained the bias on the farm power to be zero by solving

$$0 = \sum_{i=1}^{N_\mathrm{T}} \mu_{B,i} = \sum_{i=1}^{N_\mathrm{T}} \delta_{\zeta(i)} \tag{9}$$

for δ_0 and only estimating $\{\delta_\zeta\}_{\zeta=1}^{\zeta_\mathrm{max}}$ (Aerts et al., 2023). However, the value of the model error $E_{B,i}(\boldsymbol{\vartheta}_\mathrm{e}, \boldsymbol{\psi})$ for each turbine i depends on the choice of model parameters $\boldsymbol{\vartheta}_\mathrm{e}$, so simultaneously identifying both the model error and model parameters may introduce confounding of the model error with calibration parameters (Brynjarsdóttir and O'Hagan, 2014). Therefore, it is more intuitive to define the mean bias as the discrepancy $\mu_{B,i} = \mathbb{E}_\varphi[\tilde{\mathcal{P}}_i - M_i(\hat{\boldsymbol{\vartheta}}_\mathrm{e}, \boldsymbol{\varphi})]$ that remains when the model is calibrated with a "best-fit" parameter $\hat{\boldsymbol{\vartheta}}_\mathrm{e}$ (Plumlee, 2017). If that best fit is defined as $\hat{\boldsymbol{\vartheta}}_\mathrm{e} = \operatorname{argmin} \mathbb{E}_\varphi\left[\sum_{i=1}^{N_\mathrm{T}} (\tilde{\mathcal{P}}_i - M_i(\boldsymbol{\vartheta}_\mathrm{e}, \boldsymbol{\varphi}))^2\right]$, we have as a necessary condition for optimality that for every parameter $\vartheta_{\mathrm{e,m}}$

$$\mathbb{E}_\varphi\left[\sum_{i=1}^{N_\mathrm{T}} \left.\frac{\partial M_i(\boldsymbol{\vartheta}_\mathrm{e}, \boldsymbol{\varphi})}{\partial \vartheta_{\mathrm{e,m}}}\right|_{\boldsymbol{\vartheta}_\mathrm{e}=\hat{\boldsymbol{\vartheta}}_\mathrm{e}} \mu_{B,i}\right] = 0. \tag{10}$$

Since we do not know the "best-fit" parameters a priori, we can satisfy this condition trivially by requiring that $\mu_{B,i} = 0$ if $\partial M_i(\boldsymbol{\vartheta}_\mathrm{e}, \boldsymbol{\varphi})/\partial \vartheta_{\mathrm{e,m}} \neq 0$ for at least one value of $\vartheta_{\mathrm{e,m}}$ with nonzero prior probability. As a result, the farm power predicted by the calibrated model is allowed to be biased if $\partial M_i(\boldsymbol{\vartheta}_\mathrm{e}, \boldsymbol{\varphi})/\partial \vartheta_{\mathrm{e,m}} = 0$ for all $\vartheta_{\mathrm{e,m}}$ with nonzero prior probability. However, we find that also requiring $\mu_{B,i} = 0$ in that case works best in practice, but alternatives are explored in Sect. 4.1. Hence, the model bias parameters $\{\delta_\zeta\}_{\zeta=0}^{\zeta_\mathrm{max}}$ are all given a Dirac delta distribution centered at zero as marginal prior.

2.2 Bayesian updating

The objective of Bayesian uncertainty quantification is to construct and interpret the posterior distribution of the model parameters after Bayesian updating. The construction of the posterior based on the prior and the preceding description of the uncertainties is discussed in Sect. 2.2.1. The posterior distribution can be used as the constituent distribution to perform a forward uncertainty quantification or obtain a stochastic flow model. In this manner, the posterior predictive distribution is obtained, which can be used to validate the adequacy of the UQ procedure as explained in Sect. 2.2.2.

2.2.1 Posterior distribution

Given a power measurement $\tilde{\boldsymbol{\mathcal{P}}}$ and the corresponding input of the model $\boldsymbol{\varphi}$ describing the state of the wind farm and the atmosphere, the prior distribution $p(\boldsymbol{\vartheta}) = p(\boldsymbol{\vartheta}_\mathrm{e}, \boldsymbol{\vartheta}_b, \boldsymbol{\vartheta}_t)$ can be updated using Bayes' theorem to a posterior distribution:

$$p\left(\boldsymbol{\vartheta}|\tilde{\boldsymbol{\mathcal{P}}}, \boldsymbol{\varphi}\right) = \frac{p\left(\tilde{\boldsymbol{\mathcal{P}}}|\boldsymbol{\vartheta}, \boldsymbol{\varphi}\right) p(\boldsymbol{\vartheta})}{p\left(\tilde{\boldsymbol{\mathcal{P}}}|\boldsymbol{\varphi}\right)}. \tag{11}$$

The likelihood $p(\tilde{\boldsymbol{\mathcal{P}}}|\boldsymbol{\vartheta}, \boldsymbol{\varphi})$ of the power measurement $\tilde{\boldsymbol{\mathcal{P}}}$, given the parameters $\boldsymbol{\vartheta}$ and input to the model $\boldsymbol{\varphi}$, is given by

$$p\left(\tilde{\boldsymbol{\mathcal{P}}}|\boldsymbol{\vartheta}, \boldsymbol{\varphi}\right) = \mathcal{N}\big(\tilde{\boldsymbol{\mathcal{P}}}; \boldsymbol{M}(\boldsymbol{\vartheta}_\mathrm{e}, \boldsymbol{\varphi}) + \boldsymbol{\mu}_B(\boldsymbol{\vartheta}_b), \boldsymbol{\Sigma}_T(\boldsymbol{\vartheta}_t) + \boldsymbol{\Sigma}_B(\boldsymbol{\vartheta}_b)\big), \tag{12}$$

based on the description of model uncertainty in Sect. 2.1.1 and measurement uncertainty in Sect. 2.1.2. The evidence

$p(\tilde{\boldsymbol{\mathcal{P}}}|\boldsymbol{\varphi})$ does not depend on the parameters $\boldsymbol{\vartheta}$ and corresponds to a normalization factor. As a result, the posterior is fully determined by the prior and the likelihood.

For a dataset $\mathcal{D} = \{\tilde{\boldsymbol{\mathcal{P}}}_n, \boldsymbol{\varphi}_n\}_{n=1}^{N_D}$ of independent power measurements $\tilde{\boldsymbol{\mathcal{P}}}_n$ with corresponding inputs to the model $\boldsymbol{\varphi}_n$, the posterior is given by

$$p(\boldsymbol{\vartheta}|\mathcal{D}) \propto p(\boldsymbol{\vartheta}) \prod_{n=1}^{N_D} \mathcal{N}\big(\tilde{\boldsymbol{\mathcal{P}}}_n; \boldsymbol{M}(\boldsymbol{\vartheta}_e, \boldsymbol{\varphi}_n) + \boldsymbol{\mu}_B(\boldsymbol{\vartheta}_b), \boldsymbol{\Sigma}_T(\boldsymbol{\vartheta}_t) + \boldsymbol{\Sigma}_B(\boldsymbol{\vartheta}_b)\big), \tag{13}$$

where the total likelihood is a product of the individual likelihoods due to independence and the prior $p(\boldsymbol{\vartheta})$ is given by the product of the marginal priors as discussed in Sect. 2.1.3. In the limit of an infinite number of data, the posterior converges to a point mass, given that the parameters are identifiable – see Gelman et al. (2013, p. 89) for other conditions. For a finite but large number of data, the relative uncertainty of each of the model parameters in the posterior is inversely related to the sensitivity of the log-likelihood to that parameter, through the Fisher information matrix (Gelman et al., 2013, p. 88). Hence, the posterior parameter uncertainty represents epistemic uncertainty that can be reduced with more observations. Irreducible forms of uncertainty, such as measurement and model uncertainty, are quantified by their parametrization in the likelihood: $\boldsymbol{\Sigma}_T$ and $\boldsymbol{\Sigma}_B$ here. It is crucial that these forms of uncertainty are adequately quantified, as otherwise the marginal posterior for the model parameters $p(\boldsymbol{\vartheta}_e|\mathcal{D})$ may be over-confident and biased (Brynjarsdóttir and O'Hagan, 2014).

2.2.2 Posterior predictive distribution

The posterior predictive is the distribution of new (predicted) observations given all previous observations, the wind farm flow model, and the description of all sources of uncertainty in the likelihood and prior. For the Bayesian UQ analysis to be adequate, the original data should seem plausible under the posterior predictive distribution (Gelman et al., 2013, p. 143). Any systematic differences between the posterior predictions and the data indicate potential failings of the specified likelihood and prior to model the actual process that generates the data (cf. Eqs. 2 and 8). For instance, if the model error uncertainty is not included in the analysis, the posterior predictive may underestimate the variance of the data. In that case, a posterior predictive check will reveal the inadequacy of the specified likelihood and prior.

The posterior predictive distribution $p(\tilde{\boldsymbol{\mathcal{P}}}_{\text{new}}|\boldsymbol{\varphi}_{\text{new}}, \mathcal{D})$ can be rewritten as

$$p(\tilde{\boldsymbol{\mathcal{P}}}_{\text{new}}|\boldsymbol{\varphi}_{\text{new}}, \mathcal{D}) = \int p(\tilde{\boldsymbol{\mathcal{P}}}_{\text{new}}, \boldsymbol{\vartheta}|\boldsymbol{\varphi}_{\text{new}}, \mathcal{D}) \mathrm{d}\boldsymbol{\vartheta}, \tag{14a}$$

$$\overset{(1)}{=} \int p(\tilde{\boldsymbol{\mathcal{P}}}_{\text{new}}|\boldsymbol{\vartheta}, \boldsymbol{\varphi}_{\text{new}}) p(\boldsymbol{\vartheta}|\mathcal{D}, \boldsymbol{\varphi}_{\text{new}}) \mathrm{d}\vartheta, \tag{14b}$$

$$\overset{(2)}{=} \int p(\tilde{\boldsymbol{\mathcal{P}}}_{\text{new}}|\boldsymbol{\vartheta}, \boldsymbol{\varphi}_{\text{new}}) p(\boldsymbol{\vartheta}|\mathcal{D}) \mathrm{d}\boldsymbol{\vartheta}, \tag{14c}$$

where (1) requires that the new measurement is again independent from the previous ones and (2) presumes that the posterior based on the previously observed states of the wind farm and atmosphere is independent of the new state. In practice, this means that the calibrated model and quantified model uncertainty should also be adequate for the new inflow condition $\boldsymbol{\varphi}_{\text{new}}$ (more on that in Sect. 4.3). Samples from the posterior predictive for a given model input $\boldsymbol{\varphi}^*_{\text{new}}$ are obtained by first sampling the posterior distribution $\boldsymbol{\vartheta}^*_e, \boldsymbol{\vartheta}^*_b \sim p(\boldsymbol{\vartheta}|\mathcal{D})$ and then sampling from the likelihood $\tilde{\boldsymbol{\mathcal{P}}}^*_{\text{new}} \sim p(\tilde{\boldsymbol{\mathcal{P}}}_{\text{new}}|\boldsymbol{\vartheta}^*_e, \boldsymbol{\vartheta}^*_b, \boldsymbol{\phi}^*_{\text{new}})$ given the sampled parameters. Consequently, it can be interpreted as the forward UQ of the model given the epistemic uncertainty in the posterior, and the measurement and model uncertainty in the likelihood. By leaving out the measurement uncertainty, one obtains a stochastic flow model that accounts for both the epistemic uncertainty on the model parameters and the (systematic) model uncertainty.

2.3 Consequences of neglecting model error in Bayesian UQ

To illustrate what goes wrong when the model error is not properly included in the Bayesian framework, we consider a simple model for the farm power

$$P_f = \vartheta_e N_T P_\infty, \tag{15}$$

where the empirical parameter ϑ_e represents the efficiency of the wind farm and P_∞ is the power of a hypothetical undisturbed turbine upstream. As data, we take the time-averaged wind farm power, normalized by $N_T P_\infty$, $\{\tilde{\mathcal{P}}_{f,n}\}_{n=1}^{N_D}$ for $N_D = 9$ large-eddy simulations of a large wind farm operating in different atmospheric stratification regimes (introduced in more detail in Sect. 3.2). Although the wind farm efficiency is highly variable for this dataset, the model parameter ϑ_e in our simple model is assumed to be fixed. Since this assumption is clearly invalid in the present case, properly accounting for model error becomes essential.

2.3.1 Posterior distribution

The Bayesian framework yields the following joint posterior distribution for the model parameter ϑ_e and the standard deviation of the model error σ_B through Eq. (13):

$$p(\vartheta_e, \sigma_B|\mathcal{D}) \propto p(\vartheta_e) p(\sigma_B) \times \prod_{n=1}^{N_D} \mathcal{N}\left(\tilde{\mathcal{P}}_{f,n}; \vartheta_e, \sigma_T^2 + \sigma_B^2\right). \tag{16}$$

The uncertainty due to finite-time averaging σ_T is obtained for simplicity as the average of the bootstrap estimates $\sigma_{T,n}$

for each simulation. The marginal prior of the standard deviation of the model error is an exponential distribution $p(\sigma_B) = \lambda^{-1}\exp(-\sigma_B/\lambda)$ with mean $\lambda = 0.1$. The marginal prior of the wind speed reduction factor is a normal distribution with mean μ_0 and standard deviation σ_0, that is, $p(\vartheta_e) = \mathcal{N}(\vartheta_e; \mu_0, \sigma_0^2)$ with $\mu_0 = 0.5$ and $\sigma_0 = 0.1$.

We can examine the effect of neglecting model uncertainty by comparing the conditional posterior $p(\vartheta_e|\sigma_B, \mathcal{D})$ with $\sigma_B = 0$ and with σ_B equal to the mode of the marginal posterior $p(\sigma_B|\mathcal{D})$. For this simple example, the conditional posterior of the model parameter ϑ_e is a normal distribution $p(\vartheta_e|\sigma_B, \mathcal{D}) = \mathcal{N}(\vartheta_e; \mu_{N_D}, \sigma_{N_D}^2)$ with mean μ_{N_D} and variance $\sigma_{N_D}^2$, equal to

$$\mu_{N_D} = \frac{\sigma_0^2 \sum_{n=1}^{N_D} \tilde{\mathcal{P}}_{f,n} + \sigma_M^2 \mu_0}{N_D \sigma_0^2 + \sigma_M^2}, \tag{17}$$

$$\sigma_{N_D}^2 = \frac{\sigma_0^2 \sigma_M^2}{N_D \sigma_0^2 + \sigma_M^2}, \tag{18}$$

where $\sigma_M^2 = \sigma_B^2 + \sigma_T^2$ is the total variance. Note that by increasing the number of measurements N_D, the posterior indeed converges to a point mass, in this case centered at the sample mean $\langle\tilde{\mathcal{P}}_f\rangle = \sum_{n=1}^{N_D} \tilde{\mathcal{P}}_{f,n}/N_D$. Given enough data and for λ sufficiently large, the mode of $p(\sigma_B|\mathcal{D})$ becomes $\sqrt{s^2 - \sigma_T^2}$, where $s^2 = \sum_{n=1}^{N_D}(\tilde{\mathcal{P}}_{f,n} - \langle\tilde{\mathcal{P}}_f\rangle)^2/N_D$ is the sample variance. Consequently, the estimated model error variance σ_B^2 will capture the remaining variance in the dataset, after subtracting the variance related to finite-time averaging errors σ_T^2. If one assumes that $\sigma_B = 0$ when there is non-negligible model uncertainty, Eqs. (17) and (18) show that the posterior mean μ_{N_D} is biased and that the uncertainty σ_{N_D} underestimated. This is demonstrated for the example at hand in Fig. 1a, as the posterior distribution obtained by neglecting the model error is highly overconfident. Also note that the samples of the marginal posterior $p(\vartheta_e|\mathcal{D})$ obtained with UMBRA (see Sect. 2.4) agree very well with the analytical conditional posterior $p(\vartheta_e|\sigma_B, \mathcal{D})$ with σ_B equal to $\sqrt{s^2 - \sigma_T^2}$.

2.3.2 Posterior predictive distribution

The posterior predictive distribution for a new measurement $\tilde{\mathcal{P}}_{f,\text{new}}$ is given by

$$p\left(\tilde{\mathcal{P}}_{f,\text{new}}|\mathcal{D}\right) = \mathcal{N}\left(\tilde{\mathcal{P}}_{f,\text{new}}; \mu_{N_D}, \sigma_B^2 + \sigma_T^2 + \sigma_{N_D}^2\right). \tag{19}$$

The posterior predictive variance consists of the model error variance σ_B^2, the measurement variance σ_T^2, and the propagated posterior parameter variance $\sigma_{N_D}^2$. If the model error is not included in the analysis, the posterior predictive underestimates the variance of the data both by neglecting σ_B^2 and by underestimating $\sigma_{N_D}^2$. Figure 1b shows that in the current example, the data do not seem plausible under the posterior predictive that neglects model error, rendering such a Bayesian analysis inadequate. The proper inclusion of model error through our framework yields an adequate posterior predictive and Bayesian analysis. Although the current example exhibits exceptionally large model uncertainty, most, if not all, of the current wind farm flow models have non-negligible model error, and Bayesian UQ analyses of such models that neglect model error will suffer similar issues.

2.3.3 Implications for stochastic flow models

Current stochastic wake models are obtained by propagating the posterior of the parameters – obtained by ignoring model error – through the model (Zhang and Zhao, 2020). They only account for the uncertainty on their parameters due to limited calibration data, and they do not capture the uncertainty due to varying unmodeled physics, such as stratification effects and time-resolved turbulence. Given enough data (here only 9 observations), they will therefore significantly underestimate the variability of the true process, as demonstrated in Fig. 1b. Moreover, in the limit of infinite data, such "stochastic" models will in fact become deterministic, as seen in Eqs. (18) and (19) with $\sigma_{N_D} \to 0$ and $\sigma_B = 0$. Therefore, truly stochastic wind farm flow models should include both the posterior parameter uncertainty, quantified in $p(\boldsymbol{\vartheta}_e|\mathcal{D})$, and the model uncertainty, quantified through the model error covariance matrix $\boldsymbol{\Sigma}_B$ in the framework.

2.4 Sampling the posterior distribution

In practice, the analytical derivation of the moments and marginal distributions of the posterior distribution quickly becomes intractable. Instead, Markov chain Monte Carlo (MCMC) algorithms are typically used to efficiently sample the posterior distribution. With those samples, one can visualize the marginalized and joint posterior(s), consult the posterior predictive, and compute expected values. However, these algorithms are inherently serial and require $\mathcal{O}(10^5)$ likelihood evaluations to converge to the posterior and adequately represent it (Geyer, 2011). Even for engineering models with reasonable computational expense, this can become relatively time-consuming. Therefore, we employ a variant of the transitional Markov chain Monte Carlo (TMCMC) algorithm (Ching and Chen, 2007), which is inherently parallel.

Instead of directly sampling the posterior distribution, TMCMC samples a sequence of target distributions with a sequential Monte Carlo (SMC) method. This sequence $\pi_j(\boldsymbol{\vartheta})$ is obtained by tempering the likelihood: a stage exponent or inverse temperature β_j is introduced that sequentially "cools" the target

$$\pi_j(\boldsymbol{\vartheta}) \propto p\left(\mathcal{D}|\boldsymbol{\vartheta}\right)^{\beta_j} p(\boldsymbol{\vartheta}), \tag{20}$$

where $j = 1, \ldots, J$ and $0 = \beta_1 < \cdots < \beta_J = 1$, such that the algorithm transitions from the prior ($j = 1$), which is typically easy to sample from, to the posterior ($j = J$). This

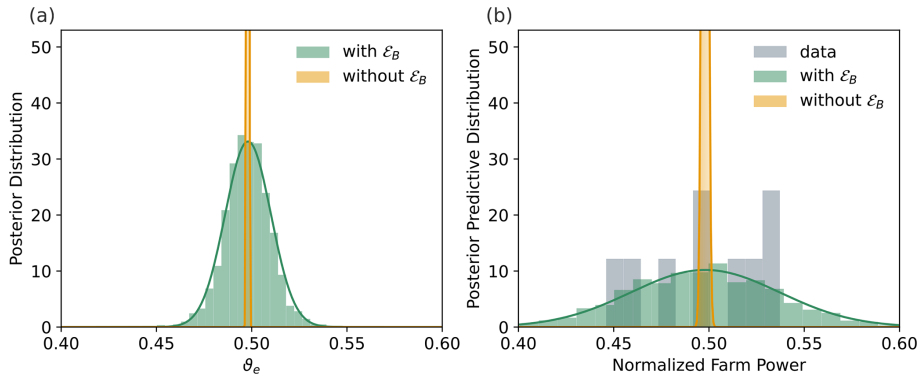

Figure 1. Comparison of the results of Bayesian UQ with and without the inclusion of model error $\mathcal{E}_B$: **(a)** posterior distribution and **(b)** posterior predictive distribution. The data and samples of the distributions obtained with UMBRA are given as histograms, and the analytically obtained probability density functions are given as full lines.

nomenclature stems from the analogy with the Boltzmann distribution $p(\varepsilon) \propto \exp(-\varepsilon/(k_B\Theta))$, which has high variance for high temperatures Θ and vice versa. The goal of the stage exponent is to gradually increase the influence of the likelihood by starting with an artificially large variance and then subsequently shrinking it. As a result, the algorithm can efficiently explore the prior range and successfully sample multimodal target distributions.

On the sequence of tempered target distributions, a particular version of the Resample–Move SMC algorithm is then used (Gilks and Berzuini, 2001; Doucet and Johansen, 2009), as depicted in Fig. 2. At every stage j, importance resampling is used to obtain N samples $\{\boldsymbol{\vartheta}_j^{(k)}\}_{k=1}^N$ that asymptotically follow the target distribution. To that end, importance weights are computed for every jth generation of particles:

$$w\left(\boldsymbol{\vartheta}_j^{(k)}\right) = \frac{p\left(\mathcal{D}|\boldsymbol{\vartheta}_j^{(k)}\right)^{\beta_{j+1}} p\left(\boldsymbol{\vartheta}_j^{(k)}\right)}{p\left(\mathcal{D}|\boldsymbol{\vartheta}_j^{(k)}\right)^{\beta_j} p\left(\boldsymbol{\vartheta}_j^{(k)}\right)} = p\left(\mathcal{D}|\boldsymbol{\vartheta}_j^{(k)}\right)^{\Delta\beta}, \tag{21}$$

where $\Delta\beta = \beta_{j+1} - \beta_j$. Then the particles $\{\boldsymbol{\vartheta}_j^{(k)}\}_{k=1}^N$ are resampled with a probability

$$\mathbb{P}\left(\boldsymbol{\Theta}_{j+1}^{(k)} = \boldsymbol{\vartheta}_j^{(k)}\right) = \frac{w\left(\boldsymbol{\vartheta}_j^{(k)}\right)}{\sum_{k=1}^N w\left(\boldsymbol{\vartheta}_j^{(k)}\right)} = \bar{w}\left(\boldsymbol{\vartheta}_j^{(k)}\right). \tag{22}$$

Due to the importance sampling steps, the algorithm works best for priors that sufficiently cover the high-likelihood region. Since we employ wide weakly informative priors, this is almost always the case.

Then N Metropolis–Hastings (MH) MCMC chains of length L are instantiated to perturb these samples again and remove the degeneracy introduced by the resampling step (Wu et al., 2018b). In the MH algorithm, the new sample $\boldsymbol{\vartheta}_{j,l+1}^{(k)}$ is sampled from a proposal density $q(\boldsymbol{\vartheta}_{j,l+1}^{(k)}|\boldsymbol{\vartheta}_{j,l}^{(k)})$ that only depends on the previous sample (Markov property). The new sample is accepted with a probability $\min(1, \alpha_{\mathrm{MH}})$,

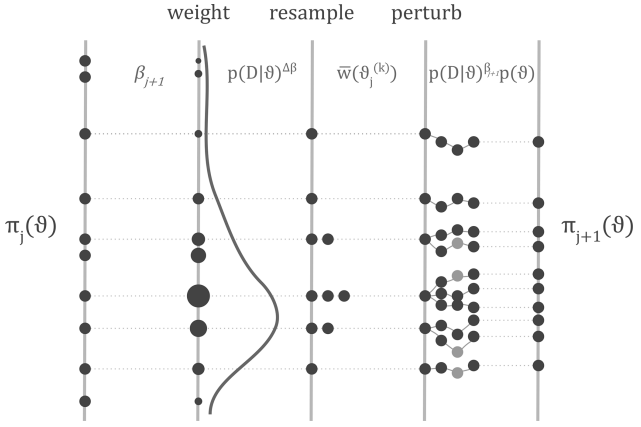

Figure 2. Schematic overview of one stage in the transitional Markov chain Monte Carlo (TMCMC) algorithm. From left to right, the samples are weighted, resampled, and perturbed. The algorithm can be parallelized in the perturbation phase. This figure is based on similar figures in the literature of sequential Monte Carlo and TMCMC (e.g., Doucet et al., 2001; Minson et al., 2013; Murphy, 2023).

where the acceptance ratio α_{MH} is defined as

$$\alpha_{\mathrm{MH}} = \frac{\pi_j(\boldsymbol{\vartheta}_{j,l+1}^{(k)})}{\pi_j(\boldsymbol{\vartheta}_{j,l}^{(k)})} \frac{q(\boldsymbol{\vartheta}_{j,l}^{(k)}|\boldsymbol{\vartheta}_{j,l+1}^{(k)})}{q(\boldsymbol{\vartheta}_{j,l+1}^{(k)}|\boldsymbol{\vartheta}_{j,l}^{(k)})}. \tag{23}$$

The Lth samples in the chains are taken as particle generation $j+1$. Since all MH chains can run simultaneously, the algorithm is inherently parallel. However, due to the MH steps, and thus similarly to (Adaptive) MH, TMCMC does not scale as well with parameter dimension as Hamiltonian Monte Carlo, which employs gradient information. If the prior does not sufficiently cover the high-likelihood region, longer MH chains are also required to compensate for the degeneracy introduced during importance sampling.

Several components of the algorithm can be tuned, such as the proposal distribution $q(\boldsymbol{\vartheta}_{j,l+1}^{(k)}|\boldsymbol{\vartheta}_{j,l}^{(k)})$ in MH, the method

to determine β_j, the number of samples per stage N, and the number of MCMC steps L. In our case, the distribution of the MH proposal is multivariate normal, that is, $q(\boldsymbol{\vartheta}_{j,l+1}^{(k)}|\boldsymbol{\vartheta}_{j,l}^{(k)}) = \mathcal{N}(\boldsymbol{\vartheta}_{j,l+1}^{(k)}; \boldsymbol{\vartheta}_{j,l}^{(k)}, \boldsymbol{\Sigma}_q)$, where $\boldsymbol{\Sigma}_q$ is the importance weighted sample covariance matrix of the previous stage scaled by $1/9 + 8R/9$, where R is the observed acceptance rate (Minson et al., 2013). The cooling rate should be fast enough to reduce the computational cost but slow enough to adequately represent the next target distribution after resampling. This adequacy can be quantified by the effective sample size (ESS) N^{eff} (Gelman et al., 2013). Its interpretation is that N weighted samples $\{w_j^{(k)}, \boldsymbol{\vartheta}_{j,L}^{(k)}\}_{k=1}^{N}$ are worth N^{eff} i.i.d. samples drawn from the target distribution π_{j+1} (Beck and Zuev, 2013). The ESS is estimated from the normalized importance weights as

$$\hat{N}_{j+1}^{\text{eff}} = \frac{1}{\sum_k \bar{w}\left(\boldsymbol{\vartheta}_{j,L}^{(k)}\right)^2} = \frac{\left(\sum_k p\left(\mathcal{D}|\boldsymbol{\vartheta}_{j,L}^{(k)}\right)^{\Delta\beta}\right)^2}{\sum_k p\left(\mathcal{D}|\boldsymbol{\vartheta}_{j,L}^{(k)}\right)^{2\Delta\beta}}. \tag{24}$$

The optimal cooling rate is obtained when β_{j+1} is chosen such that the ESS is approximately half the total number of samples per stage (Minson et al., 2013). To that end, Eq. (24) is solved for β_{j+1} with a bisection method on $(\beta_j, 1]$ (Kantas et al., 2014). Based on the study of the sensitivity of the posterior to the choice of N and L with increasing parameter dimension by Minson et al. (2013), we take $N = 1920$ and $L = 20$ for a total of (maximum) 10 parameters. For 15 stages, this requires 5.76×10^5 likelihood evaluations, of which only 300 are inherently serial. For further details on the implementation, the reader is referred to the Python toolbox UMBRA, which is released together with this paper.

3 Case setup

The Bayesian UQ framework is demonstrated with a reference large-eddy simulation (LES) dataset for wind farm blockage and atmospheric gravity waves in conventionally neutral boundary layers (CNBLs). Since it is still challenging to model these effects with state-of-the-art models, the dataset provides a controlled setting for evaluating the framework under conditions of substantial model bias and uncertainty. The wind farm flow models are introduced in Sect. 3.1, and the dataset is described in Sect. 3.2. Based on the empirical model parameters and dataset characteristics, the prior distribution is further specified in Sect. 3.3.

3.1 Wind farm flow models

The two wind farm flow models we will consider in this study are a standard wake model (Sect. 3.1.1) and an atmospheric perturbation model (APM) (Sect. 3.1.2). The wake model cannot capture wind farm blockage and is expected to exhibit large model uncertainty for the considered dataset. The second model consists of a wake model with an APM-based blockage correction. Since the blockage effect in this dataset is mainly caused by atmospheric gravity waves, we opt for the APM as a blockage correction model, but different models exist (Branlard et al., 2020).

3.1.1 Standard wake model

The primary objective of wake models is to predict the effect of the velocity deficit downstream of a turbine on the other turbines within a farm. The most well-known and widely used wake model is presumably the one originally proposed by Jensen (1983), but over the years many others have followed (Göçmen et al., 2016). In this study, the Gaussian wake model of Bastankhah and Porté-Agel (2014) will be employed. It is based on the typical self-similar Gaussian profile of the velocity deficit and mass and momentum conservation in the wake (see also Frandsen et al., 2006).

The velocity deficit of a turbine i located upstream is then given by

$$\frac{U_\infty - U_{\text{w}}(\boldsymbol{x})}{U_\infty} = \left(1 - \sqrt{1 - \frac{C_{T,i}}{8\sigma_i^{*2}}}\right) \times \exp\left(-\frac{1}{2\sigma_i^{*2}}\frac{y^2+z^2}{D_i^2}\right)\mathcal{H}(x), \tag{25}$$

where U_∞ is the free stream speed, U_{w} is the speed in the wake, and $\boldsymbol{x} = [x, y, z]^\top$ is defined in a local coordinate system at the turbine hub, with z the vertical direction, z and y spanning the rotor plane, and $x > 0$ downstream of the turbine. Furthermore, $\mathcal{H}$ is the Heaviside function, $C_{T,i}$ the turbine thrust coefficient, D_i is the rotor diameter, and $\sigma_i^* = \sigma_i / D_i$ is the normalized wake width which grows linearly with the distance downstream from the rotor as

$$\sigma_i^* = k_{\text{w}}\frac{x}{D_i} + \epsilon. \tag{26}$$

Here, k_{w} is the wake expansion rate, which is related to the local turbulence intensity at turbine with two empirical parameters as $k_{\text{w}} = k_a I + k_b$, where a previous fit yielded $k_a = 0.3837, k_b = 0.003678$ (Niayifar and Porté-Agel, 2016). Finally, ϵ is a semi-empirical parameter that represents the initial wake width, which depends on the turbine thrust coefficient (Bastankhah and Porté-Agel, 2014). The turbulence intensity at the turbine is determined with the method of Niayifar and Porté-Agel (2016) based on the original (Zehtabiyan-Rezaie and Abkar, 2023) expression for the added turbulence intensity expression by Crespo and Hernández (1996):

$$I_+ = 0.73a^{0.8325} I_\infty^{-0.0325} (x/D)^{-0.32}, \tag{27}$$

where I_∞ is the background turbulence intensity.

In order to combine multiple wakes into one flow field a wake-merging method is needed. We will use the one developed by Lanzilao and Meyers (2022), which uses a self-similarity argument to combine the wake deficits through

multiplication, without introducing any empirical parameters. Additionally, the turbines are mirrored to capture the effect of the ground plane (Lissaman, 1979). The power extracted by turbine i can then be computed as $P_i = \frac{1}{2}\rho A_i U_i^3 C_{P,i}(U_i)$, where $C_{P,i}(U_i)$ is the power coefficient of that turbine, A_i is the swept rotor area, and U_i is the disk-averaged flow speed, calculated with the same quadrature method as in Allaerts and Meyers (2019).

3.1.2 Atmospheric perturbation model

Atmospheric perturbation models aim to model wind farm–atmosphere interaction effects such as wind farm blockage in addition to the turbine-scale interactions due to wakes. They do so by solving the height-averaged and linearized Reynolds-averaged Navier–Stokes (RANS) equations for the atmospheric boundary layer (ABL) under the Boussinesq approximation. The linearization involves adding a perturbation velocity $\boldsymbol{u} = [u, v, w]^\top$ to the velocity $\boldsymbol{U} = [U, V, W]^\top$ in the ABL. The APM further divides the ABL in a wind farm layer of height H_1 and a second layer of height $H_2 = H - H_1$ where H is the ABL height. The equations and their solution procedures are derived and described in detail by Allaerts and Meyers (2019), Stipa et al. (2024a), and Devesse et al. (2024a). The three most important terms for our purposes are the farm thrust, the added turbulent momentum flux associated with the development of an internal boundary layer (IBL), and the pressure feedback induced by the upward displacement of the capping inversion layer – the interface between the neutral atmospheric boundary layer and the stably stratified free atmosphere aloft in CNBLs. Including this pressure feedback allows the APM to simulate the interaction between the ABL flow and gravity waves explicitly, thereby modeling blockage effects without introducing tuning parameters. The farm thrust and turbulent momentum flux, on the other hand, do contain parameters that will be calibrated with the Bayesian framework. Figure 3 represents these effects schematically.

The wind farm thrust $\boldsymbol{f}(x, y)$ is represented in the APM by filtering the turbine thrust forces $\boldsymbol{f}_i$ located at the turbine positions (Allaerts and Meyers, 2019)

$$\boldsymbol{f}(x,y) = \int_0^{L_x}\int_0^{L_y} G(x-x', y-y') \times \sum_{i=1}^{N_\mathrm{T}} \boldsymbol{f}_i \delta(x'-x_i, y'-y_i)\mathrm{d}x'\mathrm{d}y', \tag{28}$$

for an APM domain of $[0, L_x] \times [0, L_y]$. Here, $\delta(x, y)$ is the two-dimensional Dirac delta and $G(x, y)$ is a Gaussian filter kernel with a filter length scale of L_f

$$G(x,y) = \frac{1}{\pi L_\mathrm{f}^2} \exp\left(-\frac{x^2+y^2}{L_\mathrm{f}^2}\right). \tag{29}$$

The filter length scale is set to 1000 m by Allaerts and Meyers (2019) and Devesse et al. (2024a) and to 500 m by Stipa et al. (2024a), and it can be considered an uncertain parameter. This filtering operation must be applied to the momentum equations as a whole, resulting in dispersive stresses (Devesse et al., 2024a). However, computing the resulting dispersive stresses with the current parametrization is the most expensive part of an APM evaluation (Devesse et al., 2024b). Since these stresses are a minor perturbing force compared to the turbine thrust, we will ignore them to reduce the computational cost. Faster parametrizations for the dispersive stresses in the APM are a topic of ongoing research.

The development of an internal boundary layer is accompanied by an increase in the momentum flux from the layer above the wind farm to the wind farm layer. This added turbulent momentum flux is represented as (Devesse et al., 2024a)

$$\Delta\tau_{\mathrm{WF}}(x,y) = a_\tau C_\mathrm{F} \Pi(\mathbf{x} - d_\tau D \boldsymbol{e}_\mathrm{s}), \tag{30}$$

where a_τ is the proportionality constant to the wind farm force density $C_\mathrm{F} = \frac{1}{2} C_\mathrm{T} N_\mathrm{T} A ||\boldsymbol{U}_1||^2 / S_\mathrm{F}$ and $\boldsymbol{e}_\mathrm{s}$ is the streamwise unit vector. Here, C_T is the average turbine thrust coefficient, N_T is the number of turbines, A is the swept rotor area, $\boldsymbol{U}_1$ is the unperturbed velocity in the farm layer, and S_F is the wind farm surface. The added momentum flux is oriented along the wind farm forcing and is zero everywhere except on the wind farm footprint $\Pi(\boldsymbol{x})$. For a rectangular farm, this is a block function. To include the development of the IBL, the footprint is shifted d_τ turbine diameters downstream given that the turbines that are on average aligned with the wind. That leaves two empirical parameters a_τ and d_τ, which were previously fitted to 0.12 and 27.8 based on the computed added momentum flux from LES data (Devesse et al., 2024a).

The farm thrust will slow down the flow in the ABL. The resulting decrease in the streamwise velocity u is balanced in the continuity equation by induced spanwise flow v and thickening of the ABL. This thickening corresponds to a lifting of the capping inversion η, which leads to two distinct processes that result in the pressure feedback $p_t = p_i + p_\mathrm{fa}$ (cf. Fig. 3c). First, the lifting of the capping inversion directly corresponds to a cold anomaly, as the air below it is colder than the air above. These pressure perturbations p_i can travel horizontally along the capping inversion as two-dimensional interfacial gravity waves. Second, the changes in capping inversion height perturb the free atmosphere aloft, leading to internal gravity waves. These three-dimensional waves also lead to pressure perturbations, which are felt throughout the ABL (Smith, 2010). Combined, these two types of gravity waves cause a pressure increase upstream, leading to the blockage effect. Downstream, they also induce a favorable pressure gradient throughout the farm (cf. Fig. 3c).

The wake effects are included with a standard wake model, which can be coupled to the height-averaged RANS equations for the two layers together with the pressure feedback

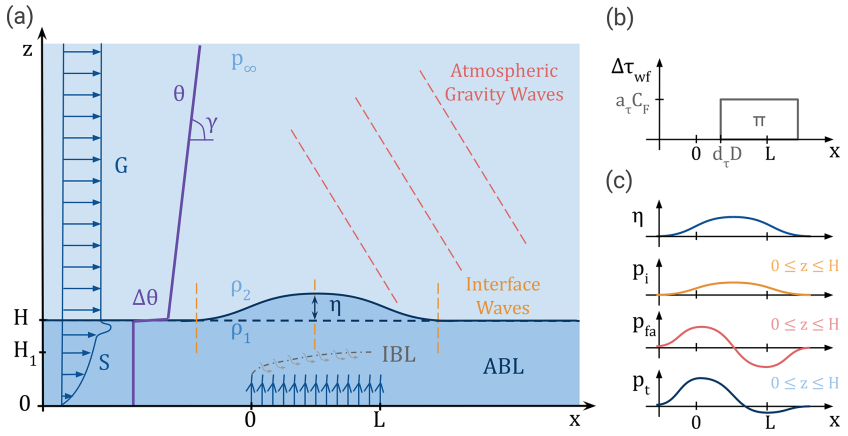

Figure 3. Schematic representation wind farm–atmosphere interaction via atmospheric gravity waves and internal boundary layer (IBL) development as modeled by the atmospheric perturbation model: **(a)** a sketch of the wind farm, vertical profiles of wind speed (blue) and potential temperature (purple), wave crests and troughs for internal gravity waves in the free atmosphere (red) and interface waves on the inversion layer (orange), and a developing IBL (gray); **(b)** the added momentum flux due to the presence of the wind farm and related to IBL growth as modeled in the APM; **(c)** a hypothetical displacement η of the inversion layer with the associated pressure feedback p_t in the atmospheric boundary layer (ABL), split up in the pressure components related to the waves on the inversion layer p_i, and the waves in the free atmosphere p_{fa}.

from the upper atmosphere (Devesse et al., 2024b). The predicted flow redirection is included with a simplified version of the bidirectional wake-merging method of Lanzilao and Meyers (2022), which is elaborated upon in Appendix A. In the current work, the wake model is coupled via the pressure (Stipa et al., 2024a), but an upstream coupling (Allaerts and Meyers, 2019) and a velocity matching approach exist as well (Devesse et al., 2024a). Figure 4 shows the obtained flow field, where the wake model provides the information on the wakes and the APM provides the background velocity and pressure information.

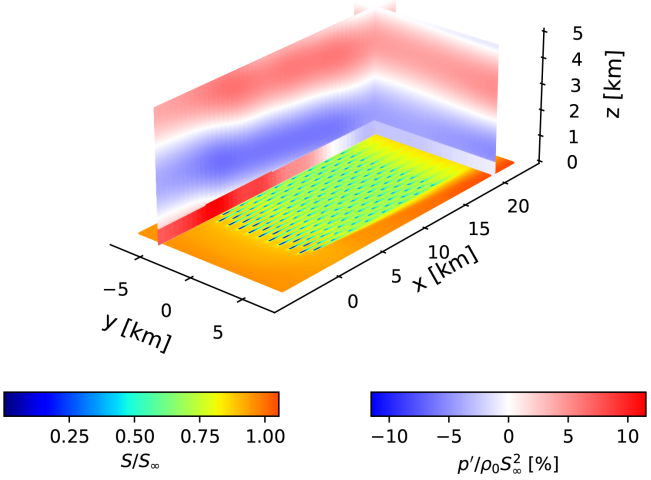

Figure 4. Wind speed and pressure field obtained with an atmospheric perturbation coupled via the pressure to a Gaussian wake model with a bidirectional wake-merging method. The staggered wind farm consists of 16 rows of 10 wind turbines with streamwise and spanwise spacings of 5 rotor diameters. The atmospheric boundary layer height amounts to 500 m with a capping inversion strength of 4 K and a free-atmosphere lapse rate of 8 K km^{-1}. The friction velocity equals 0.275 m s^{-1}. The upstream wind speed and ambient turbulence intensity at hub height are 9.24 m s^{-1} and 3.93 %.

3.2 Reference dataset

As reference data, the parametric LES study of wind farm blockage and gravity waves in CNBLs of Lanzilao and Meyers (2024) is used. They simulated 36 selected atmospheric states based on 30 years of ERA5 re-analysis data at the nearest grid point to the Belgian–Dutch offshore wind farm cluster. Figure 5 depicts the inflow conditions (the time average of the last 4 h of the precursor simulations) and the power output per turbine, normalized by the power of a hypothetical undisturbed turbine upstream, for the cases with an ABL height of 500 m. From Fig. 5b it can be seen that the ABL is always neutrally stratified, since the potential temperature θ remains constant $\mathrm{d}\theta/\mathrm{d}z = 0$. The capping inversion strength $\Delta\theta$, and the lapse rate in the free atmosphere aloft $\mathrm{d}\theta/\mathrm{d}z > 0$ are varied.

The wind farm consists of 16 rows and 10 columns ($N_{\mathrm{T}} = 160$) of 10 MW IEA reference turbines (Bortolotti et al., 2019). The streamwise and spanwise spacings are $5D$, with $D = 198$ m as the turbine diameter. The turbines have hub heights of $z_h = 119$ m, and in the study the thrust coefficient is fixed to $C_{\mathrm{T}} = 0.88$. The rows are counted in the streamwise direction and labeled with capital letters in Fig. 5b. It can be seen in Fig. 5b how the blockage effect causes large reductions in turbine power in the first rows. At the same time,

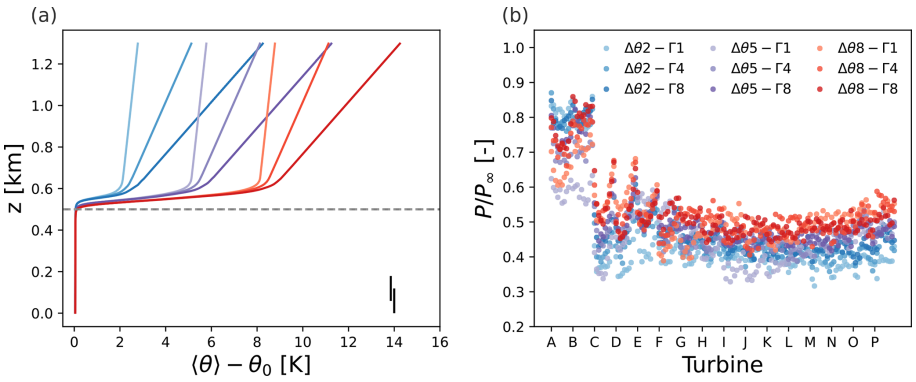

Figure 5. Large-eddy simulation data of Lanzilao and Meyers (2024): **(a)** potential temperature profile with $\theta_0 = 288.15$ K, averaged horizontally in space and over the last 4 h of the precursor simulation, and **(b)** power output per turbine normalized by the power of an "undisturbed" upstream turbine. In panel **(a)**, the ABL height is indicated with a dashed gray line and a turbine is drawn with full black lines as a reference.

the favorable pressure gradient improves power recovery in the last rows. The significant local flow redirection related to blockage misaligns the turbine wakes with the downstream turbines on the sides of the farm, resulting in a U-shaped trend in the power per row. In general, the resulting turbine power output varies significantly with atmospheric stratification.

Of the 36 available cases, we will only consider the 9 cases with an ABL height of 500 m in this study, to isolate the effect of stratification of the upper atmosphere from the height of the boundary layer. As we intend to demonstrate the framework's capability to quantify the uncertainty adequately for models of different fidelity, the APM is included as a model. However, an APM evaluation takes about 30 s, so 1000 model evaluations already require 8.3 core hours, compared to 0.83 core minutes for 1000 wake–model evaluations of 0.05 s. Although the SMC algorithm allows the many required APM evaluations to be performed in parallel to achieve a speed-up in time, the total computational cost remains the same.

3.3 Prior choices depending on the dataset and models

An overview of the marginal priors for the empirical model parameters $\boldsymbol{\vartheta}_e$ of the wake model and the APM, as well as the parameters $\boldsymbol{\vartheta}_b = \{\delta_\zeta, \sigma_{B,\zeta}\}_{\zeta=0}^{\zeta_{\max}}$ that characterize the model error distribution, is given in Table 1. The measurement error covariance can be calculated from the 90 min long turbulent power signals using the moving block bootstrap (Garcia et al., 2005). Since the upstream velocity and potential temperature profiles are available from the LES, the inflow uncertainty is considered negligible. The prior of the mean bias terms depends on the considered wind farm flow model in the adapted Bayesian framework. For the wake model, we know a priori that the sensitivity of the predicted power to the wake expansion rate is only zero for the upstream turbines, which are not waked. Therefore, only the bias in the upstream and undisturbed rows of turbines δ_0 can be nonzero based on the condition against confounding of the model bias with calibration parameters in Eq. (10). However, we find that also requiring $\delta_0 = 0$ works best in practice based on a comparison of alternatives in Sect. 4.1. For the APM, the condition for the best-fit interpretation of the model parameters in Eq. (10) directly requires that all mean bias terms $\delta_{\zeta(i)}$ are zero.

Table 1. Prior distributions for the inverse uncertainty quantification of the wake model and the atmospheric perturbation model. Uniform distributions on an interval $[a, b]$ are abbreviated as $\mathcal{U}_a^b$. The exponential distribution Exp_λ has a PDF $p(x) = \lambda^{-1}\exp(-x/\lambda)$ for $x > 0$. The Dirac delta distribution $\mathrm{Delta}(a)$ has a PDF $p(x) = \delta(x-a)$. The filter length L_f and height of the first layer H_1 are expressed in meters, and the spatial delay of the turbulent entrainment d_τ is expressed in turbine rotor diameters.

$\boldsymbol{\vartheta}_e$						$\boldsymbol{\vartheta}_b$	
k_a	k_b	L_f	H_1	a_τ	d_τ	$\delta_{\zeta(i)}$	$\sigma_{B,\zeta(i)}$
$\mathcal{U}_0^1$	$\mathcal{U}_0^{0.1}$	$\mathcal{U}_{500}^{2000}$	$\mathcal{U}_{220}^{450}$	$\mathcal{U}_0^1$	$\mathcal{U}_0^{80}$	Delta(0)	$\mathrm{Exp}_{0.1}$

The model parameters are given by $\boldsymbol{\vartheta}_e = (k_a, k_b)$ for the wake model and by $\boldsymbol{\vartheta}_e = (k_a, k_b, L_f, H_1, a_\tau, d_\tau)$ for the APM. Similarly to previous work (Aerts et al., 2023), the wake expansion rate parameter k_a gets a uniform prior over the unit interval, whereas k_b gets a stronger uniform prior between 0–0.1. These marginal priors for k_a and k_b ensure that the wake expansion rate is positive for all turbulence intensities. The filter length scale L_f is taken not smaller than the grid spacing of 500 m and not larger than 2 times the current value in Devesse et al. (2024a). The allowable farm layer height H_1 is chosen slightly larger than the turbine tip height of 218 m and smaller than 90 % of the ABL height of 500 m. The strength a_τ of the turbulent entrainment should be positive and is not expected to be more than 10 times larger than its current estimated value of ≈ 0.1 (Devesse et al., 2024a). The spatial delay of the turbulent en-

trainment d_τ should clearly be positive and cannot exceed the farm length of 80 turbine diameters.

4 Results and discussion

In Sect. 4.1, the adequacy of the Bayesian framework is verified for the wake model, which is expected to show large model error and uncertainty for the blockage dataset, and the APM, which is expected to perform better. The posterior distributions for both models are compared in Sect. 4.2. Lastly, the generalization of the results and the intended use of the framework are discussed in Sect. 4.3. The posterior distributions are sampled on the wICE supercomputing platform of the VSC (Vlaams Supercomputer Centrum), using Sapphire Rapids nodes containing 2 Intel Xeon Platinum 8468 CPUs (48 cores each).

4.1 Adequacy of the Bayesian framework

We now turn to the analysis of the wake model and the APM with turbine power data. Since the posterior distribution is intractable for those cases, we use the parallelized SMC algorithm implemented in UMBRA to sample it (Sect. 2.4). The SMC algorithm yields 1920 samples of the joint posterior distribution of $\boldsymbol{\vartheta}_e$ and $\boldsymbol{\vartheta}_b$ after convergence. In addition, we generate a sample of the posterior predictive distribution for each sample of the posterior so that the adequacy of the Bayesian framework can be assessed by comparing the posterior predictive samples with the observations (Sect. 2.2.2).

Figure 6 compares the means and standard deviations of the posterior predictive samples of each turbine with the distribution of the LES reference data for both the wake model and the APM. To isolate model uncertainty from (epistemic) parameter uncertainty and measurement uncertainty, we also show the posterior predictive obtained from the same posterior of the model parameters, but with $\{\sigma_{B,\zeta(i)}\}_{i=1}^{N_T}$ set to zero. Figure 6a shows that the wake model exhibits substantial model uncertainty for this dataset. In the first turbine rows, the model uncertainty is inflated because of the significant bias due to wind farm blockage. Further downstream, the large model uncertainty stems form the U-shaped power variations caused by flow redirection and the enhanced power recovery due to the favorable pressure gradient, which the wake model fails to capture. Notably, the data appear implausible under the posterior predictive without model error, underscoring the importance of properly accounting for model uncertainty. Figure 6b shows that the APM yields considerably lower model uncertainty, as it successfully incorporates blockage, flow redirection, and pressure gradient effects. However, consistent with previous findings (Devesse et al., 2024b), the APM underestimates blockage effects in the first rows, which inflates model uncertainty for those turbines. Since the reference data are deemed plausible under the posterior predictive distributions of both models, the Bayesian framework proves to be adequate.

To avoid overestimating model uncertainty caused by systematic model mismatch, one could also estimate the mean model bias. However, this bias must be constrained to ensure model parameter identifiability, either by enforcing unbiased predicted farm power in Eq. (9) or by trivially satisfying the condition against confounding of the model bias with model parameters in Eq. (10). One approach is to relax the condition for the "best-fit" parameter interpretation in Eq. (10), as illustrated in Fig. 7a for the wake model. This allows the mean bias δ_0 to be identified in the first turbine rows due to blockage, resulting in lower estimated model uncertainty. However, to compensate for this bias and still match farm power, the calibrated wake model systematically underestimates turbine power downstream. Alternatively, we can relax the constraint that predicted farm power must be unbiased, as shown in Fig. 7b for the wake model. This yields the exact same calibration (and posterior of the model parameters) as when the mean bias for the upstream turbines is not estimated, since upstream power contains no information about the model parameters. Nevertheless, as Fig. 6b shows, even flow models that have parameters that influence the predicted blockage effect can exhibit bias, leading to inflated uncertainty estimates for these models. Therefore, we recommend excluding mean model bias to enable the objective comparison of model uncertainty between different flow models, even though the estimated model uncertainty is conservative.

4.2 Model comparison with quantified model and parameter uncertainty

The Bayesian UQ framework can be used to perform objective model comparison with quantified parameter and model uncertainty given a reference dataset. In Sect. 4.2.1, the posterior distributions of the model parameters are presented for the wake model and the atmospheric perturbation model. The posterior distribution of the parameters describing the model error distribution are compared for both models in Sect. 4.2.2.

4.2.1 Posterior distribution of the model parameters

Figure 8 shows one- and two-dimensional histograms of the joint posterior distribution of the wake model parameters, based on samples generated with SMC in UMBRA and visualized with Corner (Foreman-Mackey, 2016). A comparison between the marginal posterior and prior distributions reveals that the parameters are well identified. Notably, the posterior median of k_a (0.52) exceeds its standard reference value of 0.384 (Niayifar and Porté-Agel, 2016), while k_b is estimated to be nearly zero. In rows A and B, where turbines are not affected by wakes, the mean posterior wake expansion rate $k_w = k_a I + k_b$ is approximately 0.018, given an ambient turbulence intensity of 0.039, compared to 0.019 based on the same turbulence intensity I and literature values for k_a and k_b. In the downstream rows, the mean posterior wake expan-

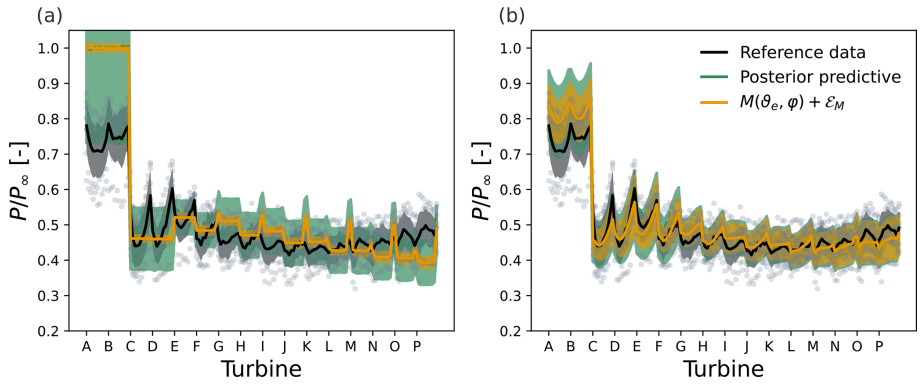

Figure 6. Comparison of the turbine power from LES with the posterior predictive distributions for **(a)** the wake model and **(b)** the atmospheric perturbation model. The mean power is shown as a solid line, with shaded regions indicating 1 standard deviation above and below the mean. The mean and standard deviation of the model outputs $\boldsymbol{M}(\boldsymbol{\vartheta}_\mathrm{e}, \varphi)$ for the posterior samples of $\boldsymbol{\vartheta}_\mathrm{e}$ are shown as well after adding the measurement error $\boldsymbol{\mathcal{E}}_\mathrm{M}$. The reference data points are also plotted as individual dots.

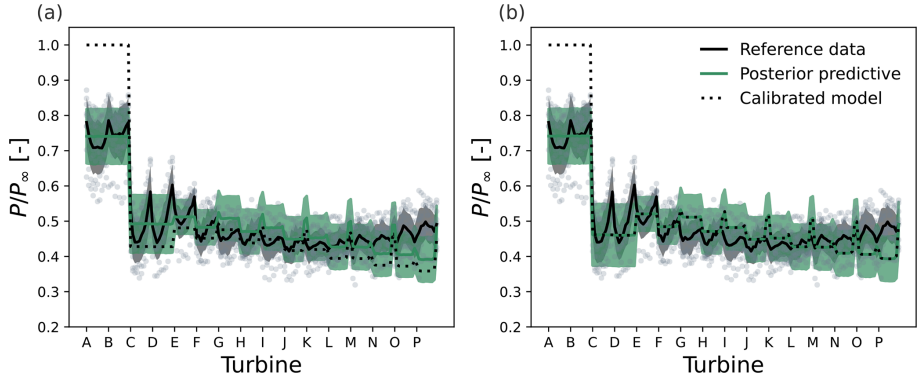

Figure 7. Comparison of the turbine power data from LES with the posterior predictive distribution for the wake model when the mean bias terms $\mu_{B,i}$ are estimated together with the model parameters. In panel **(a)**, all mean bias terms are estimated with the constraint that their sum is zero, whereas in panel **(b)** only the bias on the upstream turbines is estimated. The mean power is shown as a solid line, with shaded regions indicating 1 standard deviation above and below the mean. The reference data points are also plotted as individual dots.

sion rate increases to around 0.073 due to wake-added turbulence, compared to a literature-based value of 0.064 using the same local turbulence intensity. This suggests that wake recovery in waked turbine rows is overestimated to compensate for the favorable pressure gradient influencing the background flow field. Although the estimated wake expansion rate in the upstream rows is consistent with earlier results, the local wind speed is largely overestimated by the wake model, which makes the comparison invalid.

Figure 9 shows the samples of the joint posterior distribution of the APM parameters, which are all well identified. Compared to standard values, the posterior median of k_a is lower (0.32) and that of k_b is higher (0.0125), resulting in a larger mean wake expansion rate in rows A–B ($k_\mathrm{w} \approx 0.027$) and a similar rate in the downstream rows ($k_\mathrm{w} \approx 0.064$), given the same turbulence intensities. The increased upstream wake expansion rate suggests the need for further investigation into turbine wakes under blockage conditions (see Ndindayino et al., 2025). The estimated filter length scale L_f is smaller than the value used in Devesse et al. (2024a) but slightly larger than in Stipa et al. (2024a). The relatively high uncertainty in L_f is attributed to the limited sensitivity of the turbine power to changes in filter width between 500 m and 640 m. The height of the first layer H_1 is estimated to be close to twice the turbine hub height, consistent with findings from a previous parameter study (Allaerts and Meyers, 2019). The estimated strength of the wind farm added momentum flux a_τ reaches the upper bound of its prior, roughly 10 times its current value, and its spatial delay d_τ is larger than previously fitted. This discrepancy across the model chain indicates that the parametrization of the added momentum flux requires further refinement. In the pressure-based coupling of the wake model to the height-averaged RANS equations, noticeable increases in turbine power only occur for a_τ values near 1. Thus, the strong added momentum flux downstream helps to replicate the observed power increase in rows N–P (cf. Fig. 6b).

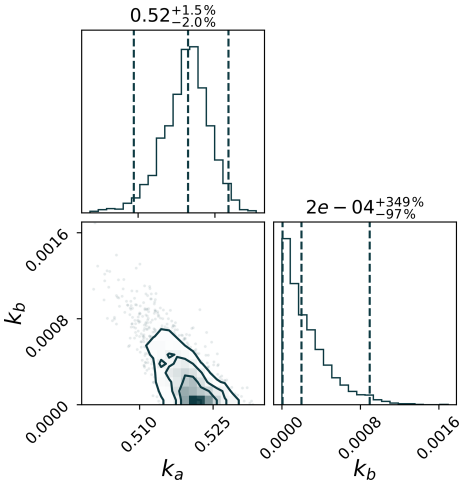

Figure 8. Joint posterior probability density of the parameters in the wake expansion rate $k_w = k_a I + k_b$. For each parameter, the median is given together with the 2.5 % and 97.5 % quantiles, expressed as relative deviations from the median.

4.2.2 Comparison of the quantified model uncertainty

Table 2 summarizes the marginal posteriors of the model error standard deviations for both wind farm flow models. It is clear that the rather large uncertainties $\sigma_{B,\zeta(i)}$ for the wake model are caused by the unobserved or unmodeled variations in atmospheric stratification in this dataset. In general, the model uncertainty is smaller for the APM, as the standard deviations of the model error $\sigma_{B,\zeta(i)}$ are smaller. Because the APM captures the blockage effect, the model uncertainty on upstream turbines is reduced by a factor of 5.5 for this dataset. In addition, variations in upstream blockage are better estimated, as the uncertainty of the turbines with one waking turbine is lowered by a factor of 1.5. Only the model uncertainty for turbines with two upstream waking turbines is larger. In fact, it is seen in Fig. 6 that the deviations from the predicted power are larger in rows E and F. In contrast, the model uncertainty $\sigma_{B,3}$ for turbines farther downstream is a factor of 2.2 lower than for the wake model. This is because the APM adequately models the increase in power in later rows due to the inclusion of the favorable pressure gradient and the wind farm added momentum flux. Note that the posterior uncertainty on $\sigma_{B,3}$ is smaller than on the other model error standard deviations. This is because $\sigma_{B,3}$ is associated with 100 turbines, while the others are associated with 20 only. Thus, its epistemic uncertainty is further reduced by a factor $\sqrt{5}$.

4.3 Generalization of the obtained results

A natural question that arises when calibrating models is to what extent the resulting performance generalizes to other datasets. Since physics-based wind farm flow models are expected to generalize well to other farm layouts, wind speeds, and wind directions, we expect that the performance should generalize well when varying those conditions. In that case the posterior distribution of the empirical parameters $p(\boldsymbol{\vartheta}_e|\mathcal{D})$ can be used together with the quantified model error distribution to obtain a stochastic wind farm flow model. However, for unobserved or unmodeled conditions $\boldsymbol{\psi}'$, we cannot expect a proper generalization, since the bias of the flow model can only (partially) be reduced by altering the empirical model parameters $\boldsymbol{\vartheta}_e$. Consequently, the results of the uncertainty quantification depend largely on the resemblance between the distribution of these unobserved conditions in the considered data $p(\boldsymbol{\psi}')$ and their true distribution $p_{\text{true}}(\boldsymbol{\psi}')$. In this case, we have by no means covered the true distribution $p_{\text{true}}(\boldsymbol{\psi}')$ – even if we intended to consider only atmospheric stratification effects in CNBLs with an ABL height of 500 m. Similarly, the posteriors of the model parameters likely depend on the ABL height. As such, the obtained results are specific to the dataset considered and are intended only as an illustration of the methodology. However, we showed that the methodology does allow objective model comparison with quantified model and parameter uncertainty given a benchmark dataset.

Table 2. Summary statistics of the marginal posterior distributions of the model error standard deviation $\sigma_{B,\zeta(i)}$ for the wake model (WM) and the atmospheric perturbation model (APM). For each parameter, the median is given together with the 2.5 % and 97.5 % quantiles, expressed as relative deviations from the median.

[%]	$\sigma_{B,0}$	$\sigma_{B,1}$	$\sigma_{B,2}$	$\sigma_{B,3}$
WM	$27^{+11\,\%}_{-10\,\%}$	$8.6^{+12\,\%}_{-10\,\%}$	$4.6^{+11\,\%}_{-11\,\%}$	$6.1^{+5\,\%}_{-5\,\%}$
APM	$4.9^{+13\,\%}_{-12\,\%}$	$5.1^{+13\,\%}_{-11\,\%}$	$6.1^{+12\,\%}_{-11\,\%}$	$2.8^{+6\,\%}_{-6\,\%}$

5 Summary and outlook

Bayesian UQ leverages data to quantify the uncertainty of model parameters and the model itself by updating a prior distribution that characterizes the available knowledge, before having seen the data, to a posterior distribution. In this article, we examined the use of Bayesian UQ for (1) obtaining stochastic wind farm flow models through Bayesian calibration and (2) objective model comparison with quantified parameter and model uncertainty. As both applications require that model inadequacy is properly taken into account, the model inadequacy formulation in a previously developed Bayesian UQ framework (Aerts et al., 2023) was improved. The framework was demonstrated with engineering models for wind farm–atmosphere interaction using a large-eddy simulation dataset for wind farm blockage due to atmospheric gravity waves. In doing so, the framework was tested for delineated data with a large anticipated model uncertainty, as current engineering models face difficulties in representing those effects. In contrast to earlier studies,

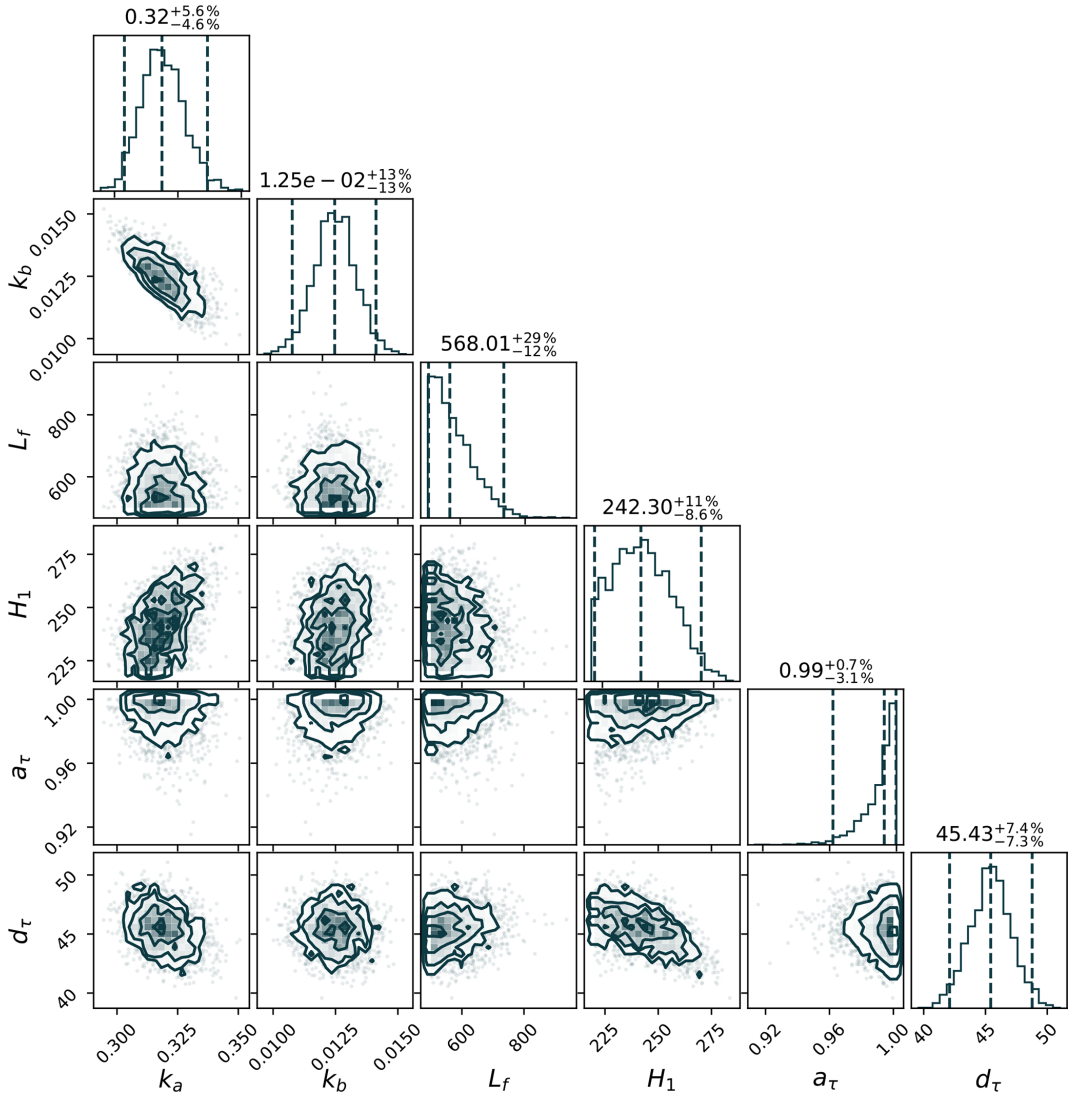

Figure 9. Joint posterior probability density of the parameters in the atmospheric perturbation model. For each parameter, the median is given together with the 2.5 % and 97.5 % quantiles, expressed as relative deviations from the median.

we used a parallelized sequential Monte Carlo algorithm based on likelihood tempering to speed up the approximation of the posterior, though at a similar computational cost. This complete workflow is made available in a Python toolbox coined Uncertainty Modeling toolbox for Bayesian data Re-Analysis (UMBRA), which can be used together with WAYVE (Devesse et al., 2023) and WIFA (Quick et al., 2024).

With a simple example model for wind farm power, the consequences of not properly including model error in Bayesian UQ are illustrated. On the one hand, the posterior distribution of the model parameters is overconfident and biased when the model error is neglected. In that case, a posterior predictive check also shows that the Bayesian analysis is inadequate. Hence, the proper inclusion of model error is also important when one is only interested in the posterior distribution of the flow model parameters. On the other hand, current stochastic flow models, which only propagate the posterior distribution of the model parameters through the model (Zhang and Zhao, 2020), may significantly underestimate the variability of the true process that the model aims to represent, as soon as there is non-negligible model error. Moreover, in the limit of infinite data, such "stochastic" models will, in fact, become deterministic. The presented framework properly includes the model error so that the posterior distribution of the model parameters and the model uncertainty are adequately quantified, also in the limit of infinite data. Since most, if not all, of the current wind farm flow models have non-negligible model error, Bayesian UQ analyses that neglect model error will suffer similar issues.

The adequacy of incorporating model error on turbine power predictions within the Bayesian framework was also

assessed. A posterior predictive check using the blockage dataset revealed that the framework is adequate both for a standard wake model, which does not capture wind farm blockage effects and has large model error, and for an atmospheric perturbation model (APM), which does capture those effects. By requiring that the calibrated model is unbiased on the farm power (Aerts et al., 2023) and that the calibration parameters are to be interpreted as "best-fit" parameters (Plumlee, 2017), the mean bias on each turbine is a priori considered to be zero. By doing so, the current approach does not suffer from the confounding of calibration parameters with model inadequacy (Brynjarsdóttir and O'Hagan, 2014). Although the quantified model uncertainty is more conservative as a result, the model uncertainty can be objectively compared for different wind farm flow models.

The Bayesian UQ of the wake model and APM showed that the framework can be used for objective model comparison with quantified parameter and model uncertainty given a reference dataset. The posterior distribution of the model parameters is significantly updated with respect to the prior distribution, indicating that the parameters are well identified. The uncertainty of the parameters characterizes the remaining uncertainty due to the limited number of data, and it is seen that the relative uncertainties of the model parameters are inversely related to their sensitivity. Posterior correlations and inconsistencies between the posterior modes throughout the model chain may inform modelers of the parts that need to be further improved. For the APM, the estimated wake expansion rate in the upstream turbine rows is higher than the rate derived from standard parameter values (Niayifar and Porté-Agel, 2016), while in the downstream rows, they align closely. This encourages further research on turbine wakes under blockage conditions (e.g., Ndindayino et al., 2025). A comparison of the quantified model uncertainties shows that the APM exhibits substantially lower uncertainty than the wake model. This applies both to the upstream turbines, which are subject to significant blockage effects, and to the downstream turbines, which benefit from the favorable pressure gradient across the farm in the considered dataset. Incorporating parameter uncertainty into model uncertainty quantification enables a more robust assessment of model performance under specific atmospheric conditions, which is relevant for applications such as production forecasting and wind farm flow control.

Further research may use the method to formally compare the model-form uncertainty for wind farm flow models of different complexity given benchmark datasets with representative atmospheric conditions. In addition, the parameter and model uncertainty quantified in the posterior can inform robust wind farm flow control and layout optimization. Moreover, the quantified model uncertainty may help evaluate the significance of the obtainable power gains with the controller or optimized layout in the benchmarking conditions. When using the framework with operational farm data, the uncertainty on the inflow conditions can also be incorporated using a similar approach to the hierarchical stochastic prior. In doing so, the model uncertainty may also be separated from the uncertainty in the inflow conditions. Lastly, further research into the accuracy of approximate Bayesian inference methods in this setting, such as the Laplace approximation, variational methods, and Gaussian process emulators, is also of interest to reduce the computational cost of the methodology.

Appendix A: Including local flow redirection in the wake–model coupling in WAYVE

The reduction in the streamwise velocity due to blockage is accompanied by both an increase in the ABL height and a spanwise velocity increase directed away from the centerline of the farm. Hence, the background flow is bidirectional and can be formulated as

$$\boldsymbol{U}_b(\boldsymbol{x}) = \|\boldsymbol{U}_b(\boldsymbol{x})\|_2 \begin{pmatrix} \cos\theta_b(x,y) & \sin\theta_b(x,y) & 0 \end{pmatrix}^\top. \quad \text{(A1)}$$

Since the height-averaged ABL equations solved in the APM do provide a spanwise velocity perturbation, it is expected that a bidirectional wake model will perform better.

Lanzilao and Meyers (2024) derived their wake-merging method for a heterogeneous background velocity field characterized by changes in direction and magnitude. By assuming that the wake only affects the velocity component perpendicular to the rotor and not the velocity component parallel to it that may develop downstream, they arrive at the recursion formula

$$\begin{aligned} \boldsymbol{U}_i(\boldsymbol{x}) =& (\boldsymbol{U}_{i-1}(\boldsymbol{x}) \cdot \boldsymbol{e}_{\perp,i})[1 - W_i(\boldsymbol{X}_i(\boldsymbol{x}))]\boldsymbol{e}_{\perp,i} \\ &+ (\boldsymbol{U}_{i-1}(\boldsymbol{x}) \cdot \boldsymbol{e}_{\parallel,i})\boldsymbol{e}_{\parallel,i}, \end{aligned} \quad \text{(A2)}$$

where $\boldsymbol{e}_{\perp,i} = (\cos\theta_i, \sin\theta_i, 0)$, $\boldsymbol{e}_{\parallel,i} = (-\sin\theta_i, \cos\theta_i, 0)$, and $\boldsymbol{U}_0(\boldsymbol{x}) = \boldsymbol{U}_b(\boldsymbol{x})$. Since the wakes are transported by the mean flow, they introduce a local coordinate system $\boldsymbol{X}_i(\boldsymbol{x}) = (X_i(\boldsymbol{x}), Y_i(\boldsymbol{x}), Z_i(\boldsymbol{x}))$ that is oriented along the streamlines of the background flow field:

$$X_i(\boldsymbol{x}) = \int_{x_i}^{x} \cos\theta_b(\bar{x}, y)\mathrm{d}\bar{x} + \int_{y_i}^{y} \sin\theta_b(x, \bar{y})\mathrm{d}\bar{y}, \quad \text{(A3)}$$

$$Y_i(\boldsymbol{x}) = -\int_{x_i}^{x} \sin\theta_b(\bar{x}, y)\mathrm{d}\bar{x} + \int_{y_i}^{y} \cos\theta_b(x, \bar{y})\mathrm{d}\bar{y}, \quad \text{(A4)}$$

$$Z_i(\boldsymbol{x}) = z - z_{h,i}. \quad \text{(A5)}$$

Since engineering models are mostly designed to be as cheap as possible (for efficient AEP evaluations, layout optimization, and wind farm flow control), it is desirable to circumvent the two integrations per turbine over the whole wake center line. Therefore, we take a similar approach to Stipa et al. (2024b) to reduce computational costs – albeit with another wake-merging method. They argue that the scale at

which the local wind direction changes is much larger than the turbine wake scale which allows ignoring the advection of wake deficits. If additionally, all turbines are aligned with the background flow,

$$\theta_i = \theta_b(\boldsymbol{x}_i) \tag{A6}$$

$$\boldsymbol{U}_b(\boldsymbol{x}_i)\cdot\boldsymbol{e}_{\perp,i} = \|\boldsymbol{U}_b(\boldsymbol{x}_i)\|_2, \tag{A7}$$

wake deflection through yaw misalignment can be neglected. The assumption of slowly varying background wind direction compared to the wake scale allows setting $\theta_b(x, y) \approx \theta_b(x_i, y_i)$ in the integrals such that

$$\boldsymbol{X}_i(\boldsymbol{x}) = \begin{bmatrix} \cos\theta_b(\boldsymbol{x}_i) & \sin\theta_b(\boldsymbol{x}_i) & 0 \\ -\sin\theta_b(\boldsymbol{x}_i) & \cos\theta_b(\boldsymbol{x}_i) & 0 \\ 0 & 0 & 1 \end{bmatrix} (\boldsymbol{x} - \boldsymbol{x}_i). \tag{A8}$$

With these analytical solutions to the integrals, the recursion formula can be made fully explicit. After some matrix manipulations,

$$\boldsymbol{U}_i(\boldsymbol{x}) = \begin{bmatrix} \boldsymbol{e}_{\perp,i} & \boldsymbol{e}_{\parallel,i} \end{bmatrix} \times \begin{pmatrix} (\boldsymbol{U}_{i-1}(\boldsymbol{x})\cdot\boldsymbol{e}_{\perp,i})[1 - W_i(\boldsymbol{X}_i(\boldsymbol{x}))] \\ \boldsymbol{U}_{i-1}(\boldsymbol{x})\cdot\boldsymbol{e}_{\parallel,i} \end{pmatrix} \tag{A9a}$$

$$= \begin{bmatrix} \boldsymbol{e}_{\perp,i} & \boldsymbol{e}_{\parallel,i} \end{bmatrix} \begin{bmatrix} \boldsymbol{e}_{\perp,i}^{\top}[1 - W_i(\boldsymbol{X}_i(\boldsymbol{x}))] \\ \boldsymbol{e}_{\parallel,i}^{\top} \end{bmatrix} \boldsymbol{U}_{i-1}(\boldsymbol{x}), \tag{A9b}$$

it is found that (Devesse et al., 2024a)

$$\boldsymbol{U}(\boldsymbol{x}) = \left(\prod_{i=1}^{N_\mathrm{T}} \mathrm{B}_i(\boldsymbol{x})\right) \boldsymbol{U}_b(\boldsymbol{x}), \tag{A10}$$

where

$$\mathrm{B}_i(\boldsymbol{x}) = \boldsymbol{e}_{\perp,i}\boldsymbol{e}_{\perp,i}^{\top}[1 - W_i(\mathrm{R}_i(\boldsymbol{x} - \boldsymbol{x}_i))] + \boldsymbol{e}_{\parallel,i}\boldsymbol{e}_{\parallel,i}^{\top} \tag{A11}$$

and R_i the rotation matrix defined in Eq. (A8). With a vectorized implementation, the bidirectional wake-merging method has a computational time similar to that of the unidirectional method.

Code and data availability. The large-eddy simulation dataset for wind farm blockage and atmospheric gravity waves in conventionally neutral boundary layers that is used in this work is publicly available (https://doi.org/10.48804/L45LTT; Lanzilao and Meyers, 2023). The code for the wake model and atmospheric perturbation model are available in the Python package WAYVE (https://doi.org/10.48804/XMNVVY; Devesse et al., 2023). The code used to perform the Bayesian uncertainty quantification with a parallelized sequential Monte Carlo algorithm is made available in a Python package coined UMBRA: Uncertainty Modeling toolbox for Bayesian data Re-Analysis (https://gitlab.kuleuven.be/TFSO-software/umbra; Aerts and Meyers, 2026).

Author contributions. FA, KD, and JM jointly defined the scope of the study. FA developed the Bayesian inference framework, implemented necessary algorithms, and performed all simulations. KD provided support for the atmospheric perturbation model and data processing. The article was written by FA and JM and edited by KD.

Competing interests. At least one of the (co-)authors is a member of the editorial board of *Wind Energy Science*. The peer-review process was guided by an independent editor, and the authors also have no other competing interests to declare.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. The authors bear the ultimate responsibility for providing appropriate place names. Views expressed in the text are those of the authors and do not necessarily reflect the views of the publisher.

Financial support. This research has been supported by the European Union Horizon Europe Framework program (HORIZON-CL5-2021-D3-03-04) under grant agreement no. 101084205. The computational resources and services in this work were provided by the VSC (Flemish Supercomputer Center), funded by the Research Foundation Flanders (FWO) and the Flemish Government department EWI.

Review statement. This paper was edited by Cristina Archer and reviewed by two anonymous referees.

References

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

Aerts, F. and Meyers, J.: UMBRA: an Uncertainty Modeling toolbox for Bayesian data Re-Analysis, Gitlab [code], https://gitlab.kuleuven.be/TFSO-software/umbra, last access: 16 March 2026.

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.

Allaerts, D. and Meyers, J.: Sensitivity and feedback of wind-farm-induced gravity waves, J. Fluid Mech., 862, 990–1028, https://doi.org/10.1017/jfm.2018.969, 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.

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

Bastankhah, M., Mohammadi, M. M., Lees, C., Diaz, G. P. N., Buxton, O. R., and Ivanell, S.: A fast-running physics-based

wake model for a semi-infinite wind farm, J. Fluid Mech., 985, https://doi.org/10.1017/jfm.2024.282, 2024.

Beck, J. L. and Zuev, K. M.: Asymptotically independent Markov sampling: a new MCMC scheme for Bayesian inference, Int. J. Uncertain. Quan., 3, 445–474, https://doi.org/10.1615/int.j.uncertaintyquantification.2012004713, 2013.

Bortolotti, P., Dykes, K., Merz, K., and Zahle, F.: IEA Wind Task 37 on systems engineering in wind energy, WP2-Reference Wind Turbines: IEA Wind Task, 37, 565, https://doi.org/10.2172/1529216, 2019.

Branlard, E., Quon, E., Forsting, A. R. M., King, J., and Moriarty, P.: Wind farm blockage effects: comparison of different engineering models, J. Phys. Conf. Ser., 1618, 062036, https://doi.org/10.1088/1742-6596/1618/6/062036, 2020.

Brynjarsdóttir, J. and O'Hagan, A.: Learning about physical parameters: the importance of model discrepancy, Inverse Probl., 30, 114007, https://doi.org/10.1088/0266-5611/30/11/114007, 2014.

Ching, J. and Chen, Y.-C.: Transitional Markov chain Monte Carlo method for Bayesian model updating, model class selection, and model averaging, J. Eng. Mech., 133, https://doi.org/10.1061/(ASCE)0733-9399(2007)133:7(816), 2007.

Clerc, A., Anderson, M., Stuart, P., and Habenicht, G.: A systematic method for quantifying wind flow modelling uncertainty in wind resource assessment, J. Wind Eng. Ind. Aerod., 111, 85–94, https://doi.org/10.1016/j.jweia.2012.08.006, 2012.

Clifton, A., Smith, A., and Fields, M.: Wind plant preconstruction energy estimates, current practice and opportunities, Tech. rep., National Renewable Energy Lab. (NREL), Golden, CO (US), https://doi.org/10.2172/1248798, 2016.

Crespo, A. and Hernández, J.: Turbulence characteristics in wind-turbine wakes, J. Wind Eng. Ind. Aerod., 61, 71–85, https://doi.org/10.1016/0167-6105(95)00033-X, 1996.

Devesse, K., Lanzilao, L., Allaerts, D., Jamaer, S., and Meyers, J.: Wind-fArm gravitY-waVe and blockagE code, KU Leuven RDR, V2 [code], https://doi.org/10.48804/XMNVVY, 2023.

Devesse, K., Lanzilao, L., and Meyers, J.: A meso–micro atmospheric perturbation model for wind farm blockage, J. Fluid Mech., 998, https://doi.org/10.1017/jfm.2024.868, 2024a.

Devesse, K., Stipa, S., Brinkerhoff, J., Allaerts, D., and Meyers, J.: Comparing methods for coupling wake models to an atmospheric perturbation model in WAYVE, J. Phys. Conf. Ser., 2767, 092079, https://doi.org/10.1088/1742-6596/2767/9/092079, 2024b.

Doekemeijer, B. M., Simley, E., and Fleming, P.: Comparison of the Gaussian wind farm model with historical data of three offshore wind farms, Energies, 15, 1964, https://doi.org/10.3390/en15061964, 2022.

Doucet, A. and Johansen, A. M.: A tutorial on particle filtering and smoothing: fifteen years later, in: The Oxford Handbook of Nonlinear Filtering, edited by: Crisan, D. and Rozovskii, B., Oxford University Press, New York, 656–704, ISBN 9780199532902, 2009.

Doucet, A., Freitas, N., and Gordon, N.: Sequential Monte Carlo Methods in Practice, Vol. 1, Springer, https://doi.org/10.1007/978-1-4757-3437-9, 2001.

Foreman-Mackey, D.: corner.py: scatterplot matrices in Python, The Journal of Open Source Software, 1, 24, https://doi.org/10.21105/joss.00024, 2016.

Frandsen, S., Barthelmie, R., Pryor, S., Rathmann, O., Larsen, S., Højstrup, J., and Thøgersen, M.: Analytical modelling of wind speed deficit in large offshore wind farms, Wind Energy, 9, 39–53, https://doi.org/10.1002/we.189, 2006.

Garcia, C. M., Jackson, P. R., and Garcia, M. H.: Confidence intervals in the determination of turbulence parameters, Exp. Fluids, 40, 514–522, https://doi.org/10.1007/s00348-005-0091-8, 2005.

Gaumond, M., Réthoré, P. E., 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, 2014.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B.: Bayesian Data Analysis, Chapman and Hall/CRC, https://doi.org/10.1201/b16018, 2013.

Gelman, A., Simpson, D., and Betancourt, M.: The prior can often only be understood in the context of the likelihood, Entropy, 19, https://doi.org/10.3390/e19100555, 2017.

Geyer, C. J.: Introduction to Markov chain Monte Carlo, in: Handbook of Markov Chain Monte Carlo, 20116022, 22, https://doi.org/10.1201/b10905, 2011.

Gilks, W. R. and Berzuini, C.: Following a moving target-Monte Carlo inference for dynamic Bayesian models, J. Roy. Stat. Soc. B Met., 63, 127–146, 2001.

Göçmen, T., Laan, P. v. d., Réthoré, P.-E., Diaz, A. P., Larsen, G. C., and Ott, S.: Wind turbine wake models developed at the Technical University of Denmark: a review, Renew. Sust. Energ. Rev., 60, 752–769, https://doi.org/10.1016/j.rser.2016.01.113, 2016.

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 Energ. Sci., 7, 1791–1825, https://doi.org/10.5194/wes-7-1791-2022, 2022.

Haario, H., Saksman, E., and Tamminen, J.: An adaptive Metropolis algorithm, Bernoulli, 7, 223–242, https://doi.org/10.2307/3318737, 2001.

Hoffman, M. D. and Gelman, A.: The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo, J. Mach. Learn. Res., 15, 1593–1623, 2014.

Howland, M. F.: Wind farm yaw control set-point optimization under model parameter uncertainty, J. Renew. Sustain. Ener., 13, https://doi.org/10.1063/5.0051071, 2021.

Jensen, N. O.: A Note on Wind Generator Interaction, Risø National Laboratory, ISBN 87-550-0971-9, 1983.

Kantas, N., Beskos, A., and Jasra, A.: Sequential Monte Carlo methods for high-dimensional inverse problems: a case study for the Navier–Stokes equations, SIAM/ASA Journal on Uncertainty Quantification, 2, 464–489, https://doi.org/10.1137/130930364, 2014.

Kennedy, M. C. and O'Hagan, A.: Bayesian calibration of computer models, J. Roy. Stat. Soc. B Met., 63, 425–464, https://doi.org/10.1111/1467-9868.00294, 2001.

Klemmer, K. S. and Howland, M. F.: Wake turbulence modeling in stratified atmospheric flows using a novel $k - l$ model, J. Renew. Sustain. Ener., 17, https://doi.org/10.1063/5.0249278, 2025.

Kwon, S. D.: Uncertainty analysis of wind energy potential assessment, Appl. Energ., 87, 856–865, https://doi.org/10.1016/j.apenergy.2009.08.038, 2010.

Lackner, M., Rogers, A., and Manwell, J.: Uncertainty Analysis in Wind Resource Assessment and Wind Energy Production Estimation, American Institute of Aeronautics and Astronautics, ISBN 978-1-62410-012-3, https://doi.org/10.2514/6.2007-1222, 2007.

Lanzilao, L. and Meyers, J.: A new wake merging method for wind farm power prediction in the presence of heterogeneous background velocity fields, Wind Energy, 25, 237–259, https://doi.org/10.1002/we.2669, 2022.

Lanzilao, L. and Meyers, J.: A reference database of wind-farm large-eddy simulations for parametrizing effects of blockage and gravity waves, KU Leuven RDR, V2 [data set], https://doi.org/10.48804/L45LTT, 2023.

Lanzilao, L. and Meyers, J.: A parametric large-eddy simulation study of wind-farm blockage and gravity waves in conventionally neutral boundary layers, J. Fluid Mech., 979, https://doi.org/10.1017/jfm.2023.1088, 2024.

Lee, J. C. Y. and Fields, M. J.: An overview of wind-energy-production prediction bias, losses, and uncertainties, Wind Energ. Sci., 6, 311–365, https://doi.org/10.5194/wes-6-311-2021, 2021.

Lissaman, P. B.: Energy effectiveness of arbitrary arrays of wind turbines, J. Energy, 3, 323–328, 1979.

LoCascio, M. J., Gorlé, C., and Howland, M. F.: Data-driven wake model parameter estimation to analyze effects of wake superposition, J. Renew. Sustain. Ener., 15, https://doi.org/10.1063/5.0163896, 2023.

McElreath, R.: Statistical Rethinking, Chapman and Hall/CRC, https://doi.org/10.1201/9781315372495, 2018.

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 Energ. Sci., 7, 2271–2306, https://doi.org/10.5194/wes-7-2271-2022, 2022.

Minson, S. E., Simons, M., and Beck, J. L.: Bayesian inversion for finite fault earthquake source models I-theory and algorithm, Geophys. J. Int., 194, 1701–1726, https://doi.org/10.1093/gji/ggt180, 2013.

Moriarty, P., Rodrigo, J. S., Gancarski, P., Chuchfield, M., Naughton, J. W., Hansen, K. S., MacHefaux, E., Maguire, E., Castellani, F., Terzi, L., Breton, S. P., and Ueda, Y.: IEA-task 31 WAKEBENCH: towards a protocol for wind farm flow model evaluation. Part 2: Wind farm wake models, J. Phys. Conf. Ser., 524, https://doi.org/10.1088/1742-6596/524/1/012185, 2014.

Murphy, K. P.: Probabilistic Machine Learning: Advanced Topics, MIT Press, ISBN 978-0-262-04843-9, 2023.

Ndindayino, O., Puel, A., and Meyers, J.: Effect of blockage on wind turbine power and wake development, Wind Energ. Sci., 10, 2079–2098, https://doi.org/10.5194/wes-10-2079-2025, 2025.

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.: Systematic quantification of wake model uncertainty, in: EWEA Offshore Conference, 10–12, 2015.

Nygaard, N. G., Poulsen, L., Svensson, E., and Pedersen, J. G.: Large-scale benchmarking of wake models for offshore wind farms, J. Phys. Conf. Ser., 2265, 022008, https://doi.org/10.1088/1742-6596/2265/2/022008, 2022.

Ørsted: Ørsted presents update on its long-term financial targets, https://orsted.com/en/company-announcement-list/2019/10/1937002 (last access: 18 February 2026), 2019.

Plumlee, M.: Bayesian calibration of inexact computer models, J. Am. Stat. Assoc., 112, 1274–1285, https://doi.org/10.1080/01621459.2016.1211016, 2017.

Quick, J., Mouradi, R.-S., Devesse, K., Mathieu, A., Paul Van Der Laan, M., Murcia Leon, J. P., and Schulte, J.: Verification and validation of wind farm flow models, J. Phys. Conf. Ser., 2767, 092074, https://doi.org/10.1088/1742-6596/2767/9/092074, 2024.

Rodrigo, J. S., Arroyo, R. A. C., Moriarty, P., Churchfield, M., Kosović, B., Réthoré, P. E., Hansen, K. S., Hahmann, A., Mirocha, J. D., and Rife, D.: Mesoscale to microscale wind farm flow modeling and evaluation, WIREs Energy Environ., 6, https://doi.org/10.1002/wene.214, 2017.

Sargsyan, K., Najm, H. N., and Ghanem, R.: On the statistical calibration of physical models, Int. J. Chem. Kinet., 47, 246–276, https://doi.org/10.1002/kin.20906, 2015.

Smith, R. B.: Gravity wave effects on wind farm efficiency, Wind Energy, 13, 449–458, https://doi.org/10.1002/we.366, 2010.

Stipa, S., Ajay, A., Allaerts, D., and Brinkerhoff, J.: The multi-scale coupled model: a new framework capturing wind farm–atmosphere interaction and global blockage effects, Wind Energ. Sci., 9, 1123–1152, https://doi.org/10.5194/wes-9-1123-2024, 2024a.

Stipa, S., Allaerts, D., and Brinkerhoff, J.: A shear stress parametrization for arbitrary wind farms in conventionally neutral boundary layers, J. Fluid Mech., 981, A14, https://doi.org/10.1017/jfm.2024.22, 2024b.

Toussaint, U. V.: Bayesian inference in physics, Rev. Mod. Phys., 83, 943–999, https://doi.org/10.1103/RevModPhys.83.943, 2011.

Trotta, R.: Bayes in the sky: Bayesian inference and model selection in cosmology, Contemp. Phys., 49, 71–104, https://doi.org/10.1080/00107510802066753, 2008.

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 Energ. Sci., 9, 1507–1526, https://doi.org/10.5194/wes-9-1507-2024, 2024.

Walker, K., Adams, N., Gribben, B., Gellatly, B., Nygaard, N. G., Henderson, A., Marchante Jiménez, M., Schmidt, S. R., Rodriguez Ruiz, J., Paredes, D., Harrington, G., Connell, N., Peronne, O., Cordoba, M., Housley, P., Cussons, R., Håkansson, M., Knauer, A., and Maguire, E.: An evaluation of the predictive accuracy of wake effects models for offshore wind farms, Wind Energy, 19, 979–996, https://doi.org/10.1002/we.1871, 2015.

Wasserman, L.: All of Statistics: a Concise Course in Statistical Inference, Springer Science and Business Media, https://doi.org/10.1007/978-0-387-21736-9, 2004.

Wu, S., Angelikopoulos, P., Beck, J. L., and Koumoutsakos, P.: Hierarchical stochastic model in Bayesian inference for engineering applications: theoretical implications and efficient approximation, ASCE-ASME Journal of Risk and Uncertainty

in Engineering Systems, Part B: Mechanical Engineering, 5, https://doi.org/10.1115/1.4040571, 2018a.

Wu, S., Angelikopoulos, P., Papadimitriou, C., and Koumoutsakos, P.: Bayesian annealed sequential importance sampling: an unbiased version of transitional Markov chain Monte Carlo, ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering, 4, https://doi.org/10.1115/1.4037450, 2018b.

Xiao, H. and Cinnella, P.: Quantification of model uncertainty in RANS simulations: a review, Prog. Aerosp. Sci., 108, 1–31, https://doi.org/10.1016/j.paerosci.2018.10.001, 2019.

Zehtabiyan-Rezaie, N. and Abkar, M.: A short note on turbulence characteristics in wind-turbine wakes, J. Wind Eng. Ind. Aerod., 240, https://doi.org/10.1016/j.jweia.2023.105504, 2023.

Zhang, J. and Zhao, X.: Quantification of parameter uncertainty in wind farm wake modeling, Energy, 196, 117065, https://doi.org/10.1016/j.energy.2020.117065, 2020.