ENERGY SOURCES, PART B: ECONOMICS, PLANNING, AND POLICY
https://doi.org/10.1080/15567249.2021.2000520

Taylor & Francis
Taylor & Francis Group

Check for updates

Uncertainty quantification of offshore wind farms using Monte Carlo and sparse grid

Pascal Richter[a], Jannick Wolters[b], and Martin Frank[b]

[a]Theory of Hybrid Systems, RWTH Aachen University, Aachen, Germany; [b]Steinbuch Centre for Computing, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

ABSTRACT

The power produced by an offshore wind farm is subject to multiple uncertainties, such as volatile wind, turbine performance wear, and availability losses. Knowledge about the propagation of these uncertainties and their effect on the produced power is crucial in the design stage of a wind farm. Due to the multitude of uncertainties, an analysis requires high-dimensional numerical integration to determine these parameter sensitivities. Such an analysis has not been done in the current literature for the full set of parameters. In this work, a thorough analysis of all uncertainties is provided, modeled from several years of collected data from the existing wind farms *Horns Rev 1, DanTysk*, and *Sandbank*. The analysis reveals four major parameters, allowing the other parameters to be neglected in future measurement data acquisitions and sensitivity analysis processes. Furthermore, the accuracy of several Uncertainty Quantification techniques is analyzed and a recommendation for future analysis is given.

KEYWORDS

Offshore wind farm; Uncertainty Quantification; Quasi-Monte Carlo; Stochastic Collocation; Sensitivity Analysis

1. Introduction

The computation of the annual energy production in offshore wind farms depends on a large set of parameters. Most of these parameters have a high level of uncertainty due to the impossibility of performing ideal measurements on volatile wind, the occurrence of material fatigue (which influences the performance wear), and other uncertainties, such as availability losses. In order to make reasonable predictions for the annual energy production, however, these uncertainties have to be taken into account.

Few publications have considered uncertainties in parameters for wind farm simulations. Lackner and Elkinton (2007) listed a set of elemental wind farm parameters which have an influence on the uncertainty of the annual energy production, that is, wind direction and speed, a wind turbine's power curve and wake losses. Murcia et al. (2015) analyzed the influence of measurement uncertainties in all of these core parameters except for the wake losses in an attempt to check for modeling errors. Foti, Yang, and Sotiropoulos (2017) on the other hand focused their analysis on the surface roughness and the induction factor which is linked to the thrust coefficient and analyzed their influence on the total power production. Rinker (2016) investigated the sensitivities of four turbulence-related parameters with respect to the load response of a single turbine. Similarly, Ashuri et al. (2016) analyzed the influence of the uncertain parameters wind shear, wind speed, Weibull shape parameter, and air density. Padrón et al. (2016) used polynomial chaos to compute the stochastic effect of wind direction and wind speed on the annual energy production in wind farms. Sarathkumar et al. (2019) consider uncertainties in the cost model for the forecasting of wind power.

CONTACT Pascal Richter pascal.richter@rwth-aachen.de RWTH Aachen University, Research Group for Theory of Hybrid Systems, 55, 52074 Aachen, Germany

© 2021 Taylor & Francis Group, LLC

As can be seen, the current literature deals with a wide variety of uncertainties, but the parameters are either analyzed in isolation or only few parameters are considered. There is a knowledge gap needed to be closed between these specialized analyses and a comprehensive analysis of all major uncertain parameters for an entire wind farm. Therefore, the novelty of this work lies in the comprehensive sensitivity analysis of all nine identified parameters at once with uncertainty estimates based on measurements from several offshore wind farms. Consequently, the existing models were combined and augmented to obtain a unified stochastic model. Although the individual models are mostly linear, the overall model, including the treatment of the measured data, is nonlinear, making a unified stochastic model essential. In addition, the evaluation of uncertainties requires efficient sampling strategies in high-dimensional probability spaces and a significant computational effort. In order to model the uncertainties, we make use of extensive data records sourced from business operations of the Vattenfall Europe Windkraft GmbH.

The outline of this paper is structured as follows: Section 2 describes the deterministic model which we use in order to compute the annual energy production. This involves sub-models for the wake computation, the power generation, and costs. In Section 3 the whole model is extended by introducing randomness into the input parameters to obtain a stochastic model. Section 4 introduces the Uncertainty Quantification methods starting with the classical Monte Carlo method, followed by refinements, such as the quasi-Monte Carlo and finally the Stochastic Collocation method. The performance of each method will be discussed. Subsequently, one of these methods will be recommended based on the presented stochastic model. The stochastic model and the Uncertainty Quantification methods from these two sections are then used in Section 5 where the parameter sensitivities with respect to different cost functions are investigated. All results will afterward be summarized in Section 7.

2. Deterministic model for the AEP and LCOE in wind farms

The annual energy production (AEP) of a wind farm is computed by three sub-models, which describe the wind, wake, and power generation. The temporal integration of gross produced power is computed by using measurements of wind speed and wind direction within one year. For each wind direction and wind speed, the power produced by the wind farm needs to be computed with the integral over all directions, resulting in the produced energy. This value is then used in a cost model to describe economic quantities, as shown in Figure 1. In the following all of the aforementioned sub-models are described.

2.1. Wind model

Starting from the raw energy source, wind is quantified by thousands of wind direction and speed measurements. Figure 2 shows a clustered wind direction distribution over a time period of eight years. Furthermore, for each wind direction sector, the distribution of wind speed is considered independently.

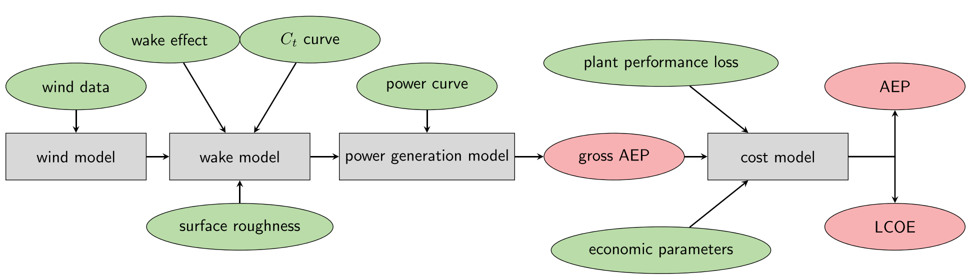
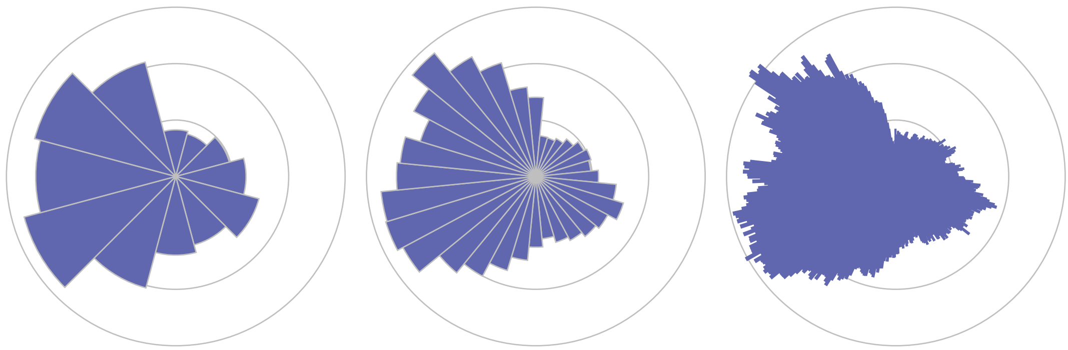


Figure 1. Structure of the offshore wind farm model. Input parameters are highlighted in green. These parameters are also later used for the uncertainty quantification. Possible outputs of the wind farm model (highlighted in red) are the annual energy production (AEP) and the levelized cost of electricity (LCOE).



(a) 12 sectors (b) 32 sectors (c) 360 sectors

Figure 2. Wind direction measurements at the FINO3 research platform in the North Sea, about 80 km away from the German Island Sylt. The measurements at a 100 m height taken over eight years from January 2010 to December 2017 are clustered into N_{dir}=12, 32, and 360 direction sectors.

The raw measurement data is commonly modeled by a Weibull distribution, see Genc et al. (2005), which is fitted to the data using a maximum likelihood estimation, see Carrillo et al. (2014),

$$\mathcal{W}(u;\lambda,k) = \left(\frac{k}{\lambda}\right)\cdot\left(\frac{u}{\lambda}\right)^{k-1}\cdot\exp\left(-\left(\frac{u}{\lambda}\right)^{k}\right), \tag{1}$$

where $\lambda > 0$ is the scale parameter and $k > 0$ is the shape parameter of the Weibull distribution, as plotted in Figure 3.

In the following sections, we will refer to the general model as deterministic, meaning no input parameters are subject to any kind of uncertainty.

For the simulation of a wind farm, we need to consider a discrete number of wind direction sectors and wind speeds. Increasing the number of discrete bins N_{dir} and N_{speed} generally results in a more precise model as long as the individual distributions can still be fitted with reasonable accuracy.

2.2. Wake model

In an offshore wind farm, several wind turbines are arranged in a grid bounded by a predefined region. The wind passing the turbine blades generates a wake behind each turbine, which is responsible for turbulence-induced wind speed reductions for subsequent wind turbines. The task of a wake model is to compute this affected velocity field behind each wind turbine.

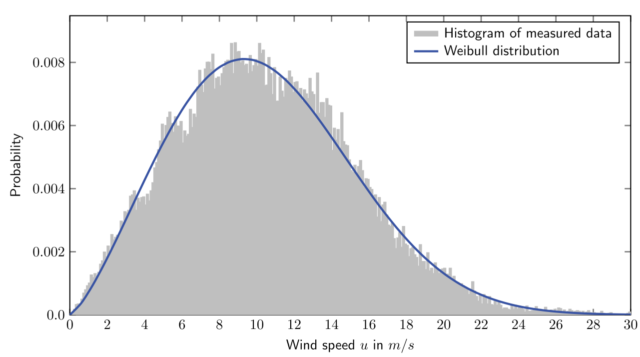


Figure 3. Wind speed measurements at the FINO3 research platform at a height of 100 m for the eight years from January 2010 to December 2017 for the western wind direction sector (255° to 285°). The data is fitted by a Weibull distribution using the maximum likelihood estimation.

In the literature, one can find many different models which describe the wake effect. The first notable wake model (*PARK model*) was developed by N. O. Jensen (1983) and Katic, Højstrup, and Jensen (1986). Further developments are the *Eddy-Viscosity wake model* by Ainslie (1988) and some extensions by Lange et al. (2003); Larsen et al. (2007), the *Deep-array wake model* by Brower and Robinson (2012), a *linearized Reynolds-averaged Navier–Stokes* model by Ott (2009), and *Large Eddy Simulations* by Stovall, Pawlas, and Moriarty (2010).

In this paper, we focus on the *PARK model*, which is mainly used in commercial software tools, because the velocity deficit calculated by the model has a top-hat shape which allows for a simple and efficient computation. This efficiency is also needed in the context of this paper as the high-dimensional integration, which is needed to obtain the statistical quantities of interest, requires a large amount of model evaluations and therefore is not feasible with full CFD Navier-Stokes simulations. Furthermore, investigations by Barthelmie et al. (2006) show that the computations are rather accurate compared to other more detailed models, such as the k-ε turbulence model which is based on a Navier–Stokes model by Schepers (2003). It is well known that PARK overestimates wake losses and does not consider deep array effects of wind farms, see Gaumond et al. (2014); Pillai, Chick, and Laleu (2014). In order to compensate for this deficiency, we consider an additional loss term ℓ_{wake} in the model.

The velocity deficit inside the wake only changes in down-stream direction. The wake radius in the PARK model grows with a constant factor $k := 0.5/\ln(z/z_0)$, which depends on the hub height z of the turbine and the surface roughness z_0. The wake diameter $D_w(x)$ grows linearly by a factor of $2k$, as illustrated in Figure 4. For an inflow velocity u_0, the wake velocity $u_w(x)$ at any point inside the wake of a turbine with rotor diameter D is given by

$$u_w(x) = (\ell_{\text{wind}} \cdot u_0) - \frac{1 - \sqrt{1 - C_t(\ell_{\text{wind}} \cdot u_0)}}{\left(1 + \frac{x}{D} \cdot \frac{1}{\ln(z/z_0)}\right)^2} \cdot \ell_{\text{wake}} \cdot (\ell_{\text{wind}} \cdot u_0), \tag{2}$$

where the last fraction describes the velocity deficit. The velocity-dependent thrust coefficient of the turbine $C_t(u_0)$ is a characteristic property of the turbine type and thus needs to be provided by the manufacturer of the wind turbine (Figure 5). The loss parameter ℓ_{wake} is used to consider wake effect losses due to internal turbine arrays, external turbines and future developments in the vicinity of the wind farm. Furthermore, we consider wind speed losses ℓ_{wind} due to turbulence, off-yaw axis winds, inclined flow, and high shear wind flow. Loss parameters are mainly based on experience and thus are used to make the developed model more realistic, by closing the gap between laboratory conditions and actual conditions at the site.

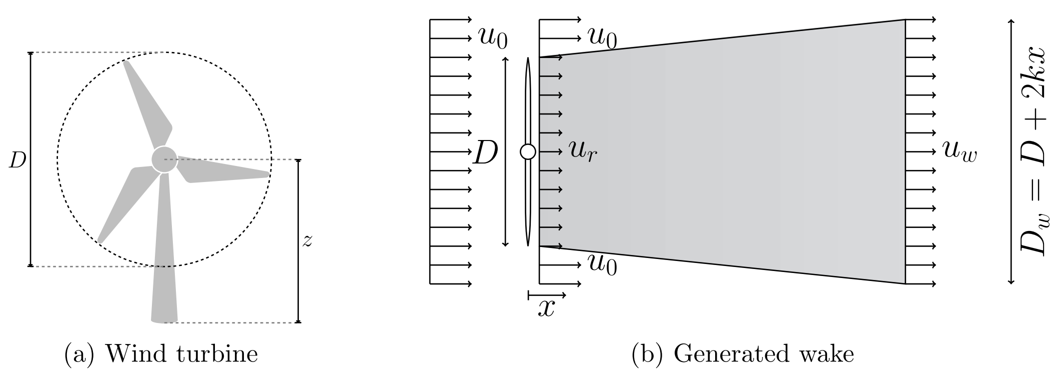


Figure 4. Wind turbine with hub height z and rotor diameter D. The wake diameter D_w grows linearly by $2k$. Inside the wake, the incident wind speed u_0 is reduced to the wake velocity u_w. The Figures are taken from Heiming (2015).

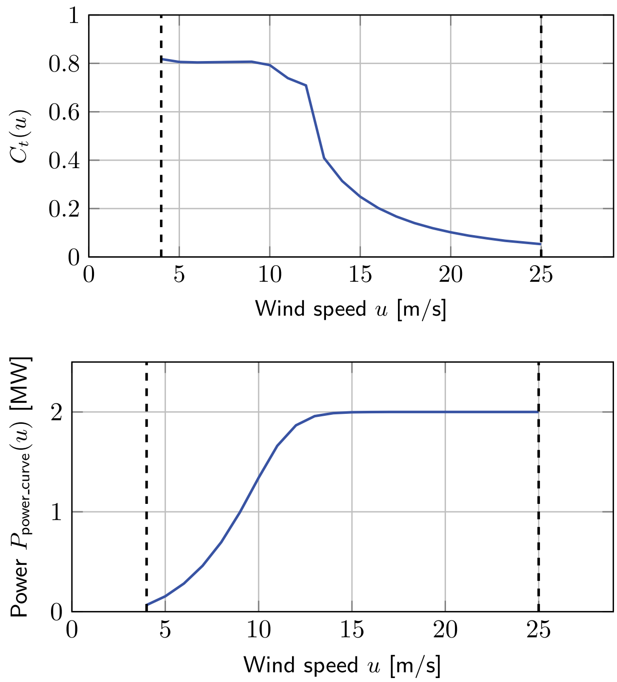


Figure 5. Thrust coefficient C_t and power production of the turbine Vestas V80 with cut-in speed of 4 m/s and cutout speed of 25 m/s (dashed vertical lines).

As introduced in the PARK model, to compute the incident velocity of a wind turbine, we simply use a weighting factor β between the free stream velocity u_0 and the wake velocity $u_w(x)$, which is computed by the circular intersection $A_{\text{Intersection}}$ of the wake cross-section with the turbine's circular area A_{Turbine}, $\beta = \frac{A_{\text{Intersection}}}{A_{\text{Turbine}}}$.

2.3. Power generation model

Turbines convert the wind's kinetic energy into electrical energy. Thus, the generated power of a wind turbine P depends on the incident wind speed u and the power curve (which is usually provided by the manufacturer as seen in Figure 5),

$$P(u) = P_{\text{power\_curve}}(u) \cdot \ell_{\text{power}}. \tag{3}$$

With the parameter ℓ_{power}, power curve losses such as material performance deviations from the expected curve are considered. The power curve depends on a cut-in speed u_{cutin} and a cutout speed u_{cutout} which specify the range of wind speeds in which the turbine generates power. If the wind speed is lower than u_{cutin} there is not enough wind for efficient power production, and for wind speeds larger than u_{cutout} the wind is too strong such that the turbine might be damaged.

The gross annual energy production is given as the produced power for the duration of one year (in hours)

$$E_{\text{AEPgross}} = (8766h) \cdot P, \tag{4}$$

while the total generated power P is computed from the generated power for each wind direction φ,

$$P := \int_0^{2\pi} P_\varphi d\varphi \approx \frac{1}{N_{\text{dir}}} \cdot \sum_{i=1}^{N_{\text{dir}}} P_{\varphi_i}, \tag{5}$$

with direction φ_i. The power P_{φ_i} for one wind direction is given by integrating along the corresponding probability function and the power curve. For a discrete number of wind speeds N_{speed} and a total of N_{turbine} turbines in the wind farm, all in all the annual production is approximated by

$$P \approx \frac{1}{N_{\text{dir}}} \cdot \sum_{i=1}^{N_{\text{dir}}} w_{\phi_i} \cdot \sum_{j=1}^{N_{\text{speed}}} w_j \cdot \mathcal{W}_{\phi_i}(u_j) \cdot \sum_{k=1}^{N_{\text{turbine}}} P(u_j, \phi_i, k) \tag{6}$$

with wind speed u_j, wind direction ϕ_i for turbine k.

2.4. Cost model

As a result of the last three sub-models we can compute the gross annual energy production by integrating over each wind direction and wind speed, computing the wake velocity for subsequent turbines and evaluating the power curve (see Figure 1). This value is now used to compute different economic indicator functions, for example, the annual energy production, the levelized cost of electricity by Lackner and Elkinton (2007), the net present value by González et al. (2009), or the internal rate of return by Ç Akar (2017). In this work, we consider the following two economic indicator functions:

- The *annual energy production*,

$$E_{\text{AEP}} = E_{\text{AEPgross}} \cdot \ell_{\text{performance}}, \tag{7}$$

is the basic quantity for most economic indicator functions. With the plant performance loss $\ell_{\text{performance}}$, we consider electrical losses due to availability of turbines, high wind hysteresis, and environmental performance degradation, such as icing and high temperatures.

- The *levelized cost of electricity* given by Lackner and Elkinton (2007) and Ioannou, Angus, and Brennan (2018),

$$K_{\text{LCOE}} = \frac{C_{\text{capital}} \cdot \frac{(1+r_{\text{rate}})^T \cdot r_{\text{rate}}}{(1+r_{\text{rate}})^T - 1} + C_{OM}}{\text{E}_{\text{AEP}}}, \tag{8}$$

with total installed capital costs C_{capital} for turbines, cabling, substation, decommission, etc., annual operation and maintenance costs C_{O}M, and discount rate r_{rate} including debt, taxes, and insurance over the time period T.

2.5. Validation of the model

The implemented model delivers same results for AEP and LCOE as the Openwind software, see Brower and Robinson (2012).

3. Stochastic model for the AEP and LCOE in wind farms

The models described in Section 2 are purely deterministic, that is, they will always compute the same results given the same input parameter. In order to consider perturbances in some parameters we need to introduce random variables and include them into the model to obtain a stochastic model.

3.1. Uncertain parameters

The deterministic model from Section 2 is extended by introducing uncertainties in the parameters. In DNV (2013) several uncertainty–inflicted parameters at high level have been identified, here listed in Table 1, Table 2.

In this work, we model all random variables as independent and normally distributed $\xi_i \sim \mathcal{N}(\mu_i, \sigma_i)$, with mean μ_i from the deterministic model. It is important to note that this is a modeling assumption that might not be justified for some of the presented uncertainties. For parameters with implied bounds, where the unbounded support of the normal distribution can lead to problems, we instead use a truncated normal distribution (e.g. the value under the square root in equation 10 can in principle become negative), see Table 1 for a complete list of all random variables and their associated distributions. Choosing normal distributions also yields benefits for the analysis of the resulting high-dimensional joint probability distribution function in the case studies presented in Section 5. The economics-based random variables are almost surely independent from the ones from the physical models and are also considered to be independent from each other as they all address separate topics. The same holds for most of the uncertain variables from the physical domain. The uncertainty in the thrust coefficient C_t for example is considered to be a geometric design property of the turbine blades and thus the uncertainty lies within the generated thrust and not the static pressure on the blades. This also holds for the power curve where the uncertainty is modeled in the mapping of the wind velocity to the produced power and thus is independent from the velocity itself. Recently, the effect of lift coefficients on the performance of wind turbine blades has been emphasized as a factor contributing to uncertainty. This influences the estimation of the power curve and can increase uncertainty sources, see Li and Caracoglia (2020). Thus, for the independence of uncertain parameters we only assume the independence of the uncertainty on the surface roughness and the uncertainty in the measured wind speed.

Each random variable can be rewritten as a multiplication between the unperturbed value and a normal distribution $\mathcal{N}(1, \sigma)$ with the mean value of one. As the standard deviation σ is a relative property, it can be chosen to resemble any observed deviation d. In this work, the deviation d of the measurement data describes the exceedance probability of 90% and therefore σ can be computed by evaluating the inverse of its distribution function for which, in case of the normal distribution, 90% of the data lies below the bell curve. The following uncertain parameters are considered:

- *Wind speed*, due to site measurements, historic wind resource or the measure-correlate-predict method, vertical extrapolation, future variability, and wind flow extrapolation.
- *Wake effect*, due to model inaccuracies and neighboring sites.
- C_t *curve*, due to impacts of atmospheric stability and site conditions for which C_t curve is not valid.
- *Surface roughness*, due to changing surface conditions caused by weather or tides.
- *Power curve*, due to impacts of atmospheric stability, and site conditions for which power curve is not valid.
- *Plant performance*, due to electrical efficiency, availability of turbines, internal and external grid, due to environmental as blade soiling, blade degradation and weather effects.
- *Capital costs*, due to steel price fluctuations.

Table 1. Uncertain parameters of the wind farm model at high level.

Uncertain parameters		Distribution
Wind speed	ξ_{wind}	Normal
Wake effect	ξ_{wake}	Normal
C_t curve	ξ_{ct}	truncated Normal
Surface roughness	ξ_{rough}	truncated Normal
Power curve	ξ_{power}	Normal
Plant performance	$\xi_{\text{performance}}$	Normal
Capital costs	ξ_{capital}	Normal
Annual O & M costs	ξ_{om}	Normal
Discount rate	ξ_{rate}	Normal

- *Annual O&M costs*, due to use of new technologies.
- *Discount rate*, due to fluctuations of the economic discount rate.

3.2. Stochastic model

The above defined random variables are now included into the model to obtain a stochastic model.

Wind Due to inaccurate measurements, vague long-term predictions, inter-annual variability, and further interference DNV (2013), the distribution of the wind speed u is a highly uncertain parameter. Therefore, we perturb the raw data of the wind speed with a normally distributed random variable ξ_{wind} such that perturbed probability density functions of the Weibull distribution are obtained (see Figure 6). In Ç Akar (2017) and in Tuzuner and Yu (2008) it is shown that a perturbance d of the wind speed corresponds to the perturbance d of the Weibull parameter λ, such that the resulting probability for each wind speed u can be formulated as a random variable. Compare with equation (1):

$$\mathcal{W}(u;\lambda,k,\xi_{\text{wind}}) = \left(\frac{k}{\lambda\cdot\xi_{\text{wind}}}\right)\cdot\left(\frac{u}{\lambda\cdot\xi_{\text{wind}}}\right)^{k-1}\cdot\exp\left(-\left(\frac{u}{\lambda}\cdot\xi_{\text{wind}}\right)^{k}\right), \tag{9}$$

where $\lambda > 0$ is the scale parameter and $k > 0$ is the shape parameter of the Weibull distribution, which are determined with the maximum likelihood estimation using the unperturbed wind speed data.

Wake Within or even outside of a wind farm strong wake effects are present from, for example, turbine in upstream direction or other close-by wind parks. The predicted wake effect in these cases is perturbed due to uncertainty in the model inputs (including wind direction), model performance, and appropriateness for the site. Furthermore, we need to include uncertainties related to any proposed neighboring sites (construction time, layout, turbine type), see DNV (2013). Therefore, we perturb the velocity deficit of a wake with a normally distributed random variable ξ_{wake}. Additionally, the wake model depends on the C_t curve and the surface roughness z_0. Due to imprecise measurements of the C_t curve, we perturb this parameter with the random variable ξ_{ct} (see Figure 7). Here, in order to prevent negative values in the root of the numerator of equation (10), we used a truncated normal distribution, which enforces $C_t(u_0)\cdot\xi_{\text{ct}} \leq 1$ as this could otherwise lead to nonphysical complex wake velocities.

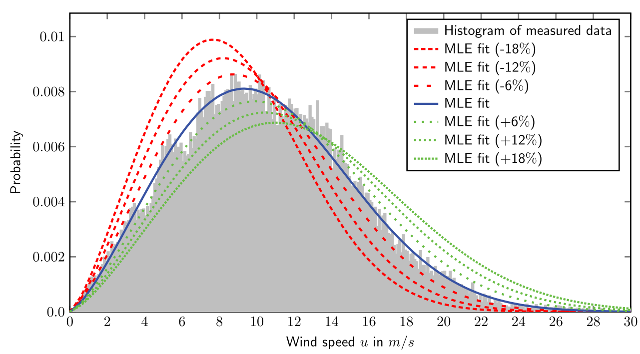


Figure 6. Fitted Weibull distribution with maximum likelihood estimation (MLE) for the wind speed distribution over the years from January 2010 to December 2017 for the western wind direction sector (255° to 285° measured at the FINO3 research platform. Red and green plots represent the fitted Weibull distribution with MLE after perturbing the wind speed data with factors of $\pm 6\%$, $\pm 12\%$, and $\pm 18\%$. The random variable $\xi_{wind}(u)$ represents the Weibull distribution between the highest red plot (MLE fit −18%) and the lowest green plot (MLE fit +18%) with a probability of 99.73%.

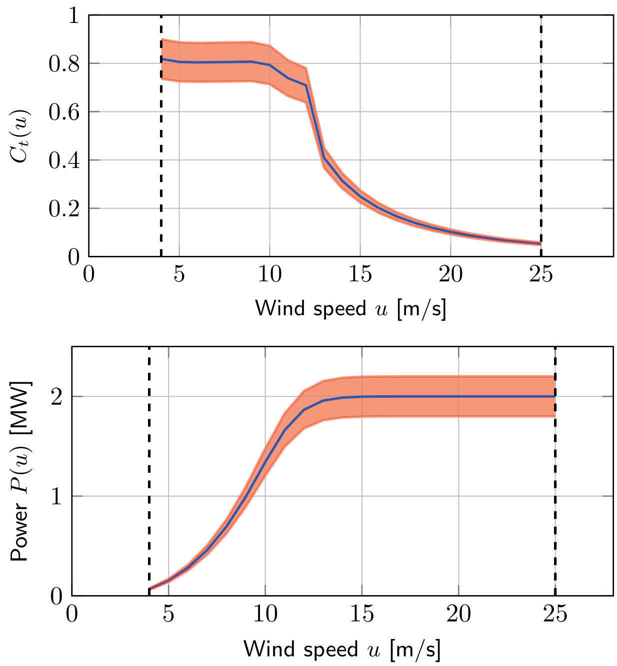


Figure 7. Exemplary perturbed C_t and power curve of a Vestas V80 turbine with a cut-in speed of 4 m/s and cutout speed of 25 m/s (dashed vertical lines). The red tube illustrates the C_t and power curve with a perturbance of $\pm 10\%$.

The surface roughness depends on the topography and flora, see Wiernga (1993), that is, for offshore wind farms it depends on the wave field, wind speed, upstream fetch, and water depth, see Lange et al. (2004). Thus, also this parameter z_0 should be stochastic and is therefore perturbed with a truncated normally distributed random variable ξ_{rough} to enforce $z_0 \cdot \xi_{\text{rough}} > 0$. Altogether, the computed wind speed behind a turbine at coordinate x from equation (2) changes as follows:

$$\widetilde{u_w}(x, \xi_{\text{wake}}, \xi_{\text{ct}}, \xi_{\text{rough}}) = (\ell_{\text{wind}} \cdot u_0) - \frac{1 - \sqrt{1 - C_t(\ell_{\text{wind}} \cdot u_0) \cdot \xi_{\text{ct}}}}{\left(1 + \frac{x}{D} \cdot \frac{1}{\ln(z/(z_0 \cdot \xi_{\text{rough}}))}\right)^2} \cdot \ell_{\text{wake}} \cdot \xi_{\text{wake}} \cdot (\ell_{\text{wind}} \cdot u_0). \tag{10}$$

Power generation In the power generation model the turbine performance is perturbed due to material fatigue which leads to uncertainty in the power curve. Furthermore, there is uncertainty in the performance under site conditions for which the power curve is not valid. This also includes the impact of atmospheric stability and icing losses as well as other environmental losses, e.g. blade soiling, blade degradation, weather effects. As it is difficult to quantify these sources of uncertainty in terms of individual standard deviations, we model them as a single normally distributed random variable ξ_{power} (see Figure 7). This changes equation (3) for the power curve as follows:

$$\widetilde{P}(u, \xi_{\text{power}}) = P_{\text{power\_curve}}(u) \cdot \ell_{\text{power}} \cdot \xi_{\text{power}}. \tag{11}$$

Annual energy production The *annual energy production* is the basic value for all other economic indicator functions. As plant performance losses due to, for example, availability or curtailment are uncertain, we perturb the performance with the normally distributed random variable $\xi_{\text{performance}}$, which results into the following formula:

$$\widetilde{E}_{\text{AEP}} = E_{\text{AEPgross}} \cdot \ell_{\text{performance}} \cdot \xi_{\text{performance}}. \tag{12}$$

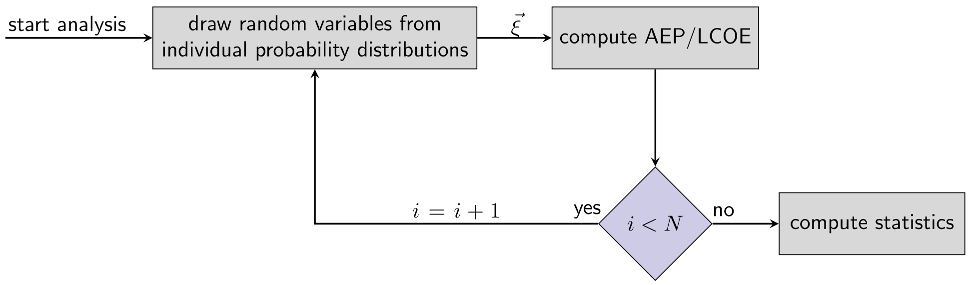


Figure 8. Flow diagram showing the basic methodology used to analyze the influence of uncertainties w.r.t. the respective quantity of interest.

Levelized cost of electricity As the capital costs C_{capital} in the LCOE formula 8 mainly depend on the price of steel, but because of a wind farms long planning stage, the calculation depends on long-term predictions for the steel price which is very volatile. In order to model these uncertainties, we perturb the capital costs with a normally distributed random variable ξ_{capital}. The same argument holds for the discount rate r_{rate}, which during an early planning stage is very tentative thus it is perturbed with a normally distributed random variable ξ_{rate}. The costs for annual operation and maintenance C_{OM} are also affected by the volatile price of steel (for the material), and other political decisions like payroll taxes. Therefore, this parameter is also perturbed with a normally distributed random variable ξ_{om}. Altogether, the levelized cost of electricity from (8) changes as follows:

$$K_{\text{LCOE}} = \frac{C_{\text{capital}} \cdot \xi_{\text{capital}} \cdot \frac{(1+\tilde{r}_{\text{rate}})^T \cdot \tilde{r}_{\text{rate}}}{(1+\tilde{r}_{\text{rate}})^T - 1} + C_{\text{OM}} \cdot \xi_{\text{om}}}{E_{\text{AEP}}}, \tag{13}$$

with perturbed discount rate $\tilde{r}_{\text{rate}} = r_{\text{rate}} \cdot \xi_{\text{rate}}$.

3.3. Methodology

In order to quantify the uncertainties of the previously introduced stochastic model sampling strategies will be used in the following analysis to compute the desired statistics. These sampling strategies draw random variables from each parameters probability distribution and use these parameters to compute the respective quantity of interest. This process is repeated until a certain amount of samples has been drawn or until a convergence criterion is fulfilled. Afterwards, the generated samples from the multidimensional stochastic space can be used to compute statistical properties of the quantity of interest as, for example, the mean or variance. A schematic representation of this process can be seen in Figure 8.

4. Methods of Uncertainty Quantification

As it is the goal of this paper to investigate the influence of uncertainties in the input parameters onto different model outputs, for example, gross annual energy production, it is necessary to introduce methods which allow for the efficient integration of the stochastic problem introduced in section 3 over the associated probability space. Performing this integration efficiently in high-dimensional probability spaces is a challenging task and this chapter introduces three methods that are commonly used in the field of Uncertainty Quantification (UQ).

From a practical point of view the methods for Uncertainty Quantification, see Smith (2013), Sullivan (2015), de Cursi and Sampaio (2015), Ghanem, Higdon, and Owhadi (2017), can be divided into two different categories: While *intrusive* methods introduce changes to the original problem as the governing equations become statistical, *non-intrusive* methods on the other hand evaluate the original problem with varying inputs and compute the statistics from the results. For complex problems, a non-intrusive method is often favored as it requires no modification to the original code. This can be seen in the literature as, for example, Lackner and Elkinton (2007) and Foti, Yang, and Sotiropoulos (2017) used a Monte Carlo simulation and Murcia et al. (2015) performed a Latin Hypercube simulation for their investigations.

As the underlying simulation is a complex data-driven procedure, we also make use of non-intrusive methods in this work. Therefore, the classical Monte Carlo method, see Smith (2013), low-discrepancy quasi-Monte Carlo methods by Niederreiter (1992), and also the Stochastic Collocation method by Babuška, Nobile, and Tempone (2007) are used. We will then provide a recommendation for the method, which performs best with respect to the underlying offshore wind farm problem setting.

4.1. Monte Carlo

In order to compute a given quantity of interest, such as the expectation of the annual energy production, the problem can mathematically be written as:

$$\mathbb{E}[u(\vec{X},\vec{\xi})] = \int_{\Theta} u(\vec{X},\vec{\xi})\rho(\vec{\xi})d\vec{\xi}, \tag{14}$$

where $u(\vec{X},\vec{\xi})$ is the disturbed model involved in computing the annual energy production, $\vec{\xi} \in \Theta \subseteq \mathbb{R}^d$ are the random variables modeling uncertainties within the inputs, $\rho(\vec{\xi})$ as the respective probability density function, and $\vec{X}$ the vector of undisturbed parameters.

The Monte Carlo method uses sampling in the probability space of the associated random variable $\vec{\xi}$ to evaluate the integral in equation 14. By computing M solutions (sampling), each starting from a different set of realizations of the uncertain parameters, M solutions of the type $u^m(\vec{X}) = u(\vec{X},\vec{\xi}^m)$ are obtained. If $\vec{\xi}^m$, $m = 1,\dots,M$ is a sequence of independent and identically distributed random variables, application of the central limit theorem yields:

$$\frac{1}{M}\sum_{m=1}^{M} u(\vec{X},\vec{\xi}^m) \xrightarrow{a.s.} \mathbb{E}[u(\vec{X},\vec{\xi})].$$

This means that, in the limit, the method converges to a fixed value for the mean and also the variance of the Quantity of Interest. The rate of convergence for the Monte Carlo method with random sampling is $\mathcal{O}(M^{-1/2})$, as shown by Caflisch (1998). $\mathcal{O}(\cdot)$ in this context describes the upper bound for the growth rate of a function. Therefore, in order to achieve one additional digit of accuracy, it is necessary to compute 100 times more samples. This slow convergence rate can cause issues in case of computationally expensive problems. Even though the convergence rate of most Monte Carlo type methods does not depend on the dimension of the underlying probability space, it is important to note that in still does in practice as can easily be demonstrated with a simple experiment.[1]

Generating random numbers for the sampling process is a difficult task in practice as computers are deterministic machines. In this work, we use pseudo-random numbers generated by the Mersenne Twister method by Matsumoto and Nishimura (1998) from the C++ standard library. A visual representation of generated samples can be seen in Figure 9a.

[1]Trying to approximate the volume of a unit sphere by drawing random samples from a unit cube and determining if the sampled point is inside the sphere works well in two or three dimensions, but for arbitrary dimension d the ratio of points inside of the sphere to the total number of samples converges to the cube's volume divided by 2^d. This leads to the conclusion that the error constant, which has been omitted in the $\mathcal{O}(\cdot)$ notation, rapidly increases with dimension d.

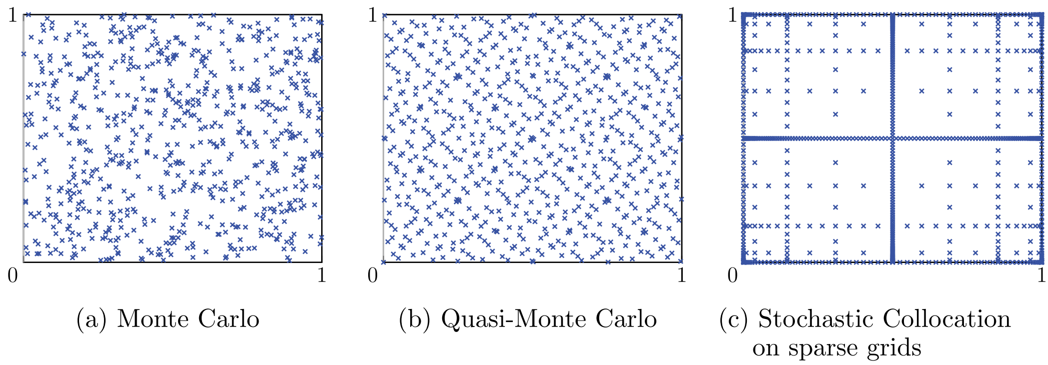


Figure 9. Visual representation of different sampling strategies on a two-dimensional unit square. (a) classical Monte Carlo with pseudo random numbers generated by the Mersenne Twister Engine from the C++ standard library, (b) Quasi-Monte Carlo with pseudo random numbers generated by the Sobol sequence, and (c) Stochastic Collocation on Smolyak sparse grids with Clenshaw Curtis nodes.

4.2. Quasi-Monte Carlo

An improvement to the classical Monte Carlo method is the quasi-Monte Carlo method (QMC). It relies on the sample principle like the classical Monte Carlo method with the difference being that it makes use of a low-discrepancy sequence in order to generate its quasi-random numbers. Morokoff and Caflisch (1995) examined three different low discrepancy sequences: the Halton, Sobol, and Faure sequence. The result indicated that Halton sequences are best for up to six dimensions and the Sobol sequence is best for all higher dimensions. As we are interested in these high-dimensional cases, the Sobol sequence is used in this paper. The Sobol sequence can be briefly explained as a sequential instruction set that fills a multi-dimensional hypercube, while trying to avoid the creation of void regions. These created values are deterministic and thus are called pseudo-random, but they evenly fill the hypercube and therefore potentially lead to a faster convergence compared to the pure Monte–Carlo method. The more equally spaced sample points can be seen in Figure 9b. For a detailed explanation regarding the generation of the sequence, see Bratley and Fox (1988). By using the Sobol sequence to generate pseudo-random numbers, the convergence of the quasi-Monte Carlo is of the order $\mathcal{O}(M^{-1}(\log M)^{d})$, see Caflisch (1998). This means that for small dimensions, the quasi-Monte Carlo simulation only needs to compute roughly five to ten times fewer samples in order to achieve one additional digit of accuracy compared to the classical Monte Carlo method.

4.3. Sparse grid stochastic collocation

Compared to the previous methods, the sparse grid Stochastic Collocation (SC) method is not a variant of the classical Monte Carlo method. The main idea of this method is to choose a set of M collocation points in probability space and then to compute the solution at these points. As the positions of the collocation points are generally arbitrary, the sparse grid Stochastic Collocation method selects them based on a quadrature rule and exploits the corresponding quadrature weights to compute the statistics of a given quantity of interest, such as for example, the annual energy production. In case of a high-dimensional probability space it is thus necessary to use an efficient quadrature rule, as for common methods the number of required quadrature points increases exponentially with the dimension of the probability space. This leads to expensive computations as a solution of the deterministic problem needs to be computed on every quadrature point.

One way to construct such an efficient rule is by using the so-called Smolyak sparse grids. These grids are constructed from nested one-dimensional quadrature rules, which results in a slower growth rate of the required quadrature points. In this paper, Smolyak sparse grids with Clenshaw Curtis nodes are used for the computations involving the sparse grid Stochastic Collocation method. As this node type yields a quadrature set for the bounded interval $[-1, 1]$ and thus would not be a suitable quadrature rule for unbounded probability density functions as, for example, the normal distribution, we use an inverse cumulative distribution function transform to expand the quadrature set to $(-\infty, \infty)$ as described in van Wyk, Gunzburger, and Burkardt (2016). For further details on the sparse grid method, see Smith (2013); Wolters (2016). Figure 9c shows an example for such a sparse grid for the domain $[0, 1]^2$.

The convergence rate of the method is of the order $\mathcal{O}(M^{-\alpha}(\log M)^{(d-1)(\alpha+1)})$, with dimension of the uncertain parameter space d, M as total number of grid points and α depending on the regularity of the solution. It has to be stressed that the number of samples M is indirectly determined by the dimensionality of the probability space and the level of the underlying quadrature rule, which can be selected by the user.

4.4. Error Comparison of UQ Methods

In order to demonstrate the convergence behavior of the described methods, we use the computation of the levelized cost of electricity (see Figure 1) as a benchmark. The uncertain parameters in this computation are equal to the maximal amount of parameters used in the results presented in Section 5. The associated probability distributions used in the sampling process for each parameter are given by Table 4, while the remaining input data is configured according to the *Horns Rev 1* wind farm dataset. As we are computing the levelized cost of electricity, all nine parameters from Table 4 have an influence on the solution and therefore the resulting joint probability space is nine-dimensional.

The convergence rates of all presented methods are summarized in Table.

As these values only show the theoretical order and neglect the influence of any constant factors, Figures 10 and 11 show the numerically studied error evolution patterns. This is especially interesting for the sparse grid Stochastic Collocation method, as its convergence rate depends on the unknown smoothness α of the underlying problem. The methods are compared in terms of the relative error in

Table 2. Order of the convergence rates of the presented UQ methods.

Monte Carlo	quasi-Monte Carlo	Sparse Grid Stochastic Collocation
$\mathcal{O}(M^{-1/2})$	$\mathcal{O}(M^{-1}(\log M)^{d})$	$\mathcal{O}(M^{-\alpha}(\log M)^{(d-1)(\alpha+1)})$

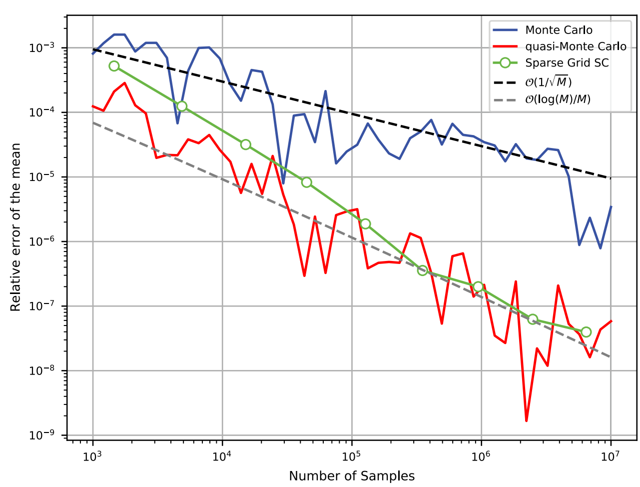


Figure 10. Convergence behavior of the presented methods in terms of the relative error in the mean w.r.t to the quasi-Monte Carlo reference solution.

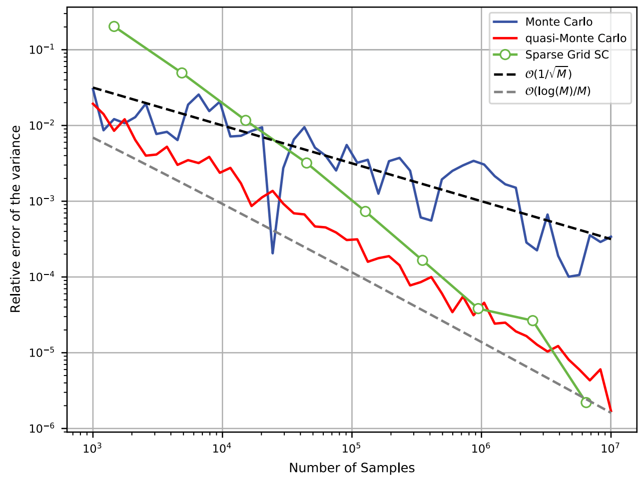


Figure 11. Convergence behavior of the presented methods in terms of the relative error in the variance w.r.t to the quasi-Monte Carlo reference solution.

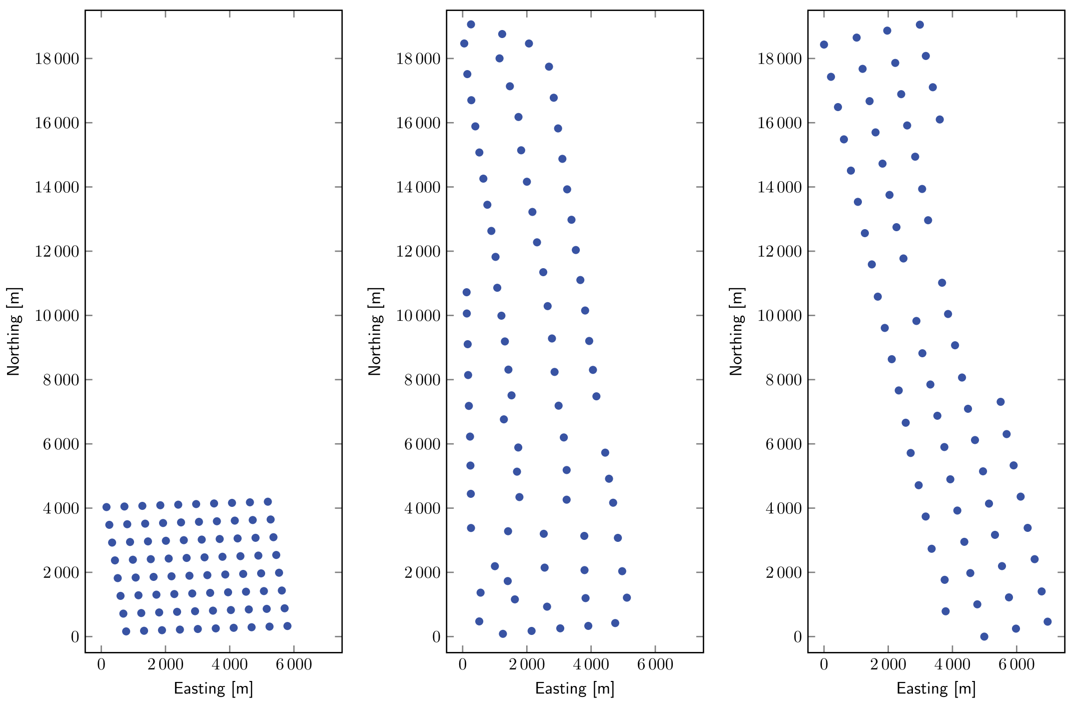


Figure 12. Turbine positions for the wind farms Horns Rev 1 (left), DanTysk (center), and Sandbank (right). The constructed wind farm *Horns Rev 1* is located in the North Sea close to the Danish coast. This wind farm has been built in 2002 and is commonly used as a test case for all kinds of offshore wind farm related research, e.g. Murcia et al. (2015), Gaumond et al. (2014), Barthelmie et al. (2009). The *DanTysk* wind farm is located close to the German shore in the North Sea, approximately 70 km west of the Island Sylt. Build in 2014, it covers an area of 70 square meters and can theoretically produce 288 MW of power, if each of its 80 turbines would operate at full load. The *Sandbank* wind farm is located right next to the DanTysk wind farm and is the newest of all presented wind farms as it was recently built in 2017. Containing 72 turbines, laid out in rows across an area of 60 square kilometers, each capable of producing 4 MW of electrical power, the wind farm has a theoretical peak electrical power output of 288 MW. The turbines also have a much higher cutout speed compared to the turbines of the other two wind farms.

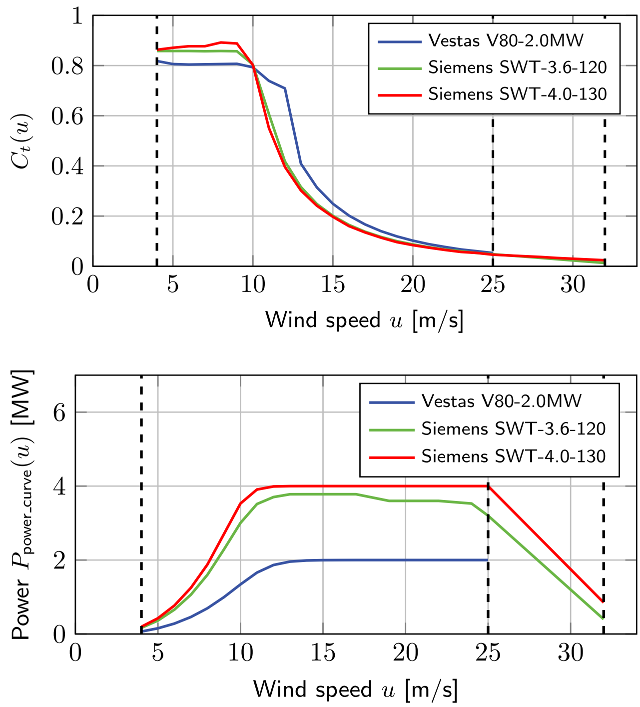


Figure 13. Thrust coefficient C_t and power production of the turbines Vestas V80 by L. Jensen et al. (2004), Siemens SWT-3.6–120, and Siemens SWT-4.0–130. The dashed vertical lines show the cut-in and cut-out speeds.

the expectation and variance with respect to a quasi-Monte Carlo simulation computed with a sufficiently large sample size of $M = 1 \cdot 10^8$. Normally, the classical Monte Carlo simulation should be used as a reference as it will always converge toward the correct result for $M \to \infty$, but the high dimensionality of the problem combined with the slow convergence rate of the method demanded an unreasonable sample size.

The presented results show that Monte Carlo is the least accurate method in terms of the observed relative error across all sample sizes which is to be expected from its theoretical convergence rate. The sparse grid Stochastic Collocation method shows a higher convergence rate compared to the quasi-Monte Carlo method, despite having a higher initial error for lower grid levels. Both methods achieve similar relative errors of $7 \cdot 10^{-7}$ for the expectation and $9 \cdot 10^{-5}$ for the variance with respect to the reference quasi-Monte Carlo solution. The dashed lines in Figures 10 and 11 show the theoretical convergence rates of the classical Monte Carlo and the quasi-Monte Carlo methods. While the classical Monte Carlo method shows the expected convergence rate, it is worth noting that the quasi-Monte Carlo method, shows a comparable error decline despite the much higher dimensionality of the problem. This implies that the probability space of the underlying problem might be dominated by a single dimension and that the low discrepancy sequence of the method can successfully exploit this fact for an accelerated convergence behavior. This observation becomes even more apparent in the sensitivity analysis presented in Section 5.

Considering the computational costs of both methods, the cost of computing the sparse grid for the Stochastic Collocation method scales nonlinearly with the dimension of the probability space and the grid level, while the sequence generation of the quasi-Monte Carlo method is linear in the number of sample points and thus potentially faster for large numbers of samples/dimensions. As the grid calculations could also be done once and stored prior to the actual Uncertainty Quantification analysis, this does not pose a significant drawback.

Based on the obtained results, it can be concluded that the quasi-Monte Carlo method is favorable over the other methods. This verdict is based on the three criteria: error control, computational overhead, and ease of implementation. While Stochastic Collocation and quasi-Monte Carlo yield comparable results for the demonstrated sample sizes, it has to be stressed that the reference solution has also been computed using the quasi-Monte Carlo method and the relative error thus has a bias in favor of quasi-Monte Carlo. Purely based on the observed relative convergence rate, the Stochastic Collocation method converges at a higher rate. Quasi-Monte Carlo allows for an arbitrary sample size and thus a fine-grained error control, while the used sparse grids in the Stochastic Collocation method only allow discrete sample sizes, which exponentially increase with the selected level (see marker in the Stochastic Collocation plots in Figures 10 and 11) due to the inherent nestedness of quadrature nodes. Furthermore, quasi-Monte Carlo does not required any precomputations, as computing the next sample from the Sobol sequence has negligible computational cost. Lastly, numerical codes for the Sobol sequence have better availability across multiple programming languages compared to arbitrary dimension and level sparse grid generators.

The presented results in Section 5 are therefore solely obtained by the application of the quasi-Monte Carlo method. For the sample size, we choose $1 \cdot 10^6$ samples as the obtained results show a relative error of $\mathcal{O}(10^{-6})$ for the LCOE, which also includes the computation of the AEP (see Figure 1) and thus errors for the mean should be of the order $\mathcal{O}(1)$ [GWh] for all presented test cases.

5. Case study results

This section describes the evaluation of the stochastic models introduced in Section 3 by using the methods of Uncertainty Quantification from Section 4. The influence of the input variables on the computed annual energy production and levelized cost of electricity will be determined and their computed statistics will be analyzed. This offline data analysis provides a better insight on the impact of each uncertain parameter and will help engineers with more accurate estimates of the expected annual energy production's 90%-quantile during the design stage of a wind farm.

For the model, we will use the parameters of the three wind farms *Horns Rev 1*,[2] *DanTysk*,[3] and *Sandbank*.[4] As each of these wind farms differs significantly in the used turbine types and grid layout, they will also have different losses and sensitivities with respect to the presented uncertainties. Because the wind farms are all closely positioned on a global scale, the same wind data will be used in each of the evaluations. The source of this wind data is the FINO3[5] met mast, positioned about 50 km southwest of Horns Rev, 2 km west from DanTysk, and 20 km east from Sandbank. We use all data recorded between January 2010 and December 2017 at a height of 100 meters.

For each wind farm, turbine positions are given in Figure 12, its turbine C_t and power curves are shown in Figure 13, while its general settings, losses, and uncertainties are listed in Tables 3 and 4. The values presented in these tables originate from data which was gathered over multiple years by the Vattenfall Europe Windkraft GmbH.

Before analyzing the propagation of uncertainties, we introduce some of the notation used by the variance-based sensitivity analysis for the joint probability space spanned by the uncertainties w.r.t. all nine variables given in Table 1 as described by Saltelli et al. (2010). This notation is especially suitable to express how the uncertainty in the model output can be linked to uncertainties in the inputs. The stochastic model derived previously will be represented by Y in the following, while the model inputs will be referred to as ξ_i. The uncertainty propagation of ξ_i through Y is best described in the context of this notation by the sensitivity measure by Saltelli et al. (2010)

[2]More information: https://powerplants.vattenfall.com/horns-rev
[3]https://powerplants.vattenfall.com/dantysk
[4]https://powerplants.vattenfall.com/sandbank
[5]https://www.fino3.de/en/

Table 3. Simulation settings for the wind farms Horns Rev 1, DanTysk, and Sandbank.

Parameter		Horns Rev 1	DanTysk	Sandbank
Positions		see Figure 12	see Figure 12	see Figure 13
Wind data		FINO3 (2010–2017)	FINO3 (2010–2017)	FINO3 (2010–2017)
Wind speed losses	ℓ_{wind}	98.5%	99.2%	99.5%
Turbine type		Vestas V80-2.0 MW	Siemens SWT-3.6–120	Siemens SWT-4.0–130
Rotor diameter	D	80 m	120 m	130 m
Hub height	z	70 m	88 m	95 m
Cut-in speed	u_{cutin}	4 m/s	4 m/s	4 m/s
Cutout speed	u_{cutout}	25 m/s	32 m/s	32 m/s
Surface roughness	z_0	$0.2 \cdot 10^{-3}$ m	$0.2 \cdot 10^{-3}$ m	$0.2 \cdot 10^{-3}$ m
Wake effect losses	ℓ_{wake}	99.9%	99.9%	99.9%
Power curve	$P(u)$	see Figure 13	see Figure 13	see Figure 13
C_t curve	$C_t(u)$	see Figure 13	see Figure 13	see Figure 13
Power curve losses	ℓ_{power}	98.8%	98.9%	99%
Total capital costs	C_{capital}	278 000 000 €	1 000 000 000 €	1 200 000 000 €
Annual operation and maintenance costs	C_{OM}	52 000 000 €	43 200 000 €	43 200 000 €
Discount rate	r_{rate}	2%	3.75%	0.15%
Project lifetime	T	20 years	20 years	20 years
Plant performance losses	$\ell_{\text{performance}}$	88.5%	88.8%	89%

Table 4. Uncertain parameters for the wind farms Horns Rev 1, DanTysk, and Sandbank. The surface roughness parameter is given by using the maximum approximation of Foti, Yang, and Sotiropoulos (2017), $\sigma_{rough} = 1.5 \cdot 10^{-5}/z$. The values presented in this table originate from data which was gathered over multiple years by the Vattenfall Europe Windkraft GmbH. These unpublished values base on recorded data of several existing wind farms.

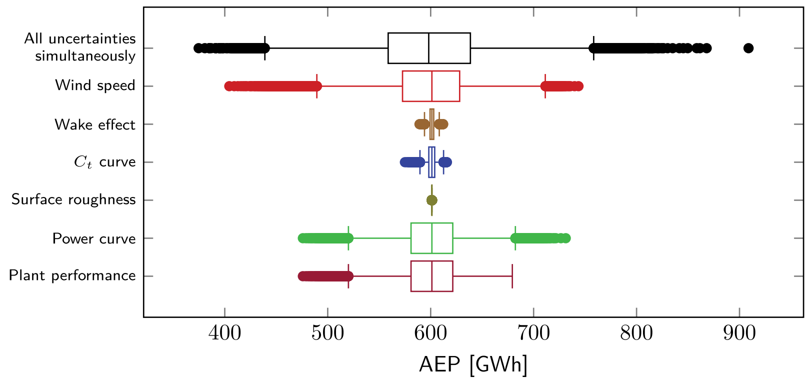
Uncertainty		Horns Rev 1		DanTysk		Sandbank	
		deviation	σ	deviation	σ	deviation	σ
Wind speed	ξ_{wind}	8.0%	0.0486	6.0%	0.0365	7.5%	0.0456
Wake effect	ξ_{wake}	3.0%	0.0182	3.0%	0.0182	3.0%	0.0182
C_t curve	ξ_{ct}	2.0%	0.0122	2.0%	0.0122	2.0%	0.0122
Surface roughness	ξ_{rough}	0.000021%	$1.277 \cdot 10^{-7}$	0.000017%	$1.034 \cdot 10^{-7}$	0.000016%	$9.737 \cdot 10^{-8}$
Power curve	ξ_{power}	2.0%	0.0122	2.0%	0.0122	2.0%	0.0122
Plant performance	$\xi_{\text{performance}}$	2.3%	0.0140	2.1%	0.0128	2.1%	0.0128
Capital costs	ξ_{capital}	3.0%	0.0182	3.0%	0.0182	3.0%	0.0182
Annual O&M costs	ξ_{om}	3.0%	0.0182	3.0%	0.0182	3.0%	0.0182
Discount rate	ξ_{rate}	4.0%	0.0234	1.0%	0.0061	0.5%	0.0030

$$S_i = \frac{\mathrm{var}_{\xi_i}(E_{\xi_{\sim i}}(\mathrm{Y}|\xi_i))}{\mathrm{var}(\mathrm{Y})}, \tag{15}$$

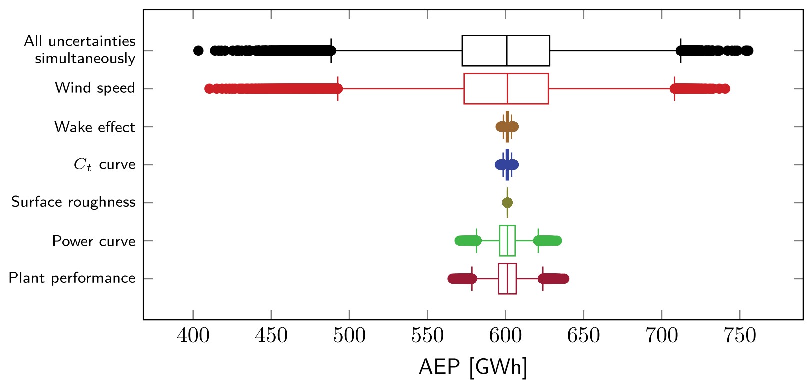
which is the first-order sensitivity coefficient that, for example, measures the additive effect of ξ_i on the model output. S_i can also be interpreted in terms of expected reduction of variance. This interpretation allows an easier understanding of the factors involved in the computation of S_i:

- $\mathrm{var}(\mathrm{Y})$: output variance with all inputs modeled as random variables
- $\mathrm{var}_{\xi_i}(E_{\xi_{\sim i}}(\mathrm{Y}|\xi_i))$: expected reduction in variance that would be obtained if ξ_i could be fixed

In the following, we use box plots to visualize statistical characteristics of the solution. The inside of the box, bounded by the lower and upper quartiles, represents 50% of the computed samples. The vertical line inside the box stands for the median of the data set. The lower whisker is the smallest data value, which is larger than: $\mathrm{lowerquartile} - 1.5 \cdot \mathrm{inter} - \mathrm{quartile} - \mathrm{range}$, where the "inter–quartile–range" is the difference between the upper and lower quartiles. The upper whisker is the largest data value which is smaller than: $\mathrm{upperquartile} + 1.5 \cdot \mathrm{inter} - \mathrm{quartile} - \mathrm{range}$.



(a) Horns Rev 1 test case with equal deviations of 5 %.



(b) Horns Rev 1 test case with different deviations according to Table 4.

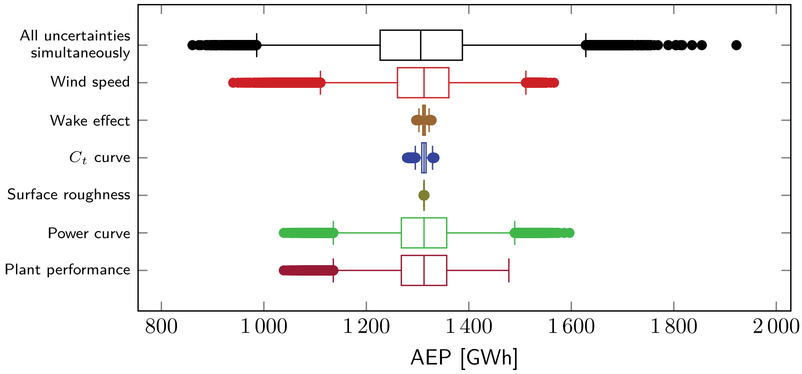
Figure 14. Graphical visualization of the sensitivity analysis with all six uncertainty parameters wind speed, wake effect, C_t curve, surface roughness, power curve, and plant performance at the same time (simultaneously) compared to single disturbances of the uncertainty parameters computed with the quasi-Monte Carlo method with 10^6 samples for the Horns Rev 1 wind farm. The mean AEP corresponds to the given values in literature, namely 599.5 GWh.

5.1. Annual energy production

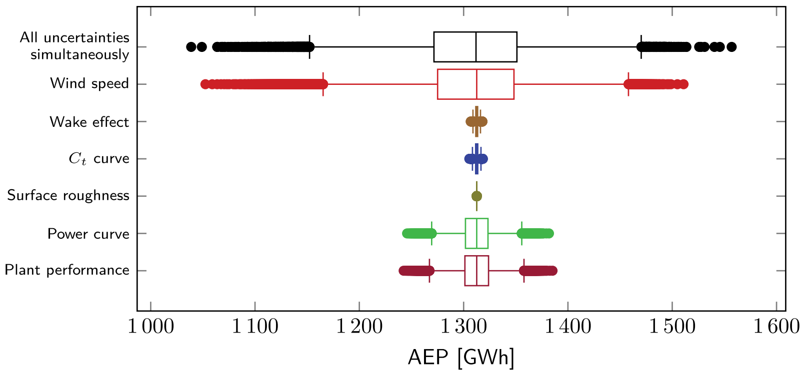
In the Figures 14, 15 and 16 the results regarding the annual energy production for all six influencing uncertainty parameters, that is, wind speed, wake effect, C_t curve, surface roughness, power curve, and power loss are shown. For their deviation, we choose two test cases: an *academic test case* for which we choose equal deviations of 5% for all six uncertainty-inflicted parameters, while for the *realistic test case* we choose the deviations as given in Table 4 by the Vattenfall Europe Windkraft GmbH.

Each box plot represents the results of an uncertain parameter. The topmost box plot represents the model output for the case that the six uncertainties are disturbed at the same time. In the box plot, it can be seen that the deviation from the expected annual energy production value behaves similarly in both directions of the AEP. The deviations of the wind speed, power curve, and plant performance have the largest impact on the annual energy production in comparison to all other uncertainties. This holds for the academic and realistic test case.

The results in terms of the variance-based sensitivity analysis are shown in Table 5 (academic test case) and Table 6 (realistic test case). In the realistic test case, for all three wind farms the sensitivity is higher than 84%. The sensitivities of the power curve and plant performance follow with about 3 to 9%. All remaining uncertainties (wake effect, C_t curve, and surface roughness) are below 0.3% and thus have a rather negligible influence on the variance of the model output.



(a) DanTysk test case with equal deviations of 5 %.



(b) DanTysk test case with different deviations according to Table 4.

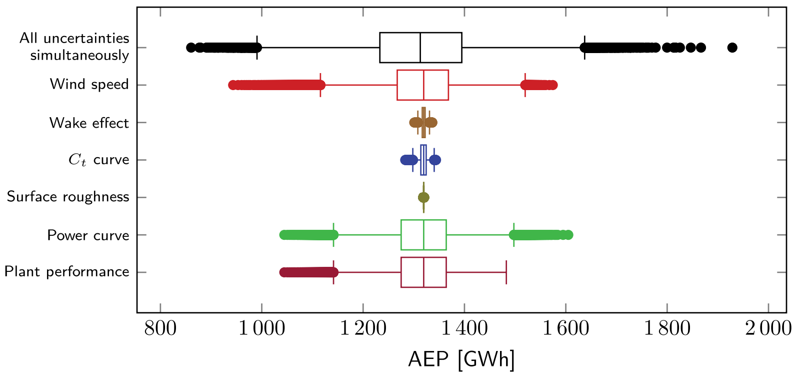
Figure 15. Graphical visualization of the sensitivity analysis with all six uncertainty parameters wind speed, wake effect, C_t curve, surface roughness, power curve, and plant performance at the same time (simultaneously) compared to single disturbances of the uncertainty parameters computed with the quasi-Monte Carlo method with 10^6 samples for the DanTysk wind farm. The mean AEP corresponds to the given values in literature, namely 1310.6 GWh.

5.2. Levelized costs of electricity

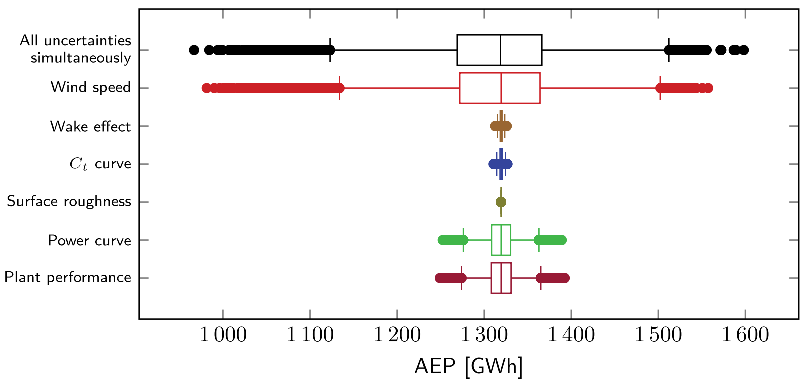
In Figures 17, 18 and 19 the results regarding the levelized costs of electricity for all nine influencing uncertainty parameters are shown, according to an equal deviation of 5% (*academic test case*) and according to the deviations given in Table 4 (*realistic test case*).

As before, it can be seen that the uncertainty of wind speed causes the main impact. Another interesting result is that the outliers of the combined and wind speed box plot only spread in the direction of a higher LCOE, which again speaks for the high sensitivity of the wind speed.

Tables 7 and 8 show the results in terms of the variance-based sensitivity analysis for the academic and realistic test case. Because the LCOE highly depends on the AEP, the uncertainties show a similar behavior as before. In the realistic test case, for all three wind farms the sensitivity is higher than 78%, while the sensitivities of the power curve and plant performance follow with about 3 to 8%. Furthermore, the sensitivities for the capital costs and the annual O&M costs are in-between 1.4 and 7.7%. All remaining uncertainties (wake effect, C_t curve, surface roughness, and discount rate) are below 1.2% and thus have a rather negligible influence on the variance of the model output.



(a) Sandbank test case with equal deviations of 5 %.



(b) Sandbank test case with different deviations according to Table 4.

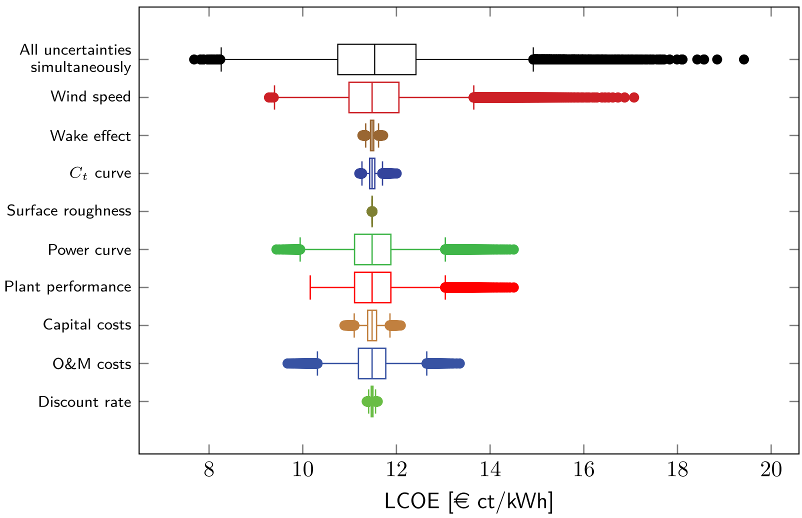
Figure 16. Graphical visualization of the sensitivity analysis with all six uncertainty parameters wind speed, wake effect, C_t curve, surface roughness, power curve, and plant performance at the same time (simultaneously) compared to single disturbances of the uncertainty parameters computed with the quasi-Monte Carlo method with 10^6 samples for the Sandbank wind farm. The mean AEP corresponds to the given values in literature, namely 1316.4 GWh.

Table 5. Results of the variance-based sensitivity analysis where all uncertainties are fixed to 5% of AEP calculation with 10^6 samples using the quasi-Monte Carlo Method.

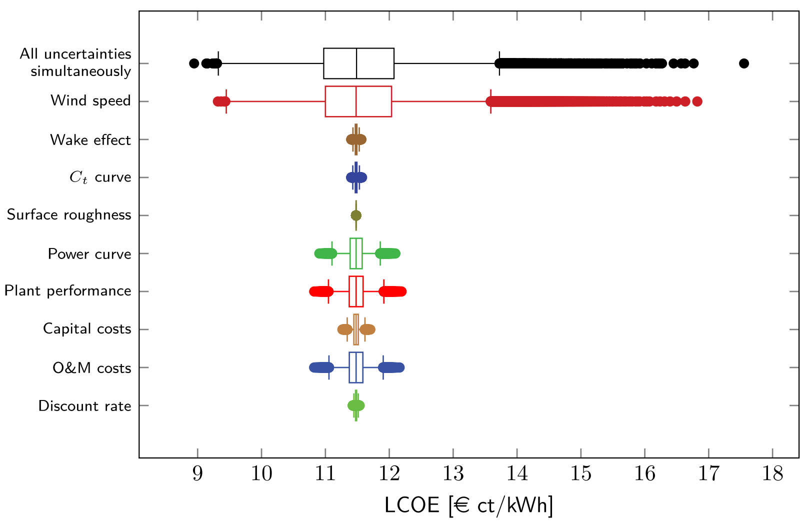
Uncertainty		Horns Rev 1		DanTysk		Sandbank	
		variance	sensitivity	variance	sensitivity	variance	sensitivity
Wind speed	ξ_{wind}	$1.6887 \cdot 10^9$	48.12%	$5.5619 \cdot 10^9$	39.3%	$5.6467 \cdot 10^9$	39.36%
Wake effect	ξ_{wake}	$1.1276 \cdot 10^7$	0.32%	$3.125 \cdot 10^7$	0.22%	$3.8256 \cdot 10^7$	0.27%
C_t curve	ξ_{ct}	$2.1235 \cdot 10^7$	0.61%	$5.1429 \cdot 10^7$	0.36%	$7.266 \cdot 10^7$	0.51%
Surface roughness	ξ_{rough}	$4.2003 \cdot 10^6$	0.12%	$1.7947 \cdot 10^7$	0.13%	$2.0629 \cdot 10^7$	0.14%
Power curve	ξ_{power}	$9.0764 \cdot 10^8$	25.86%	$4.3241 \cdot 10^9$	30.55%	$4.3729 \cdot 10^9$	30.48%
Plant performance	$\xi_{performance}$	$8.9968 \cdot 10^8$	25.63%	$4.2771 \cdot 10^9$	30.22%	$4.3179 \cdot 10^9$	30.1%

Table 6. Results of the variance-based sensitivity analysis of AEP calculation with 10^6 samples using the quasi-Monte Carlo Method.

Uncertainty		Horns Rev 1		DanTysk		Sandbank	
		variance	sensitivity	variance	sensitivity	variance	sensitivity
Wind speed	ξ_{wind}	$1.5966 \cdot 10^9$	92.81%	$2.9557 \cdot 10^9$	84.61%	$4.6948 \cdot 10^9$	89.7%
Wake effect	ξ_{wake}	$3.6257 \cdot 10^6$	0.21%	$5.7514 \cdot 10^6$	0.16%	$1.2469 \cdot 10^7$	0.24%
C_t curve	ξ_{ct}	$3.6569 \cdot 10^6$	0.21%	$6.0823 \cdot 10^6$	0.17%	$1.3101 \cdot 10^7$	0.25%
Surface roughness	ξ_{rough}	$2.7336 \cdot 10^6$	0.16%	$3.9914 \cdot 10^6$	0.11%	$1.0173 \cdot 10^7$	0.19%
Power curve	ξ_{power}	$5.6519 \cdot 10^7$	3.29%	$2.6036 \cdot 10^8$	7.45%	$2.6928 \cdot 10^8$	5.14%
Plant performance	$\xi_{performance}$	$7.3561 \cdot 10^7$	4.28%	$2.862 \cdot 10^8$	8.19%	$2.9539 \cdot 10^8$	5.64%



(a) Horns Rev 1 test case with equal deviations of 5 %.

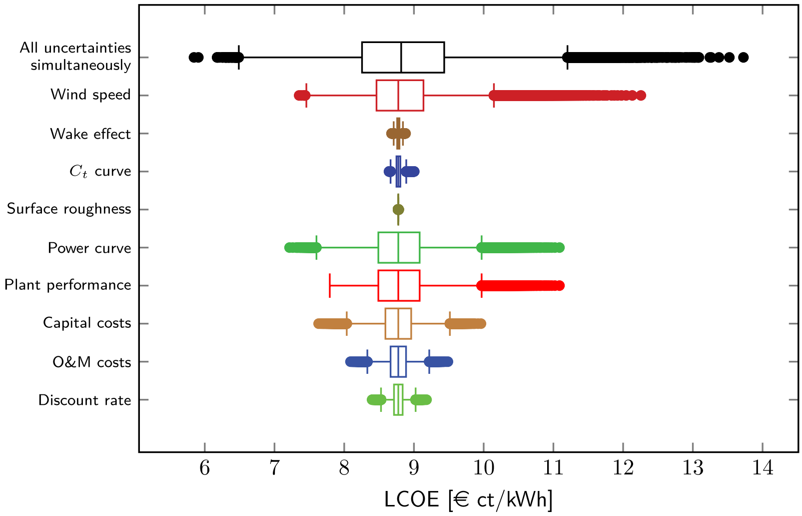


(b) Horns Rev 1 test case with different deviations according to Table 4.

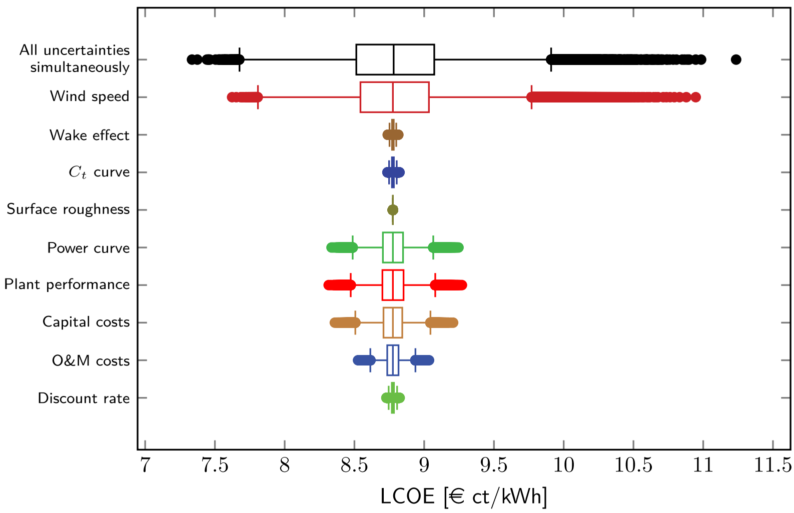
Figure 17. Graphical visualization of the sensitivity analysis with all nine uncertainty parameters wind speed, wake effect, C_t curve, surface roughness, power curve, plant performance, capital costs, O&M costs, and discount rate compared at the same time (simultaneously) to single effects of the uncertainty parameters computed with the quasi-Monte Carlo method with 10^6 samples for the Horns Rev 1 wind farm.

6. Discussion of the results

Analyzing the obtained the results, the wind speed has the largest impact with a sensitivity of at least 85% for the AEP and 78% for the LCOE. This insight in itself is not surprising and can be seen as more of a sanity check due to the wind being the governing factor in the power production. However, the Dan Tysk wind farm shows a roughly 8% lower sensitivity compared to Horns Rev 1. The obtained results are generally in line with the recent literature, but a comprehensive Uncertainty Quantification to this extent has not yet been done and the sensitivity measure by Saltelli et al. (2010) has also, despite its intuitive nature, not been used in the reviewed literature. This makes the results difficult to compare, even though the obtained results of the deterministic model were validated against



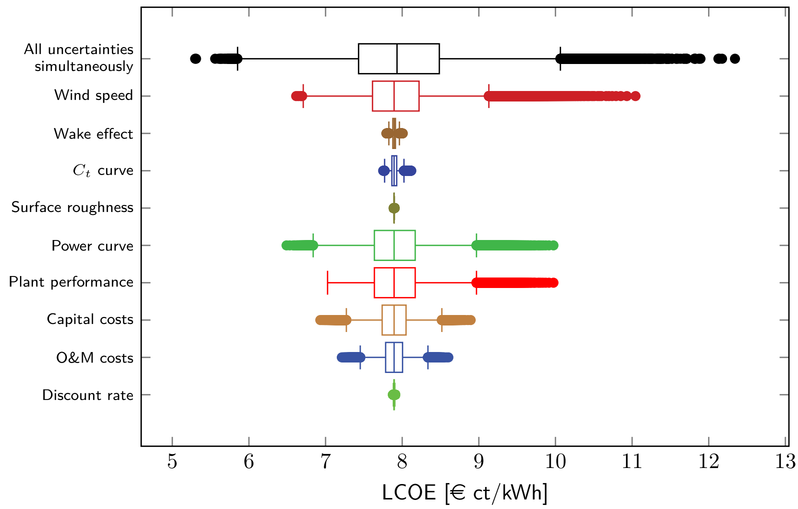
(a) DanTysk test case with equal deviations of 5 %.



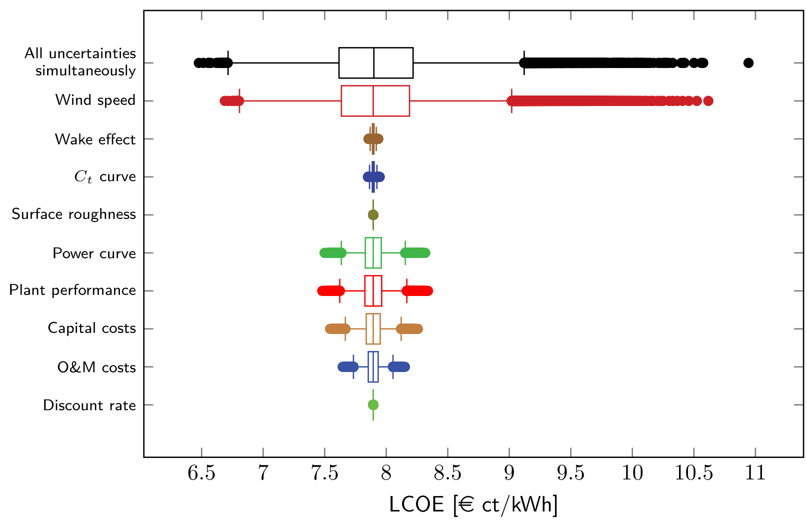
(b) DanTysk test case with different deviations according to Table 4.

Figure 18. Graphical visualization of the sensitivity analysis with all nine uncertainty parameters wind speed, wake effect, C_t curve, surface roughness, power curve, plant performance, capital costs, O&M costs, and discount rate compared at the same time (simultaneously) to single effects of the uncertainty parameters computed with the quasi-Monte Carlo method with 10^6 samples for the DanTysk wind farm.

Openwind and the Uncertainty Quantification methods have also been validated to compute correct statistics given enough samples. Therefore, these results reveal some key differences in the wind farm configurations. As the performance characteristics of all three turbine types are similar (see Figure 13) and the used database for wind speed and direction is the same, this observation likely is caused by the turbine placements. Where the layout of Horns Rev 1 resembles a slightly skewed square grid, Sandbank's turbines are also placed on a grid but elongated in northern direction. DanTysk on the other hand has a very similar layout compared to Sandbank, but the turbines are no longer placed on a fixed grid but their position are altered using numerical optimization methods to minimize wake obstruction.



(a) Sandbank test case with equal deviations of 5 %.



(b) Sandbank test case with different deviations according to Table 4.

Figure 19. Graphical visualization of the sensitivity analysis with all nine uncertainty parameters wind speed, wake effect, C_t curve, surface roughness, power curve, plant performance, capital costs, O&M costs, and discount rate compared at the same time (simultaneously) to single effects of the uncertainty parameters computed with the quasi-Monte Carlo method with 10^6 samples for the Sandbank wind farm.

Table 7. Results of variance-based sensitivity analysis where all uncertainties are fixed to 5% of LCOE calculation with 10^6 samples using the quasi-Monte Carlo Method.

Uncertainty		Horns Rev 1		DanTysk		Sandbank	
		variance	sensitivity	variance	sensitivity	variance	sensitivity
Wind speed	ξ_{wind}	$6.7186 \cdot 10^{-1}$	42.13%	$2.6737 \cdot 10^{-1}$	33.94%	$2.1755 \cdot 10^{-1}$	34.53%
Wake effect	ξ_{wake}	$2.6063 \cdot 10^{-2}$	1.63%	$1.1135 \cdot 10^{-2}$	1.41%	$9.4152 \cdot 10^{-3}$	1.49%
C_t curve	ξ_{ct}	$2.7456 \cdot 10^{-2}$	1.72%	$1.1165 \cdot 10^{-2}$	1.42%	$9.6657 \cdot 10^{-3}$	1.53%
Surface roughness	ξ_{rough}	$2.3825 \cdot 10^{-2}$	1.49%	$1.0576 \cdot 10^{-2}$	1.34%	$8.8553 \cdot 10^{-3}$	1.41%
Power curve	ξ_{power}	$3.4543 \cdot 10^{-1}$	21.66%	$1.9984 \cdot 10^{-1}$	25.37%	$1.6189 \cdot 10^{-1}$	25.69%
Plant performance	$\xi_{performance}$	$3.433 \cdot 10^{-1}$	21.53%	$1.9832 \cdot 10^{-1}$	25.18%	$1.6045 \cdot 10^{-1}$	25.47%
Capital costs	$\xi_{capital}$	$4.3827 \cdot 10^{-2}$	2.75%	$8.5743 \cdot 10^{-2}$	10.88%	$6.2208 \cdot 10^{-2}$	9.87%
Annual O&M costs	$\xi_{o\&m}$	$2.1092 \cdot 10^{-1}$	13.23%	$3.7667 \cdot 10^{-2}$	4.78%	$3.5661 \cdot 10^{-2}$	5.66%
Discount rate	ξ_{rate}	$2.4519 \cdot 10^{-2}$	1.54%	$1.8883 \cdot 10^{-2}$	2.4%	$8.8718 \cdot 10^{-3}$	1.41%

Table 8. Results of variance-based sensitivity analysis of LCOE calculation with 10^6 samples using the quasi-Monte Carlo Method.

Uncertainty		Horns Rev 1		DanTysk		Sandbank	
		variance	sensitivity	variance	sensitivity	variance	sensitivity
Wind speed	ξ_{wind}	$6.3759 \cdot 10^{-1}$	90.15%	$1.3795 \cdot 0^{-1}$	78.52%	$1.7992 \cdot 10^{-1}$	85.92%
Wake effect	ξ_{wake}	$8.1103 \cdot 10^{-3}$	1.15%	$1.0691 \cdot 10^{-3}$	0.61%	$1.9739 \cdot 10^{-3}$	0.94%
c_t curve	ξ_{ct}	$8.0655 \cdot 10^{-3}$	1.14%	$1.0721 \cdot 10^{-3}$	0.61%	$1.9698 \cdot 10^{-3}$	0.94%
Surface roughness	ξ_{rough}	$7.8208 \cdot 10^{-3}$	1.11%	$9.9291 \cdot 10^{-4}$	0.57%	$1.8982 \cdot 10^{-3}$	0.91%
Power curve	ξ_{power}	$2.6861 \cdot 10^{-2}$	3.8%	$1.2308 \cdot 10^{-2}$	7.01%	$1.0983 \cdot 10^{-2}$	5.24%
Plant performance	$\xi_{performance}$	$3.2909 \cdot 10^{-2}$	4.65%	$1.3451 \cdot 10^{-2}$	7.66%	$1.19 \cdot 10^{-2}$	5.68%
Capital costs	$\xi_{capital}$	$1.0469 \cdot 10^{-2}$	1.48%	$1.0952 \cdot 10^{-2}$	6.23%	$8.9661 \cdot 10^{-3}$	4.28%
Annual O&M costs	$\xi_{o\&m}$	$3.2609 \cdot 10^{-2}$	4.61%	$4.5816 \cdot 10^{-3}$	2.61%	$5.4487 \cdot 10^{-3}$	2.6%
Discount rate	ξ_{rate}	$7.9749 \cdot 10^{-3}$	1.13%	$1.1177 \cdot 10^{-3}$	0.64%	$1.8975 \cdot 10^{-3}$	0.91%

The uncertain parameters describing the wake effect, C_t curve, surface roughness and discount rate have a sensitivity below 1.2% across all results. Therefore, in further investigations, these values could be aggregated and modeled as a single uncertain parameter or even completely neglected if simulation time is an issue or one is only interested in changes of the AEP/LCOE uncertainty on a grand scale. The other sensitivities, namely power curve for the AEP and plant performance, capital costs and annual O&M costs for the LCOE, show sensitivities between 1.4 and 9%. Especially the economical parameters are greatly varying across the different wind farms. The capital costs for DanTysk have a sensitivity of 6.2% whereas Horns Rev 1 only has a sensitivity of 4.6%. The O&M costs, however, show the opposite behavior with Horns Rev 1 having a sensitivity value of 4.6% and DanTysk 2.6%. This observation again seems to be linked to the turbine positioning, but might require further research in order to gain certainty.

Whereas the results using a 5% standard deviation across all parameters show systematic sensitivities, the input parameters for the *real* deviations are based upon gathered data and experience from the Vattenfall Europe Windkraft GmbH and are thus subject to steady adjustments from internal experts and may require a new analysis provided updated estimates deviate significantly.

As the obtained results have been computed using a rather simple wake model to allow a comprehensive analysis of all major uncertain parameters in the resulting high dimensional stochastic space, more refined models should now be used to focus on subsets of these parameters based on this work. For future analysis, we recommend using quasi-Monte Carlo methods due to ease of implementation and fine-grained error control, but more advanced methods like the multilevel-Monte Carlo method can also be considered.

7. Conclusion

Within this work, we investigated the sensitivity of certain input parameters for the estimated economics of offshore wind farms. We developed a model describing the wind, wake, and power generation of an offshore wind farm which delivers an output of two economic target functions, tjat is, the AEP and the LCOE. These target functions are usually used for site evaluation and construction feasibility. We extended the deterministic model by introducing nine uncertain model parameters that are crucial in the computation of, for example, the produced power of an offshore wind farm. As this results in a high-dimensional probability space, requiring a numerically expensive integration, the full space has not been tested previously. Furthermore, the convergence of several numerical methods commonly used in Uncertaintainty Quantification with a focus on performance and ease of implementation. In future projects, more complicated methods, for example, multi-level Monte Carlo or intrusive methods of any kind could be investigated. Similarly, additional stochastic models based on different wake model or partial differential equations like the Navier-Stokes equations could be developed and used for an even more accurate sensitivity analysis similar to the one presented in this work. As the obtained results reveal that only four out of the nine investigated uncertainties have significant impact on the LCOE, more complex models could also be researched with reasonable computation effort without utilizing large HPC systems.

The presented case studies show similar behavior throughout all tested scenarios, but due to the wind data being the same on all three sites the respective layouts, turbines types and geological locations appear to make some wind farms more sensitive to specific parameters than others. This might also be interesting for future works, as a generally more robust wind farm configuration would greatly improve predictions with respect to average power production and financial quantities like the expected rate of return.

Acknowledgments

The authors thank the Federal Ministry for Economic Affairs and Energy BMWi (Bundesministerium für Wirtschaft und Energie) and the Project Executing Organization PTJ (Projektträger Jülich) for making the FINO3 data available. The authors thank Adam Verhoeven-Mrosek from Vattenfall Europe Windkraft GmbH for making offshore wind farm data available.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Pascal Richter http://orcid.org/0000-0002-0453-6780

References

Ainslie, J. F. 1988. Calculating the flowfield in the wake of wind turbines. *Journal of Wind Engineering and Industrial Aerodynamics* 27 (1–3):213–24. doi:10.1016/0167-6105(88)90037-2.

Ashuri, T., T. Zhang, D. Qian, and M. Rotea 2016. Uncertainty quantification of the levelized cost of energy for a 20 MW research wind turbine model. 34th Wind Energy Symposium. San Diego, USA, 1998. doi:10.2514/6.2016-1998.

Babuška, I., F. Nobile, and R. Tempone. 2007. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM Journal on Numerical Analysis* 45 (3):1005–34. doi:10.1137/050645142.

Barthelmie, R., G. Larsen, S. Frandsen, L. Folkerts, K. Rados, S. Pryor, B. Lange, and G. Schepers. 2006. Comparison of wake model simulations with offshore wind turbine wake profiles measured by sodar. *Journal of Atmospheric and Oceanic Technology* 23 (7):888–901. doi:10.1175/jtech1886.1.

Barthelmie, R., S. Frandsen, K. Hansen, J. Schepers, K. Rados, W. Schlez, and S. Neckelmann 2009. Modelling the impact of wakes on power output at Nysted and Horns Rev. European Wind Energy Conference. Marseille, France, 1–10.

Bratley, P., and B. L. Fox. 1988. Algorithm 659: Implementing Sobol's quasirandom sequence generator. *ACM Transactions on Mathematical Software (TOMS)* 14 (1):88–100. doi:10.1145/42288.214372.

Brower, M., and N. Robinson 2012. The Openwind deep-array wake model. Tech. Rep. 463, AWS Truepower, Albany, NY.

C¸ Akar, R. 2017. Uncertainty quantification of wind farm models. Bachelor Thesis, RWTH Aachen University.

Caflisch, R. E. 1998. Monte Carlo and quasi-Monte Carlo methods. *Acta Numerica* 7:1–49. doi:10.1017/s0962492900002804.

Carrillo, C., J. Cidrás, E. Díaz-Dorado, and A. F. Obando-Montaño. 2014. An approach to determine the Weibull parameters for wind energy analysis: The case of Galicia (Spain). *Energies* 7 (4):2676–700. doi:10.3390/en7042676.

de Cursi, E. S., and R. Sampaio. 2015. Uncertainty quantification and stochastic modeling with Matlab. *Elsevier*. doi:10.1016/c2014-0-04713-2.

DNV. 2013. Framework for the categorisation of losses and uncertainty for wind energy assessments. Tech. Rep., KEMA, Høvik, Bærum, Akershus, Norway.

Foti, D., X. Yang, and F. Sotiropoulos. 2017. Uncertainty quantification of infinite aligned wind farm performance using non-intrusive polynomial chaos and a distributed roughness model. *Wind Energy* 20 (6):945–58. doi:10.1002/we.2072.

Gaumond, M., P.-E. Réthoré, S. Ott, A. Pena, A. Bechmann, and K. S. Hansen. 2014. Evaluation of the wind direction uncertainty and its impact on wake modeling at the Horns Rev offshore wind farm. *Wind Energy* 17 (8):1169–78. doi:10.1002/we.1625.

Genc, A., M. Erisoglu, A. Pekgor, G. Oturanc, A. Hepbasli, and K. Ulgen. July, 2005. Estimation of wind power potential using Weibull distribution. *Energy Sources* 27 (9):809–22. doi:10.1080/00908310490450647.

Ghanem, R., D. Higdon, and H. Owhadi. 2017. *Handbook of uncertainty quantification*. Switzerland: Springer International Publishing. doi:10.1007/978-3-319-11259-6.

González, J. S., A. G. Rodríguez, J. C. Mora, J. R. Santos, and M. B. Payán 2009. A new tool for wind farm optimal design. PowerTech. Bucharest, Romania, 1–7. doi:10.1109/ptc.2009.5281977.

Heiming, G. 2015. Modeling and simulation of offshore wind farms. Bachelor Thesis, RWTH Aachen University.

Ioannou, A., A. Angus, and F. Brennan. May, 2018. Parametric CAPEX, OPEX, and LCOE expressions for offshore wind farms based on global deployment parameters. *Energy Sources, Part B: Economics, Planning, and Policy* 13 (5):281–90. doi:10.1080/15567249.2018.1461150.

Jensen, L., C. Mørch, P. Sørensen, and K. Svendsen 2004. Wake measurements from the Horns Rev wind farm. European Wind Energy Conference. London, UK, 1–9.

Jensen, N. O. 1983. A note on wind generator interaction. No. 2411, Risø National Laboratory, Rosklide.

Katic, I., J. Højstrup, and N. O. Jensen 1986. A simple model for cluster efficiency. European Wind Energy Association Conference and Exhibition. Rome, Italy, 407–10.

Lackner, M. A., and C. N. Elkinton. 2007. An analytical framework for offshore wind farm layout optimization. *Wind Engineering* 31 (1):17–31. doi:10.1260/030952407780811401.

Lange, B., H. -P. Waldl, A. G. Guerrero, D. Heinemann, and R. J. Barthelmie. 2003. Modelling of offshore wind turbine wakes with the wind farm program FLaP. *Wind Energy* 6 (1):87–104. doi:10.1002/we.84.

Lange, B., S. Larsen, J. Højstrup, and R. Barthelmie. 2004. Importance of thermal effects and sea surface roughness for offshore wind resource assessment. *Journal of Wind Engineering and Industrial Aerodynamics* 92 (11):959–88. doi:10.1016/j.jweia.2004.05.005.

Larsen, G. C., H. A. Madsen, F. Bingöl, J. Mann, S. Ott, J. N. Sørensen, Okulov, V., Troldborg, N., Nielsen, N. M., and K. Thomsen 2007. *Dynamic wake meandering modeling*. Roskilde, Denmark: Risø National Laboratory, Technical University of Denmark.

Li, S., and L. Caracoglia 2020. Wind tunnel experimental variability of aerodynamic loads for wind turbine blades. *Journal of Physics: Conference Series*, Vol. 1452, 012054. doi:10.1088/1742-6596/1452/1/012054.

Matsumoto, M., and T. Nishimura. 1998. Mersenne twister: A 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Transactions on Modeling and Computer Simulation (TOMACS)* 8 (1):3–30. doi:10.1145/272991.272995.

Morokoff, W. J., and R. E. Caflisch. 1995. Quasi-Monte Carlo integration. *Journal of Com- Putational Physics* 122 (2):218–30. doi:10.1006/jcph.1995.1209.

Murcia, J. P., P.-E. Réthoré, K. S. Hansen, A. Natarajan, and J. D. Sørensen (2015). A new method to estimate the uncertainty of AEP of offshore wind power plants applied to Horns Rev 1. EWEA Annual Conference and Exhibition Paris, France, 161–65.

Niederreiter, H. 1992. *Random number generation and quasi-Monte Carlo methods*. Philadelphia, USA: Siam. doi:10.1137/1.9781611970081.

Ott, S. (2009). Fast linearized models for wind turbine wakes.Euromech Colloquium Madrid, Spain , 11–13.

Padrón, A., A. P. Stanley, J. Thomas, J. Alonso, and A. Ning. 2016. Polynomial chaos to efficiently compute the annual energy production in wind farm layout optimization. *Journal of Physics: Conference Series* 753:032021. doi:10.5194/wes-4-211-2019.

Pillai, A., J. Chick, and V. Laleu 2014. Modelling wind turbine wakes at Middelgrunden wind farm. European Wind Energy Conference and Exhibition. Barcelona, Spain, 1–10.

Rinker, J. M. 2016. Calculating the sensitivity of wind turbine loads to wind inputs using response surfaces. *Journal of Physics: Conference Series* 753:032057. doi:10.1088/1742-6596/753/3/032057.

Saltelli, A., P. Annoni, I. Azzini, F. Campolongo, M. Ratto, and S. Tarantola. 2010. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications* 181 (2):259–70. doi:10.1016/j.cpc.2009.09.018.

Sarathkumar, T. V., A. Banik, A. K. Goswami, S. Dey, A. Chatterjee, S. Rakshit, . . . J. Saloi. September, 2019. Uncertainty borne balancing cost modeling for wind power forecast- ing. *Energy Sources, Part B: Economics, Planning, and Policy* 14 (7–9):291–303. doi:10.1080/15567249.2019.1693664.

Schepers, J. 2003. *ENDOW: Validation and improvement of ECN's wake model*. Petten: Energy research Centre of the Netherlands ECN.

Smith, R. C. 2013. *Uncertainty quantification: Theory, implementation, and applications*. Philadelphia, USA: Siam.

Stovall, T., G. Pawlas, and P. Moriarty 2010. Wind farm wake simulations in OpenFOAM. 48th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition. Orlando, USA, 825. doi:10.2514/6.2010-825.

Sullivan, T. J. 2015. *Introduction to uncertainty quantification*. Switzerland: Springer International Publishing. doi:10.1007/978-3-319-23395-6.

Tuzuner, A., and Z. Yu 2008. A theoretical analysis on parameter estimation for the Weibull wind speed distribution. Power and energy society general meeting- conversion and delivery of electrical energy in the 21st century, 2008 IEEE. Pittsburgh, USA, 1–6). doi:10.1109/pes.2008.4596167

van Wyk, H.-W., M. D. Gunzburger, and J. Burkardt. 2016. Clenshaw-Curtis Type Rules for Statistical Integrals.

Wiernga, J. 1993. Representative roughness parameters for homogeneous terrain. *Boundary- Layer Meteorology* 63 (4):323–63. doi:10.1007/bf00705357.

Wolters, J. 2016. Uncertainty quantification of wind farm models. Master Thesis, RWTH Aachen University.