Center for Computational Engineering Science Mathematics Division

RWTH AACHEN UNIVERSITY

**Diese Arbeit wurde vorgelegt am**
**Lehrstuhl für Mathematik (MathCCES)**

# Uncertainty Quantification für Windpark Modelle

## Uncertainty quantification of wind farm models

Masterarbeit
Computational Engineering Science

November 2016


| | |
|---|---|
| Vorgelegt von<br>Presented by | Jannick Wolters<br>Im Johannistal 5, 52064 Aachen<br>Matrikelnummer: 302517<br>jannick.wolters@rwth-aachen.de |
| Erstprüfer<br>First examiner | Prof. Dr. Martin Frank<br>Lehrstuhl für Mathematik (MathCCES)<br>RWTH Aachen University |
| Zweitprüfer<br>Second examiner | Prof. Dr. Erika Ábrahám<br>Lehr- und Forschungsgebiet: Theorie der hybriden Systeme<br>RWTH Aachen University |
| Koreferent<br>Co-supervisor | Dipl.-Math. Dipl.-Inform. Pascal Richter<br>Lehrstuhl für Mathematik (MathCCES)<br>RWTH Aachen University |

# Eigenständigkeitserklärung

Hiermit versichere ich, dass ich diese Masterarbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Die Stellen meiner Arbeit, die dem Wortlaut oder dem Sinn nach anderen Werken entnommen sind, habe ich in jedem Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht. Dasselbe gilt sinngemäß für Tabellen und Abbildungen. Diese Arbeit hat in dieser oder einer ähnlichen Form noch nicht im Rahmen einer anderen Prüfung vorgelegen.

Aachen, im November 2016

Jannick Wolters

# Contents

**Nomenclature** V

**Acronyms** VI

**List of Figures** VII

**List of Tables** VIII

**1. Introduction** 1

**2. Problem description** 2
2.1. Turbine properties . . . 2
2.2. Wind data . . . 3
2.3. Wake modeling . . . 3
2.3.1. PARK model . . . 4
2.3.2. Eddy-viscosity model . . . 7
2.4. Annual energy production . . . 9

**3. State of the art** 10

**4. Uncertainty Modeling** 11
4.1. Basic concepts of probability . . . 11
4.1.1. Random variables . . . 11
4.1.2. multiple RVs . . . 12
4.1.3. Statistics . . . 12
4.2. Generalized polynomial chaos . . . 12
4.3. Stochastic Spectral Methods . . . 13
4.3.1. Monte Carlo Sampling . . . 13
4.3.2. Stochastic Collocation . . . 14
4.4. Smolyak Sparse grids . . . 16

**5. Verification** 20

**6. Results** 24
6.1. Evaluation . . . 24
6.1.1. Wind speed . . . 24
6.1.2. Power curve . . . 25
6.1.3. Thrust curve . . . 25
6.2. Runtime analysis . . . 26

**7. Conclusion** 28
7.1. Summary . . . 28
7.2. Outlook . . . 28

**A. Simulation code input parameters** **29**

**References** **31**

# Nomenclature

## Symbols

| | | |
|---|---|---|
| $\beta_k$ | Shadowing factor | |
| $C_t$ | Thrust coefficient | |
| $D$ | Diameter of the turbine rotor | [m] |
| $\delta u$ | Velocity deficit | |
| $\varepsilon$ | Eddy viscosity | [$m^2/s$] |
| $I_a$ | Ambient turbulence intensity | |
| $I_{tot}$ | Total turbulence intensity | |
| $I_w$ | Wake-induced turbulence intensity | |
| $\kappa$ | Von Kármán constant | |
| $k$ | Wake decay constant | |
| $K_M$ | Eddy diffusivity | [$m^2/s$] |
| $r$ | Radial distance, radius | [m] |
| $\rho$ | Air density | [$kg/m^3$] |
| $u$ | Velocity in downstream direction | [m/s] |
| $u_0$ | Free stream velocity | [m/s] |
| $v$ | Velocity in radial direction | [m/s] |
| $x$ | Downstream distance | [m] |
| $z$ | Hub height of turbine | [m] |
| $z_0$ | Surface roughness of the site | [m] |

## Subscripts

| | |
|---|---|
| 0 | Undisturbed, concerning the site |
| $a$ | Ambient |
| $c$ | Centerline |
| $inc$ | Incident |
| $n$ | Near wake |
| $r$ | Directly behind rotor |
| $w$ | Wake |

# Acronyms

| | |
|---|---|
| **AEP** | Annual energy production |
| **BC** | Boundary condition |
| **CC** | Clenshaw Curtis |
| **CDF** | Cumulative distribution function |
| **gPC** | Generalized polynomial chaos |
| **HAWT** | Horizontal-axis wind turbine |
| **i.i.d.** | Independet and indentical distributed |
| **IC** | Initial condition |
| **MC** | Monte Carlo |
| **PDE** | Partial differential equation |
| **PDF** | Probability density function |
| **QoI** | Quantity of interest |
| **RK** | Runge-Kutta |
| **RMSE** | Root-mean-squared error |
| **RV** | Random variable |
| **SC** | Stochastic collocation |
| **SG** | Stochastic Galerkin |
| **UQ** | Uncertainty quantification |
| **WFLOP** | Wind farm layout optimization problem |

# List of Figures


1. Turbine measurements . . . . . 2
2. Enercon E-82 characteristics, from Openwind [13] . . . . . 3
3. Wind direction distribution at FINO3 for the years 2011 to 2014 . . . . 4
4. Comparison of annual wind speed distributions . . . . . 5
5. Visualization of the wake as described in PARK . . . . . 6
6. Original filter function, see equation (20) . . . . . 8
7. Visualization of different Smolyak sparse grids with nested Clenshaw Curtis nodes . . . . . 19
8. Solutions of the Burgers equation (red) with small differences in the initial condition (IC) (blue) . . . . . 20
9. $\Delta$ between both methods with increasing $n_{MC}$ . . . . . 22
10. Results for the Monte Carlo (MC) method . . . . . 23
11. Results for the stochastic collocation (SC) method . . . . . 23
12. Wind speed: statistics for different sample sizes/orders; MC(red), SC(blue) 25
13. Power curve: statistics for different sample sizes/orders; MC(red), SC(blue) 26
14. $C_t$ curve: statistics for different sample sizes/orders; MC(red), SC(blue) 27

# List of Tables


1. Probability density function (PDF) and matching basis polynomials . . 13
2. Product of monomials up to R=4 . . . . . 17
3. Number of Clenshaw-Curtis nodes for the sparse and tensor grids . . . 18
4. Parameters used in the computation . . . . . 21
5. Stochastic parameters . . . . . 21
6. Runtime comparison between MC and SC . . . . . 26
7. Input parameters for the simulation code . . . . . 30

# 1. Introduction

Renewable energy sources have become one of the most used generators for electrical energy providing 19.2% [11] of humanities global energy consumption. Compared to fossil and nuclear sources, they have the advantage of being environmentally friendly and unlimited as the providing sources are based on earths steady or periodical natural phenomena.
The most common provider of renewable energies are the solar and wind energy, but other sources as e.g. tides and geothermal heat are also very common. Solar energy though, if compared to wind energy, has an obvious disadvantage. It is not available at night and not equally efficient over the whole globe due to the earths axis orientation and its orbit around the sun. Wind on the other hand is a more steady resource, where these drawbacks are not existent and energy production is possible throughout the entire day.
Wind energy in general is harvested by turbines, which convert the kinetic energy of the free flowing wind into electrical current. After grid specific transformations, this current can directly be provided into the power grid. No additional transformations are needed and the efficiency is not diminished any further. The turbines are usually clustered into wind farms as this minimizes the costs per turbine. These wind farms can either be built on- or offshore, each having its own advantages and disadvantages. Here the focus is laid on offshore wind farms, which profit from unrestricted air flow and only occupy widely unused space, but suffer from higher construction costs.
Due to the dense clustering, the turbines influence each others performance, as wind passing one turbine creates a wake behind it from which less energy can be generated. Therefore it is desirable to choose an optimal positioning that minimizes this effect and thus maximizes the farms energy output.
For the solution of this so called 'wind farm layout optimization problem (WFLOP)', measurement data is required to gain information about different unknown parameters such as weather data at the construction ground and turbine properties. Even though this measurement data is provided, these measurements are always subject to uncertainties from e.g. the inaccuracies of measurement processes themselves. Depending on the physical process of the wake generation, these uncertainties can possibly have a big impact on the quality of final prediction of the optimal turbine positioning. Therefore modeling these uncertainties and tracing them through the wake models provides a better understanding to the WFLOP problem and primarily also to the statistics of the generated power output.
In this thesis, the propagation of uncertainties through two different wake models is examined. The parameters considered as uncertain are the wind speed, power curve and thrust curve, where the latter two are a characteristic parameter of the turbine. To propagate and evaluate these uncertainties, stochastic spectral methods of uncertainty quantification (UQ) are used to determine the stochastics of the quantity of interest (QoI), namely the annual energy production (AEP).

# 2. Problem description

This section explains the WFLOP problem, covering the most important details.
The first part is about the setup of the offshore wind park and more specifically about the turbine properties. This is followed by a description of the weather data used for the simulation. Knowledge about both of the first parts is required as the turbines thrust and power curve as well as wind speed will be considered uncertain in the later sections. Models for the wake generation as well as for the calculation of the AEP are described afterwards to wrap up the overview of problem modeling. The complete modeling process is based on the work of Heiming [4].

## 2.1. Turbine properties

Wind turbines can be categorized into different types, where the most common type is a 3-bladed horizontal-axis wind turbine (HAWT). For this type of turbine, propulsion is mainly generated from the lift force perpendicular to the blades. As the HAWT turbine is the most common type, it is the only one considered in the modeling process.

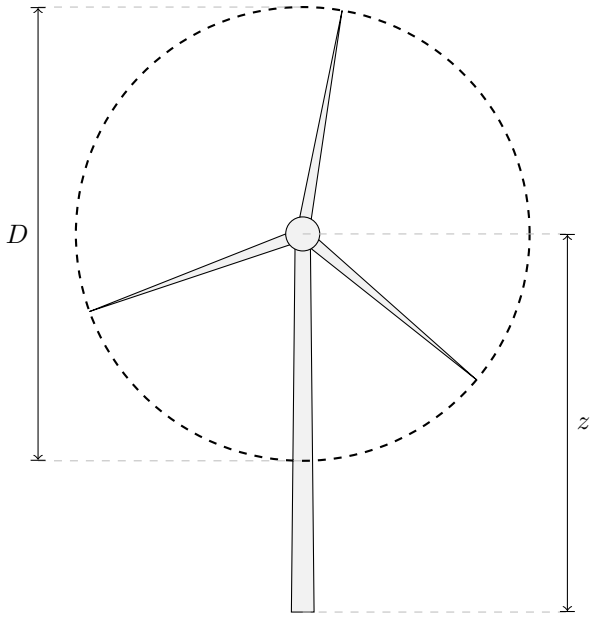
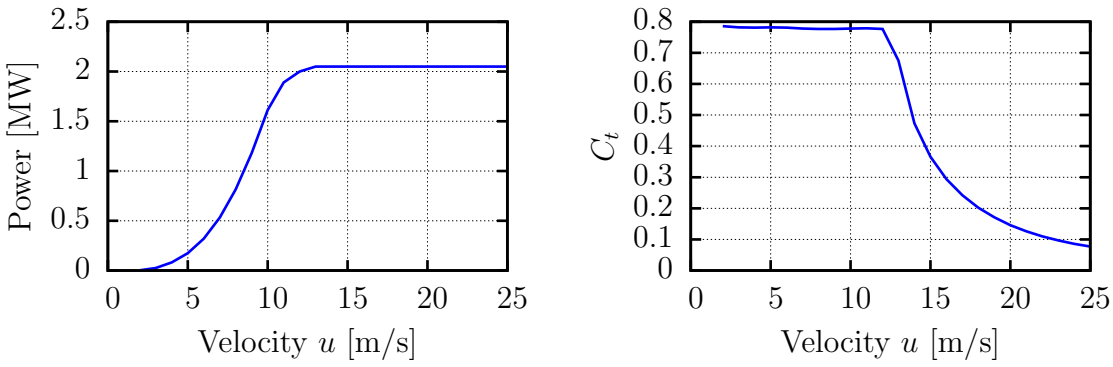



Figure 1: Turbine measurements

The general HAWT can be characterized by the hub height $z$ and the rotor diameter $D$ (see figure 1). Additionally the cut-in speed $u_{cutin}$ and cut-out speed $u_{cutout}$ are of interest, as they determine the range of wind speeds for which the turbine will be generating energy efficiently and without risking damage.
Furthermore, additional knowledge of the relations between the incident wind speed $u$ and the thrust curve $C_t$ as well as the turbines power output $P$ is needed. Usually these curves have a dependence on the airs density $\rho_{air}$, but nevertheless the density

is assumed constant in this work with $\rho_{air} = 1.225 \frac{\text{kg}}{\text{m}^3}$.
In figure 2 these curves can be seen for the example for the E-82 Enercon turbine, with $u_{cutin} = 2 \frac{\text{m}}{\text{s}}$ and $u_{cutout} = 25 \frac{\text{m}}{\text{s}}$. The thrust coefficient $C_t$ influences the overall thrust, that acts on the turbine rotor. Therefore it also is a measure for the velocity deficit at the rotor.




(a) Power curve

(b) $C_t$ curve

Figure 2: Enercon E-82 characteristics, from Openwind [13]

## 2.2. Wind data

The wind data used in this work has been gathered by the FINO3 research station, which is located at the german part of the north sea. The data recorded spans over the years 2011 to 2014 and provides roughly 190 thousand measurement points. A Weibull probability function has then been fitted to the data as described in [4]. The Weibull distribution has been used due to its good characteristics in approximating natural velocity distributions and also because of having a simple form that is less difficult to fit to the data, than other similarly good distribution fits. Figure 3 shows the wind directions where the data is clustered in 12 different direction sectors. Figure 4 shows each annual wind speed distribution and compares it to the four-year mean speed distribution.

## 2.3. Wake modeling

As the wind turbines individual wake generation is the key in maximizing the wind farms energy output, modeling the wake is a very important task. In this section two models will be highlighted. These are the same models, which will later also be used for the uncertainty quantification.

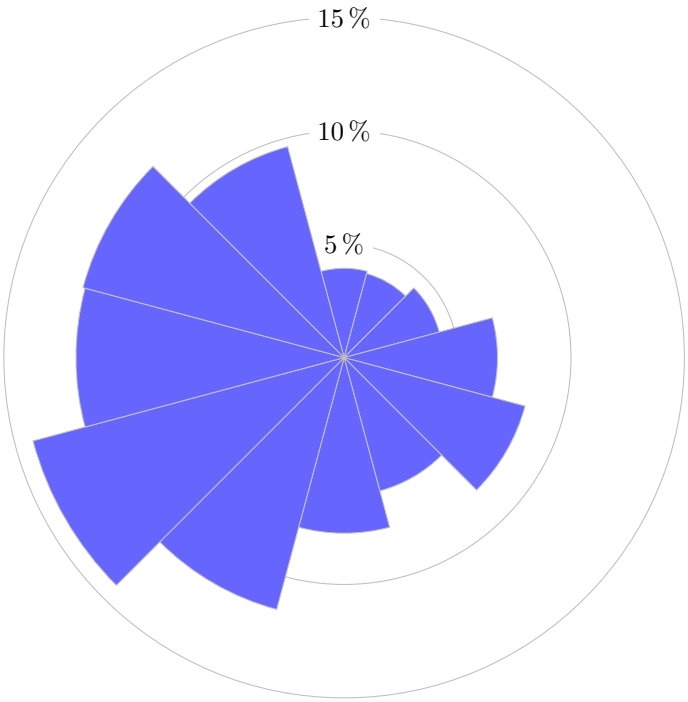



Figure 3: Combined wind direction distribution at FINO3 for the years 2011 to 2014, clustered into 12 direction sectors of with a total of 360 degrees

### 2.3.1. PARK model

The first model is the so called PARK model. It was originally developed by Jensen [5] and Katic et al [7]. The velocity deficit in this model is assumed to only change in stream direction. Therefore the velocity deficit

$$\delta u = \frac{u_o - u_w}{u_0} = 1 - \frac{u_w}{u_0}, \tag{1}$$

with wake disturbed velocity $u_w$ and $u_0$ as free stream velocity, is solely depend on x:

$$\delta u = \delta u(x) \tag{2}$$

Due to this assumption the model is not suitable for the calculation of an exact velocity distribution and is therefore only valid for the far wake case of a downstream distance of three rotor diameters ($D$) or higher. In figure 5 it can also be seen that the wake width increases linearly by $2k$. The corresponding wake growth factor $k$ is defined as:

$$k = \frac{0.5}{ln\frac{z}{z_0}}, \tag{3}$$

where $z$ is the turbines hub height and $z_0$ refers to the surface roughness, which is dependent on the on the building ground. The surface roughness will also be considered as constant in this thesis with $z_0 = 0.03$m.

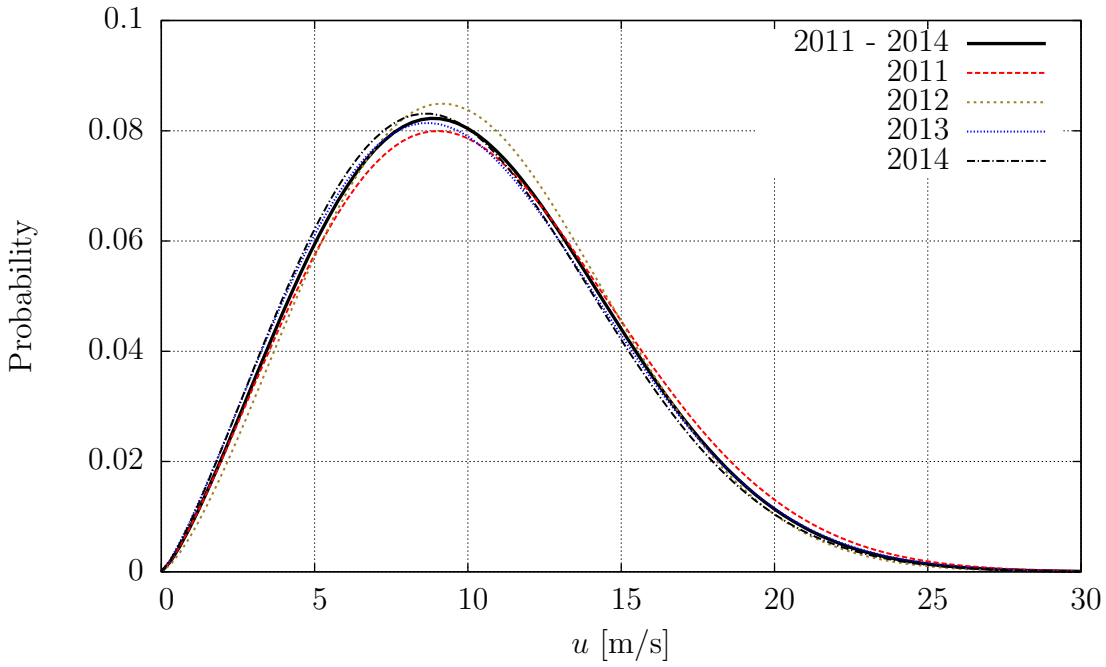



Figure 4: Comparison of annual wind speed distributions at FINO3 for the years 2011 to 2014

To calculate the velocity deficit inside the wake a momentum balance is applied for the system in figure 5. This leads to:

$$-\rho\pi\left(\frac{D}{2}\right)^2 u_r - \rho\pi\left(\left(\frac{D_w}{2}\right)^2 - \left(\frac{D}{2}\right)^2\right) u_0 + \rho\pi\left(\frac{D_w}{2}\right)^2 u_w = 0 \tag{4}$$

Under the previous assumption of incompressibility ($\rho = \textbf{const}$) for the air, the equation can be simplified to:

$$D^2 u_r + \left(D_w^2 - D^2\right) u_0 = D_w^2 u_w \tag{5}$$

For the velocity behind the rotor $u_r$, the initial velocity deficit is defined as $\delta u_r = 1 - \frac{u_r}{u_0}$. Together with $D_w = D + 2kx$ this yields:

$$\frac{u_w}{u_0} = 1 - \delta u_r \left(\frac{D}{D + 2kx}\right)^2 \tag{6}$$

According to Peña and Rathmann [10] the initial velocity deficit can be replaced by the axial induction factor, such as:

$$a(u_0) = 1 - \sqrt{1 - C_t(u_0)} \tag{7}$$

It is linked to the thrust coefficient $C_t$ and can be interpreted as the relative velocity loss at the turbine. Now it is possible to derive a formulation for the velocity deficit at any point inside the wake from 1 as:

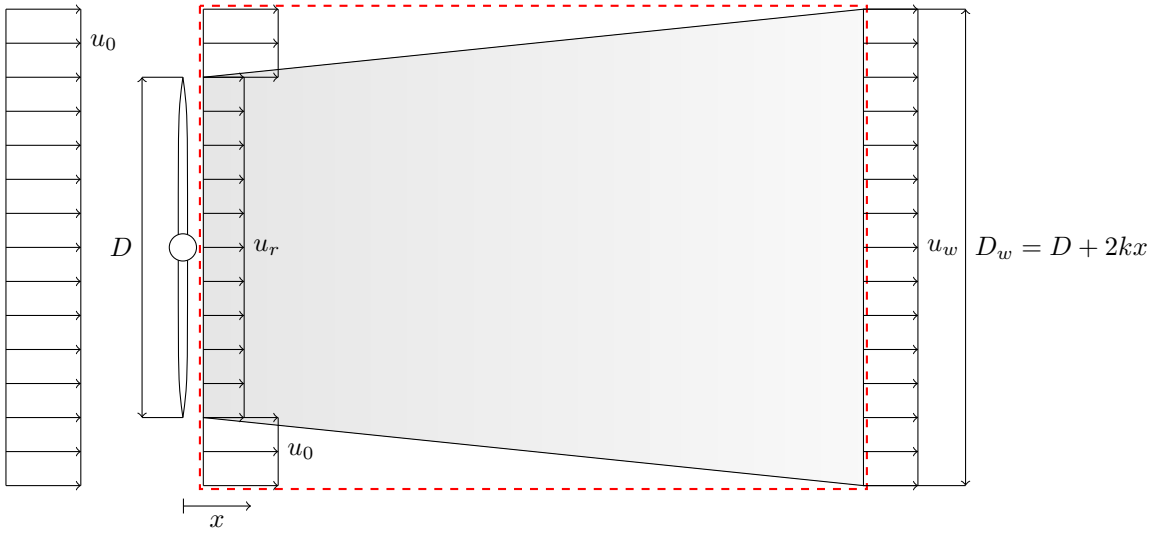



Figure 5: Visualization of the wake as described in PARK, cf. [7]

$$\delta u(x) = 1 - \frac{u_w(x)}{u_0} = \frac{1 - \sqrt{1 - C_t(u_0)}}{\left(1 + \frac{2kx}{D}\right)^2} \tag{8}$$

It is important to note that the derived equation is only valid for the velocity deficit behind a turbine in free stream. In practice this is not the case as only the wind turbines at the front of the wind park (stream-wise) will meet this condition. To address the issue a shadowing factor $\beta_k \in [0, 1]$ is introduced as presented in [3] and [13] as:

$$\beta_k = \frac{A_{\text{Intersection}}}{A_{\text{Turbine}}} \tag{9}$$

$A_{\text{Intersection}}$ describes the circular intersection of the wake cross section between wake generating turbine $i$ and wake affected turbine $j$. The introduction of this shadowing factor changes the formula for the velocity deficit as follows:

$$1 - \frac{u_w}{u_{inc,i}} = \beta_k \frac{\left(1 - \sqrt{1 - C_t(u_{inc,i})}\right)}{\left(1 + \frac{2kx}{D}\right)^2}, \tag{10}$$

with $u_{\text{inc}}$ as the incident velocity from turbine $i$. As this relation is only valid for the incident velocity from turbine it needs to be transformed in order to be depending on the free stream velocity $u_0$:

$$\delta u_{ij} = 1 - \frac{u_{w,i}}{u_0} = 1 - \frac{u_{inc,j}}{u_0} = \frac{u_0}{u_{inc,i}} \left( \frac{\beta_k \left(1 - \sqrt{1 - C_t(u_{inc,i})}\right)}{\left(1 + \frac{2kx}{D}\right)^2} \right) \tag{11}$$

For the case of multiple interacting wakes, velocity wakes are simply added up in a least squares sense:

$$\delta u_j = \sqrt{\sum_{i=1}^{N} \delta u_{ij}^2} \tag{12}$$

### 2.3.2. Eddy-viscosity model

The eddy-viscosity model originates in fluid dynamics as a turbulence closure for the stress tensor in the Reynolds averaged Navier-Stokes (RANS) equations. The model was first introduced in 1988 by Ainslie [1], based on numerical solutions of the shear layer approximations of the NS equations. The wake is assumed axisymmetric, turbulent and without any circumferential velocities. Additionally the flow field is stationary in time and pressure gradients outside of the wake are negligible. From these assumptions the momentum equation of the NS equations yields:

$$u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial r} = -\frac{1}{r}\frac{\partial\left(r\overline{uv}\right)}{\partial r}, \tag{13}$$

with $u$ as downstream velocity, $v$ radial velocity.
In general the equations is not closed as the stress tensor $\overline{uv}$ is unknown. The eddy viscosity approximation provides the closure term as

$$\overline{uv} = -\varepsilon\frac{\partial u}{\partial r}, \tag{14}$$

with $\varepsilon$ as eddy viscosity. Equation 13 combined with equation 14 further yields:

$$u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial r} = \frac{\varepsilon}{r}\left(\frac{\partial u}{\partial r} + r\frac{\partial^2 u}{\partial r^2}\right) \tag{15}$$

As the eddy viscosity $\varepsilon$ is also unknown for now, the term needs to be closed by further assumptions. Therefore the eddy viscosity is modeled to consist of an ambient eddy viscosity and a wake generated viscosity:

$$\varepsilon = l_w(x)\, u_w(x) + \varepsilon_a \tag{16}$$

The characteristic length $l_w$ in this context is proportional to the wake width $r_w$ with proportionality factor $k_l$ ($\approx 0.015$ [1]) as:

$$l_w(x)\, u_w(x) = k_l r_w(x)(u_0 - u_c(x)) \tag{17}$$

and

$$r_w = \sqrt{\frac{3.56C_t}{4\delta u_c\,(2-\delta u_c)}} \cdot D, \tag{18}$$

where $u_c$ corresponds to the wake centerline velocity. The ambient eddy viscosity can also be defined by the eddy diffusivity for neutral conditions:

$$\varepsilon_a = K_M = \frac{\kappa^2}{\log\left(\frac{z}{z_0}\right)} u_0 D, \tag{19}$$

with $\kappa$ as von Kármán constant.

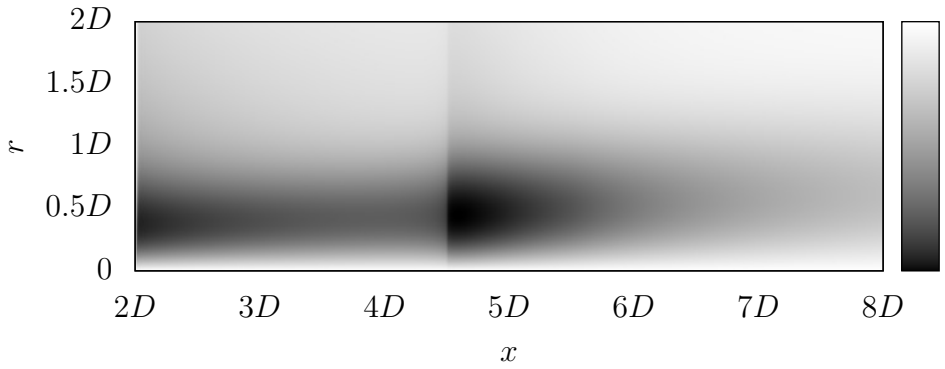



Figure 6: Original filter function, see equation (20)

Wind tunnel experiments showed a discrepancy in the eddy viscosity, which was lower than the calculated value for downstream distances of five rotor diameters or closer. To compensate this flaw, Ainslie [1] introduced a filter function F(x) as follows:

$$F(x) = \begin{cases} 0.65 + \left(\frac{x/D - 4.5}{23.32}\right)^{1/3} & x \leq 5.5\,D \\ 1 & x > 5.5\,D \end{cases} \tag{20}$$

The proposed filter function has an infinite slope at $x = 4.5D$ (see figure 6). This jump also leads to a jump in the velocity distribution as it can be seen in figure 6. This potentially unphysical effect can be avoided by choosing a different filter function with a smoother transition from near to far-field as shown by Heiming [4].

## 2.4. Annual energy production

Finally, to optimize the wind park layout means to maximize the AEP. To be able to do this, the expected power value needs to be calculated. Therefore the wind speeds $u$ get categorized by wind directions $\varphi_i$ similar to figure 3 and the expectation of the power value is expressed by:

$$\begin{aligned} E_{\varphi_i} &:= E_{\varphi_i}\left[P(\varphi_i, u)\right] \\ &= \int_0^{\infty} P(\varphi_i, u) \cdot F_{\varphi_i}(u)\,\mathrm{d}u \\ &= \int_{u_{cutin}}^{u_{cutout}} P(\varphi_i, u) \cdot F_{\varphi_i}(u)\,\mathrm{d}u \\ &\approx \sum_{j=1}^{N_{speeds}} w_j \cdot P(\varphi_i, u_j) \cdot F_{\varphi_i}(u_j) \end{aligned} \tag{21}$$

$F_{\varphi_i}(u)$ describes the probability density function (PDF) of the wind speed distribution and $w_j$ is the weight for speed class $j$. Further

$$P(\varphi, u) = \sum_{k=1}^{N_{turbines}} P(u_{inc,\varphi,k}) \tag{22}$$

denotes the total power output of the wind farm at wind speed $u$ and direction $\varphi$. The incident velocity at turbine $k$ in wind direction $\varphi$ can be determined by computing the wake velocity deficit by using the wake models derived in the previous section.
The expected power output for all directions then computes as follows:

$$\begin{aligned} E &:= \int_0^{2\pi} E_{\varphi}\mathrm{d}\varphi \\ &\approx \sum_{i=1}^{N_{directions}} w_{\varphi_i} \cdot E_{\varphi_i}, \end{aligned} \tag{23}$$

where $w_{\varphi_i}$ is the weight for direction $\varphi_i$.

As $E$ describes the mean power for a given wind distribution in MW, this must be projected to the length of one year to get the AEP measured in MWh:

$$\begin{aligned} AEP &= (8760\,\mathrm{h} + 6\,\mathrm{h}) \cdot E \\ &\approx (8760\,\mathrm{h} + 6\,\mathrm{h}) \cdot \sum_{i=1}^{N_{directions}} w_{\varphi_i} \cdot \sum_{j=1}^{N_{speeds}} w_j \cdot W_{\varphi_i}(u_j) \cdot \sum_{k=1}^{N_{Turbines}} P(u_{inc,\varphi_i,j,k}) \end{aligned} \tag{24}$$

'8760 + 6' corresponds to the total number of hours in a single year of 365.25 days, thereby also accounting for leap years.

# 3. State of the art

As previously introduced, the process of turbine induced wake generation is potentially depending on a large set of parameters. Most of these parameters have a high level of uncertainty due to e.g. the shear inability of performing ideal measurements or factoring in situational performance decreases of the wind farm.
As described by Lackner et al. [8] in 2007 the most important uncertainties can be clustered in three groups, with the most crucial uncertainty being the wind resource. Wind speed and direction are in the case of the Horns Rev 1 wind farm measured by the FINO3 weather station and can therefore not fully represent data for each turbine.
The second set of uncertainties is the turbine power production. Here the most significant role can be linked to the turbine power curve. When the power curve is determined by the manufacturer the curve is fitted to wind speed, but it is uncertain how the mapping is performed as turbulence and shear stress will have a significant impact on the curve. An additional uncertainty of this group is of the air's density as it will vary under different influences like wind velocity and humidity.
Finally the last set of uncertainties are the energy loss factors. Here, availability losses and losses due to fouling and icing play a role and can only be predicting to a limited extend, thus being highly uncertain.
Lackner et al. [8] modeled these uncertainties as Weibull distributed for the wind resources, while the other uncertainties are supposed to be normally distributed. Despite omitting any details on the used wake model, it is stated that a Monte-Carlo simulation has been performed and a method has been derived to calculate the statistics of the AEP.
In a newer approach by Murcia et al. [9] from 2015 the focus was laid on creating a framework for the validation of stationary wake models with the goal to quantify the AEPs uncertainty. Murcia et al. [9] also used the same Horns Rev 1 offshore wind power plant as a test case. The setup will therefore be the same as listed in section 2. In the paper a smaller set of uncertainties is considered, including wind velocity and direction. The turbines power curve is set to the official curve provided by the manufacturer. Again the simulation was performed by a Monte-Carlo simulation, this time based on LHS sampling. The wake model is a modified NOJ model by [7], a simple model, which is yet broadly used in the industry. [9] could show through the calculations that the uncertainties in the model inputs are too large to draw a conclusion on possible model inadequacies. Meaning uncertainties in the measurements outweigh uncertainties induced by the model. It is lastly adviced to also consider uncertainties in the power curve.
It is worth noting that the derived results only calculate uncertainties in the AEP for a single wind turbine. A comparable work for computing a full wind power plant AEP uncertainty prediction could not be found at the point in time of creating this work.

# 4. Uncertainty Modeling

In this section the basic principles of uncertainty quantification will be explained and methods will be introduced to evaluate uncertainties for a specific quantity of interest.

## 4.1. Basic concepts of probability

In order to describe uncertainties, it is first of all necessary to define basic concepts of stochastics and distributions.

### 4.1.1. Random variables

To begin with, a random variable (RV) $X$ has a cumulative distribution function (CDF) $F_X(x)$ and a PDF $f_X(x)$. The cumulative distribution function is defined as:

$$F_X(x) = P(X \leq x) = P(\omega : X(\omega) \leq x), \tag{25}$$

where by definition of probability $P$: $0 \leq F_X \leq 1$. The probability density function is linked to the cumulative distribution function through the relation:

$$F_x(a) = \int_{-\infty}^{a} f_X(x)dx \text{ and } f_X(x) = \frac{d}{dx}F_X(x) \tag{26}$$

The cumulative distribution function therefore describes the probability of $X$ having a value less or equal to $x$, where the probability density function describes the relative likelihood of $X$ to take value $x$.
An important example for a distribution is the Gaussian distribution which is also known as normal distribution $\mathcal{N}(\mu, \sigma^2)$. The parameter $\mu$ is the mean and $\sigma^2$ the variance. The PDF of this distribution can be formulated as:

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x-\mu}{2\sigma^2}}, x \in \mathbb{R} \tag{27}$$

Another example for an important distribution is the so called uniform distribution. For this distribution all outcomes within its support have the same probability, while being zero outside of the support. The support is defined as a interval $[a, b]$ on $\mathbb{R}$ with $a, b \in \mathbb{R}$. For the PDF this yields:

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{for } x \in [a, b], \\ 0 & \text{else} \end{cases} \tag{28}$$

### 4.1.2. multiple RVs

In case of multiple random variables the definition of a joint density function, as well as the marginal density function, is required.
The joint density function in case of two continuous RVs is defined as function f(x,y) which fulfills:

$$f_{X,Y}(x,y) = f_{X|Y}(x|y)f_Y(y) = f_{Y|X}(y|x)f_X(x) \tag{29}$$

As this also is a density function

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = 1 \tag{30}$$

has to hold as well. As stated, these are formulations for the two-dimensional case, but can be extended to an n-dimensional case.
The undefined terms from equation 29 are the conditional distribution e.g. $f_{X|Y}(x|y)$ and the marginal distribution e.g. $f_Y(y)$.

### 4.1.3. Statistics

The most important statistics in dealing with uncertainty quantification are the expectation and the variance. The expectation of RV X with PDF $f_X$ can be computed as:

$$\mu_x = \mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x)dx \tag{31}$$

The variance is defined as

$$\sigma_x^2 = Var(X) = \mathbb{E}[(X-\mu_x)^2], \tag{32}$$

which is the squared standard deviation $\sigma_x$ of $X$. As the expectation and the standard deviation are only the first and second standardized moments of the distribution, one could also compute higher moments, but these are not of interest in this problem setting.

## 4.2. Generalized polynomial chaos

In this section the generalized polynomial chaos (gPC) is outlined. This method refers to an expansion which can be used to describe stochastic processes $u(x,t,\omega)$. The expansion is based on a sum of orthogonal polynomials, like the Hermite or Lagrange polynomials, depending on the distribution of the RV. These orthogonal polynomials form a basis for the random components in the solution. For the general case, the described process is a function of the random vector $Q(\omega) = [Q_1(\omega), \ldots, Q_p(\omega)] : \Omega \rightarrow \mathbb{R}$, $\Omega$ being the finite sample space:

$$u^N(t,x,\omega) = \sum_{i=0}^{N} u_i(t,x)\psi_i(Q) \tag{33}$$

| Distribution | GPC basis polynomial | Support |
|---|---|---|
| Gaussian | Hermite | $(-\infty, \infty)$ |
| Uniform | Legendre | $[a, b]$ |

Table 1: PDF and matching basis polynomials

where $u_i(x,t)$ are deterministic coefficients and $\psi_i$ the referred orthogonal polynomials. As previously mentioned the used polynomials have to be orthogonal to the PDF of the specific random process. Table 1 shows matching pairs of distributions and polynomials relevant within the context of this thesis.

## 4.3. Stochastic Spectral Methods

This section introduces commonly used methods in dealing with uncertainty quantification. These methods enable dealing with the uncertain parameters of the problem setting and propagating their stochastics through the used mathematical models. Ultimately the uncertainties are modeled with the intend to have a measure of uncertainty for a specific QoI. This measure is usually the mean and variance of the QoI. The provided methods in this section will therefore give expressions for these statistics.

### 4.3.1. Monte Carlo Sampling

In order to compute a QoI such as:

$$\mathbb{E}[u(x,t,Q(\omega))] = \int u(x,t,q)\rho(q)dq \tag{34}$$

the MC method uses point-wise integration in probability space to evaluate the integral in equation 34. By computing $M$ deterministic solutions (sampling), with each starting from a different set of realizations of the uncertain parameters one obtains $M$ solutions of the type $u^m(x,t) = u(x,t,Q^m)$. If $Q^m$, $m = 1 \dots M$ is a independet and indentical distributed (i.i.d.) RV, application of the central limit theorem yields:

$$\frac{1}{M}\sum_{m=1}^{M} u(x,t,Q^m) \xrightarrow{a.s.} \mathbb{E}[u(x,t,Q)] \tag{35}$$

and for the convergence rate:

$$Var[\frac{1}{M}\sum_{m=1}^{M} u(x,t,Q^m) - \mathbb{E}[u(x,t,Q)] = \mathcal{O}(\frac{1}{\sqrt{M}}) \tag{36}$$

The rather slow convergence rate of $\mathcal{O}(\frac{1}{\sqrt{M}})$ can be a disadvantage for all computationally expensive problems, as for every digit of accuracy the sample size has to be increased by a factor of 100. This can be a deal-breaker for certain applications, but it is worth mentioning that the so called "Quasi Monte Carlo methods" offer similar

properties with an increased efficiency. On the other hand the convergence rate of the MC sampling is not dependent on the dimensions of the probability space. This is a unique feature to the MC method, which in case of high dimensional probabilities often makes it the only viable solution.

In case of multiple RV the MC sampling method stays the same. Random numbers are drawn from the PDF of each RV and the function is evaluated at these points. Therefore, formulations for the mean and variance also stay the same.

### 4.3.2. Stochastic Collocation

The idea behind the SC method is it to choose a set of $M$ collocations points $q_M$ in probability space and to enforce the solution at these points. In a more mathematical formulation this means:

$$\text{find } \mathcal{I}(Q) \text{ s.th.} \quad u(x,t,Q^m) = \mathcal{I}(Q^m) \tag{37}$$

Note that $\mathcal{I}$ is function valued, which makes the task non-trivial. As the problem is deterministic for each sample $m$, the SC method involves solving the problem $M$ times. As the positions of the collocation points are unrestricted, it is advantageous to choose them based of a quadrature rule and thus exploit the corresponding weights to compute the statistics. The solution of the stochastic problem can be interpolated from the collocation points by using Lagrange interpolation. This yields the following expression for the solution:

$$u^N(x,t,Q) = \sum_{i=1}^{M} u_i(x,t,q_i) L_i(q), \tag{38}$$

where $L_i$ are the Lagrange polynomials. The reason to choose Lagrange polynomials is due to their ability to match the exact solution at all $M$ collocation points, as:

$$L_i(q^m) = \delta_{im} \tag{39}$$

The statistics of the interpolation can then be determined by using the solutions at the collocation points. For the mean of the QoI this for example yields:

$$\mathbb{E}[u] = \sum_{i=1}^{M} u(x,t,q_i) \int_\Gamma L_i(q)\rho(q)dq \tag{40}$$

Where $\Gamma$ is the probability space on which $Q$ is defined and $\rho(q)$ is the PDF of $Q$'s distribution. By cleverly choosing the collocation points as described above and using the Lagrange property in 39 equation 40 can be simplified to:

$$\mathbb{E}[u] = \sum_{i=1}^{M} u(x,t,q_i)\rho(q_i)w_i \tag{41}$$

$w_i$ are the quadrature weights of the quadrature rule which was chosen according to the CDF of the uncertainty. The integral in equation 40 can be exactly evaluated by quadrature due to the fact that since the used Lagrange polynomials are of order M the quadrature is exact for polynomials of this order. Equation 40 could also be solved by setting up a Vandermonde-like matrix and inverting it, which is a non-trivial task and can sometimes even be impossible. A formulation for the variance can also be found:

$$\begin{aligned}\mathbb{E}[(u-\mathbb{E}[u])^2] &= \int_\Gamma (u-\mathbb{E}[u])^2\rho(q)dq \\ &= \int_\Gamma (\sum_{i=1}^{M} u(q_i)L_i(q) - \mathbb{E}[u])^2\rho(q)dq \\ &= \sum_{j=1}^{M}(\sum_{i=1}^{M} u(q_i)\delta_{ki} - \mathbb{E}[u])^2\rho(q_j)w_j \\ &= \sum_{j=1}^{M}(u(q_j) - \mathbb{E}[u])^2\rho(q_j)w_j \end{aligned} \tag{42}$$

The interpolation error for dimension $p$ with $M$ collocation points is:

$$||f - \mathcal{I}_M f|| \sim \mathcal{O}(\frac{1}{M^{\frac{-\alpha}{p}}}), \tag{43}$$

where $\alpha$ is a constant that is based on the smoothness of the interpolated function $f$. Compared to the MC method, the convergence rate of $(M^{\frac{-\alpha}{p}})^{-1}$ is depending on the dimension $p$ and thus the convergence rate is low for higher dimensions. This is the so called "curse of dimensionality", which restricts the usability of most non-MC type methods to at most medium dimension size.

In the case of multiple RV the interpolation of the solution $u$ to a multivariate partial differential equation (PDE) can be formulated as:

$$\begin{aligned} u(Q) = \mathcal{I}(u) &= (\mathcal{I}_{m_1} \otimes \cdots \otimes \mathcal{I}_{m_d})(u) \\ &= \sum_{j_1}^{m_1} \cdots \sum_{j_d}^{m_d} u(q_1^{j_1}, \ldots, q_d^{j_d})(L_1(q_1^{j_1}) \ldots L_d(q_d^{j_d})), \end{aligned} \tag{44}$$

with $d$ as the dimension (number of RVs) and $L_i$ again as Lagrange polynomial in i-th dimension. Using a multi-index $i$ equation 44 can be rewritten as:

$$\mathcal{I}_m(u) = \sum_{|i|<N} u(q_i)L_i(q) \tag{45}$$

A quadrature rule $\mathcal{Q}_l^{(p)} f$ based on 1D quadratures can also defined and statistics can be computed similarly as in the one-dimensional case. For the derivation of an efficient

quadrature rule see chapter 4.4 where sparse grids are discussed. Given the quadrature rule, the mean and variance can be calculated by:

$$\mathbb{E}[u] = \sum_{|i|<N} u(\mathbf{Q}_i) W_i \tag{46}$$

and

$$\text{var}[u] = \sum_{|i|<N} (u(\mathbf{Q}_i) - \mathbb{E}[u])^2 W_i, \tag{47}$$

with $\mathbf{Q}$ as tensor product of the RV and $W_i = \{\rho_i w_i^1, \dots, \rho_i w_i^d\}$.

It is worth noticing that the presented method, similar to the MC method, is also a non-intrusive method. This means, assuming a suitable code structure, no adjustments the solver itself have to be done. The MC and SC can work as wrapper function for the original problem and evaluate statistics of the QoI after taking enough samples. This is a big advantage compared to other available methods such as the stochastic Galerkin (SG) method where the original code has to be altered extensively.

## 4.4. Smolyak Sparse grids

In the attempt to weaken the curse of dimensionality occurring in multivariate problems, sparse grids can be used to lower the computational effort required. These sparse grids refer to a technique for integrating or interpolating a high dimensional function. They were first introduced by Smolyak and are therefore also known as Smolyak sparse grids.
In general the approximation of an multidimensional integral can be formulated as:

$$I^{(p)} f = \int_\Gamma f(q)\rho(q)dq, \tag{48}$$

with $\Gamma$ as $p$-dimensional hypercube $\Gamma = [0,1]^p$. Assuming a quadrature rule $\mathcal{Q}_{l_i}^{(j)}$ for the i-th integration direction, a tensor product rule can be defined by:

$$\begin{aligned} \mathcal{Q}_l^{(p)} f &= (\mathcal{Q}_{l_1}^{(1)} \otimes \cdots \otimes \mathcal{Q}_{l_p}^{(1)}) \\ &= \sum_{r_1=1}^{R_{l_1}} \cdots \sum_{r_p=1}^{R_{l_p}} f(q_1^{r_1}, \dots, q_p^{r_p}) w_{l_1}^{r_1} \dots w_{l_p}^{r_p} \end{aligned} \tag{49}$$

The overall quadrature therefore consists of several nested 1D quadratures $\mathcal{Q}_l^{(1)}$ each defined as:

$$\mathcal{Q}_{l_1}^{(1)} = \sum_{r=1}^{R_{l_1}} f(q_l^r) w_l^r, \tag{50}$$

with $q_l^r$ and $w_l^r$ as points and weights at level $l$ and $R_{l_i}$ as the corresponding number of nodes.

| R | |
|---|---|
| 0 | $1$ |
| 1 | $x \quad y$ |
| 2 | $x^2 \quad xy \quad y^2$ |
| 3 | $x^3 \quad x^2y \quad xy^2 \quad y^3$ |
| 4 | $x^4 \quad x^3y \quad x^2y^2 \quad xy^3 \quad y^4$ |

Table 2: Product of monomials up to R=4

On a full tensorized grid the number of points can be calculated as:

$$R = \prod_{i=1}^{p} R_{l_i} \tag{51}$$

As it is required for the collocation to compute a sample for each quadrature node, keeping the total node count as low as possible thus lowers the computational cost. Smolyak came up with a sparse quadrature formula which lowers the total number of nodes required. The underlying idea of his sparse grid construction can best be explained by comparing it to the quantification of a quadrature rules accuracy. The accuracy is usually calculated by considering products of monomials and checking, which monomials of degree $R$ can be integrated exactly. To specify monomials of e.g. degree four a normal full tensor product will have 25 terms, whereas one would only need to consider the five terms which can be seen in table 2. The full tensor product includes terms of higher order than necessary such as e.g. $x^4y$. Smolyak sparse grids work similarly. Constructed grids and weights are supposed to yield the same accuracy while using significantly less points.
To construct a sparse grid quadrature rule a difference relation is introduced as:

$$\Delta_l^{(1)} f = (\mathcal{Q}_l^{(1)} - \mathcal{Q}_{l-1}^{(1)}) f, \tag{52}$$

where $\mathcal{Q}_l^{(1)}$ is a 1D quadrature rule (eq. 50). Note that $\mathcal{Q}_0^{(1)} \equiv 0$.
$\Delta_l^{(1)} f$ is also a quadrature formula which uses, due to the nested property of the nodes, the same nodes as $\mathcal{Q}_l^{(1)}$. The weights of this quadrature are also simply the difference of the weights on level $l$ and $l-1$. Since nodes are nested in a sense that nodes on level $l-1$ are a subset of the nodes on level $l$, negative weights are also possible.
The sparse quadrature formula at level $l$ is then given as:

$$\mathcal{Q}_l^{(p)} f = \sum_{|l'| \leq l+p-1} (\Delta_{l_1}^{(1)} \otimes \cdots \otimes \Delta_{l_p}^{(1)}) f, \tag{53}$$

with $l' = (l_1, \ldots, l_p) \in \mathbb{N}^p$ as multi-index where $|l'| = \sum_{i=1}^{p} l_i$.
The full tensor product on the other hand can be written as:

$$\mathcal{Q}_l^{(p)} f = \sum_{\max l' \leq l} (\Delta_{l_1}^{(1)} \otimes \cdots \otimes \Delta_{l_p}^{(1)}) f, \tag{54}$$

| $p$ | $R_l$ | sparse grid nodes | tensor grid nodes |
|---|---|---|---|
| 2 | 5 | 13 | 25 |
| | 9 | 29 | 81 |
| 5 | 5 | 61 | 3125 |
| | 9 | 241 | 59049 |
| 10 | 5 | 221 | 9765625 |
| | 9 | 1581 | $> 3 \times 10^9$ |
| 50 | 5 | 5101 | $> 8 \times 10^{34}$ |
| | 9 | 171901 | $> 5 \times 10^{47}$ |
| 100 | 5 | 20201 | $> 7 \times 10^{69}$ |
| | 9 | 1353801 | $> 2 \times 10^{95}$ |

Table 3: Number of Clenshaw-Curtis nodes for the sparse and tensor grids

with $\max l' = \max \{l_1 \dots l_p\}$. Even though differences seem small and depending the choice of nodes the overall node count still has an exponential growth factor, it can be seen in table 3, that the sparse grid offers a huge reduction in required nodes.

According to [12] the quadrature error satisfies:

$$||I^{(p)} f - \mathcal{Q}_l^{(p)} f|| = \mathcal{O}(R^{-\alpha} \log(R)^{(p-1)(\alpha+1)}), \tag{55}$$

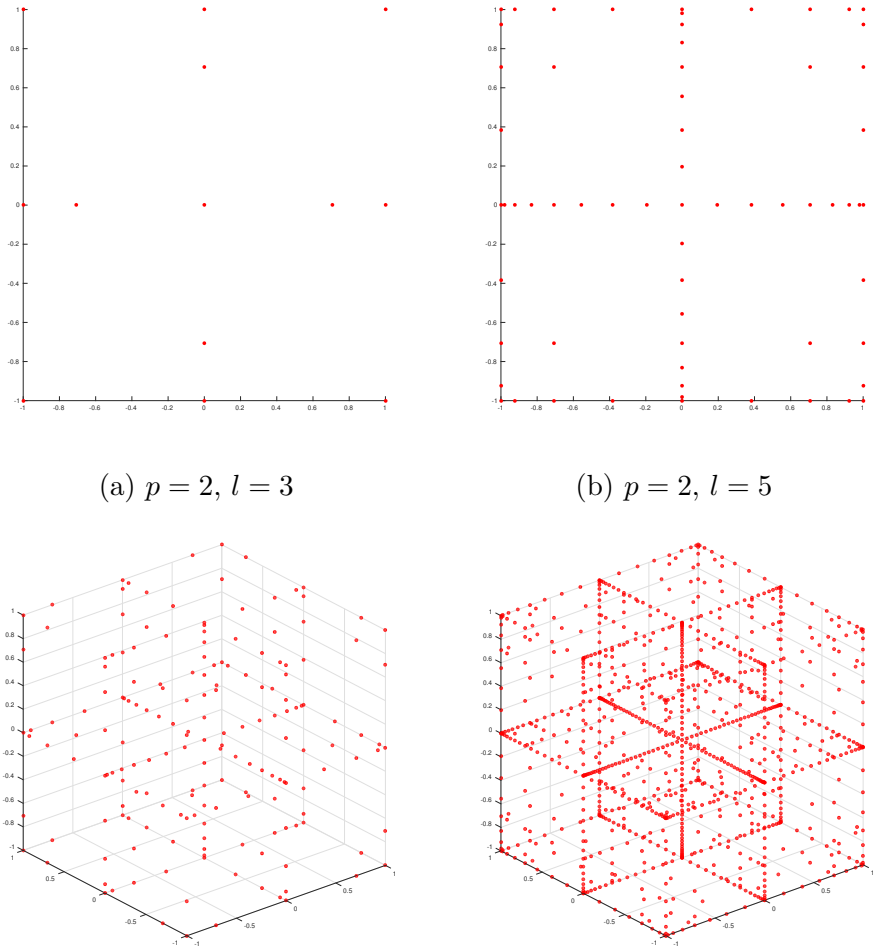
with $\alpha$ as constant dependent on the smoothness of function $f$. For a very high dimension $p$ the $R^{-\alpha}$ term dominates and therefore the convergence rate can no longer compete with the MC method, which has a convergence rate of $R^{-1/2}$. To conclude, the collocation method on sparse Smolyak grids is suitable low to medium dimension $p$, but will suffer from the curse of dimensionality if the dimension is too high.

As mentioned earlier the node count is depending on the node type. In this work the Clenshaw Curtis (CC) nodes are used for all sparse grids. These nodes are the extrema of the Chebychev polynomials typically defined on the interval $[-1, 1]$. The nodes can be computed by:

$$q_l^r = -cos \frac{\pi(r-1)}{R_l - 1} \quad , r = 1, \dots, R_l, \tag{56}$$

with $R_1 = 1$ and $R_l = 2^{l-1} + 1$ for $l > 1$.

A visualization of the sparse Smolyak grids with CC nodes can be seen in figure 7. The different refinement levels $l$ and dimensions $p$ give an impression on the overall structure of the nodes and also their nestedness. From the scatter plots it also becomes obvious that more points are added at the central axes and the boundaries of the domain. This leads to an exponential increase in nodes on higher levels. Alterations of the CC nodes exist which provide a lower growth factor. See Burkardt and Webster [2] for further reading.



(a) $p = 2$, $l = 3$ (b) $p = 2$, $l = 5$

(c) $p = 3$, $l = 5$ (d) $p = 3$, $l = 7$

Figure 7: Visualization of different Smolyak sparse grids with nested Clenshaw Curtis nodes

# 5. Verification

In order to verify the implementations of the methods outlined in the previous modeling section 4, the Burgers equations will be tested for each of the methods. The Burgers equation is well suited to test the derived methods of UQ on. The obtained results are highly dependent on the boundary condition (BC) and are sensitive even to very small changes.
The one-dimensional Burgers equation reads:

$$\begin{cases} \frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0, \quad x \in [-1, 1] \\ u(-1, t) = 1 \\ u(1, t) = -1 \\ u(x, 0) = u_0 \end{cases}, \tag{57}$$

with velocity $u$ and viscosity $\nu$. The Burgers equation (57) is of the hyperbolic and non-linear PDE type and has spatial as well as temporal derivatives. The deterministic solutions of the equation is plotted in figure 8 for $t = 0$ with two different BCs at the left boundary.

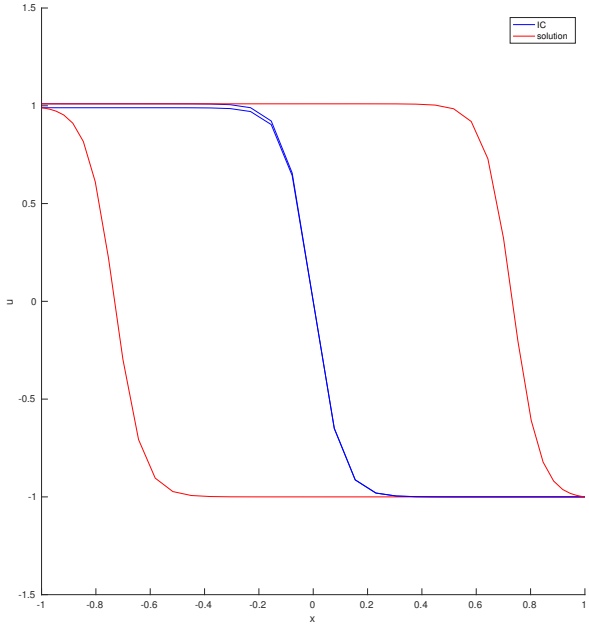



Figure 8: Solutions of the Burgers equation (red) with small differences in the IC (blue)

| viscosity $\nu$ | time step size $\Delta t$ | convergence parameter $\epsilon$ | spatial steps | domain |
|---|---|---|---|---|
| 0.05 | 5E-04 | 1E-06 | 46 | $[0, 1]$ |

Table 4: Parameters used in the computation

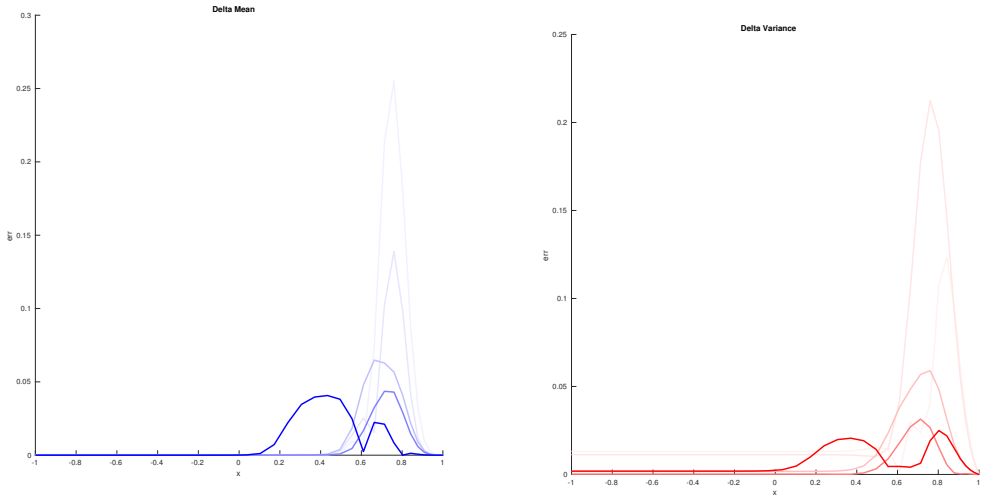
| Distribution | Samples (MC) | Collocation points SC |
|---|---|---|
| $q \sim \mathcal{U}(0, 0.1)$ | 100 | 10 |

Table 5: Stochastic parameters

In general, the solution is almost constant at each boundary, mirroring the BC, but has shock right at the center. Position of the shock is highly sensitive to the BCs, as it can be seen in figure 8. A small distortion in the left BC of $\pm 1\%$ leads to a significant change of the shocks location. This makes the Burgers equation especially suitable to test different UQ methods.
To test the derived methods, the original Burgers equation needs to be modified. An uncertainty parameter $\delta(q)$ is added to the left boundary condition and the equation now reads:

$$\begin{cases} \frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0, & x \in [-1, 1] \\ u(-1, t) = 1 + \delta(q) \\ u(1, t) = -1 \\ u(x, 0) = u_0 \end{cases}, \tag{58}$$

where $\delta(q)$ is a random perturbation based on a PDF.
Despite the introduced MC and SC methods are non-intrusive and basically can handle the deterministic equations as a black-box problem, it is worth mentioning that a 3rd order Runge-Kutta (RK) method was used to solve the equation 58 in an iterative manner until reaching equilibrium. All used computational parameters are listed in table 4 and 5. The uncertainty is assumed to be uniformly distributed and within the interval $[0, 0.1]$ s.th. $q \sim \mathcal{U}(0, 0.1)$.
The obtained results for the MC and the SC can be seen in figure 10 and 11. As mentioned, each solution was iterated until equilibrium was reached, i.e. the difference between two time steps was smaller than $\epsilon$. For the MC method a sample size of $n_{MC} = 100$ were needed in order to be comparable to results of the SC method which made use of $n_{SC} = 10$ collocation points (samples) to determine the results. Both obtained solutions show comparable characteristics, where the variance is high in regions where the shocks are occurring. Due to the high sensitivity w.r.t to the BC this behavior is exactly as expected. For the MC method the variance is larger for $x \in (0, 0.5)$. This effect comes from a drawn random number close to zero for which the shock is almost located at $x = 0$. Due to the deterministic Legendre nodes used in the SC method, these point are not considered by this method. The MC sampling can therefore show its advantage of being able to integrate arbitrary functions with e.g. high input sensitivities like the Burgers equation.
A delta plot between the SC method with increasing sample size for the MC method can be seen in figure 9. The sample sizes are $n_{MC} = [5, 10, 25, 50, 100]$, where the color

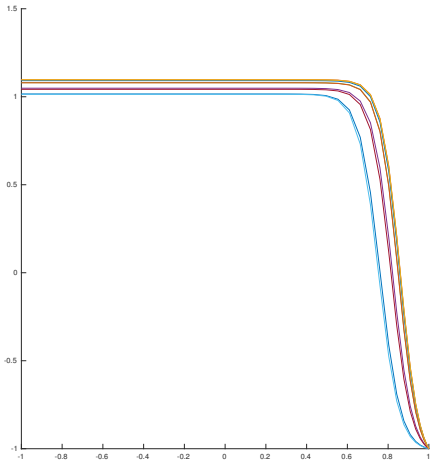



(a) Expectation

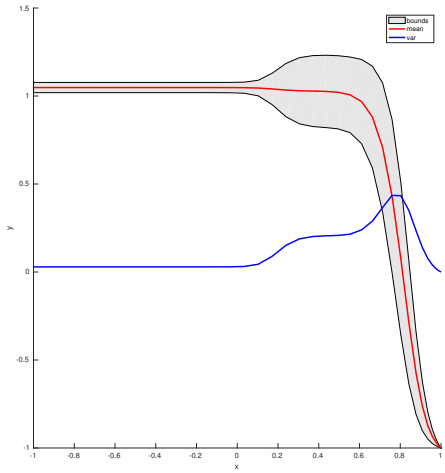
(b) Variance

Figure 9: $\Delta$ between both methods with increasing $n_{MC}$

intensity in the plots increases with $n_{MC}$. The plots show that for the expectation as well as the variance MC method slowly converges towards the SC methods solution with increasing sample size.
Because runtimes vary greatly due to the sensitive dependence of the solution on the IC / BC, a speed comparison is not reasonable for this problem. Nevertheless, as expected the SC method is faster in practice when achieving comparable accuracy ($SC_{10}$: 525.9s to $MC_{100}$: 4190.2s).
The obtained results for the SC are identical as found in Kærgaard [6]. For the MC sampling no such verification could be found, but the results are close to the ones computed by the SC method and the deviations can easily be explained.

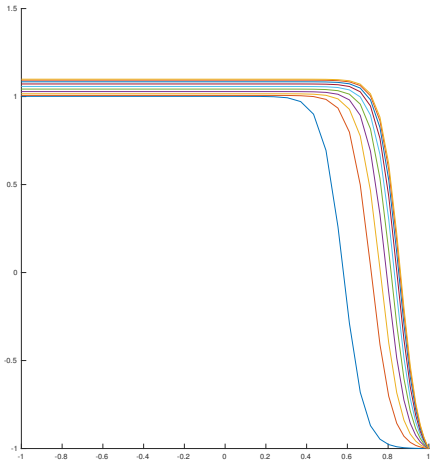


(a) Solutions for $n_{MC} = 10$ samples

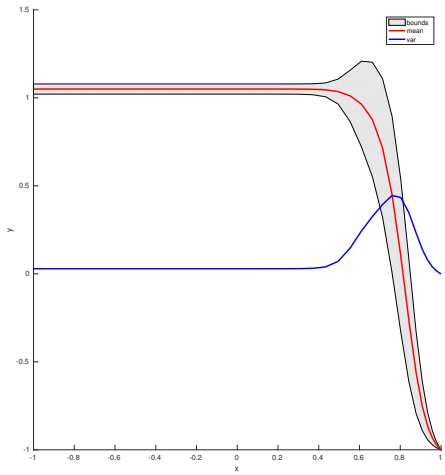



(b) Statistic results ($n_{MC} = 100$)

Figure 10: Results for the MC method



(a) Solutions for $n_{SC} = 10$ samples




(b) Statistic results

Figure 11: Results for the SC method

# 6. Results

In this final section, the methods introduced in chapter 4 will be used for the problem setting in chapter 2. The uncertainties for the wind speed, power curve and the thrust curve will be modeled and their impact upon the AEP, the QoI, will be measured. Runtimes for the various methods will then be compared afterwards.

## 6.1. Evaluation

Sadly the evaluation could only be performed for the PARK model. The eddy viscosity model caused the simulations to crash and thus made it impossible to obtain any results. Because the deterministic simulation was part of a previous work, this made it more difficult to revise the method and ultimately the time to do it was not sufficient. Additionally only univariate uncertainties are examined. The derived Smolyak sparse grids therefore won't be used in this section. Nevertheless, the sparse grid is implemented and has been verified to work by integrating over the hypercube $\Omega = [-1, 1]^d$.

### 6.1.1. Wind speed

The uncertainty in the wind speed is supposed to be modeled as a Weibull distribution, as this kind of asymmetric PDF is usually well suited to mimic natural wind resources. The probability density function for this kind of distribution reads:

$$f_{A,k}(x) = (\frac{k}{A})(\frac{x}{A})^{k-1}\exp(-(\frac{x}{A})^k), \tag{59}$$

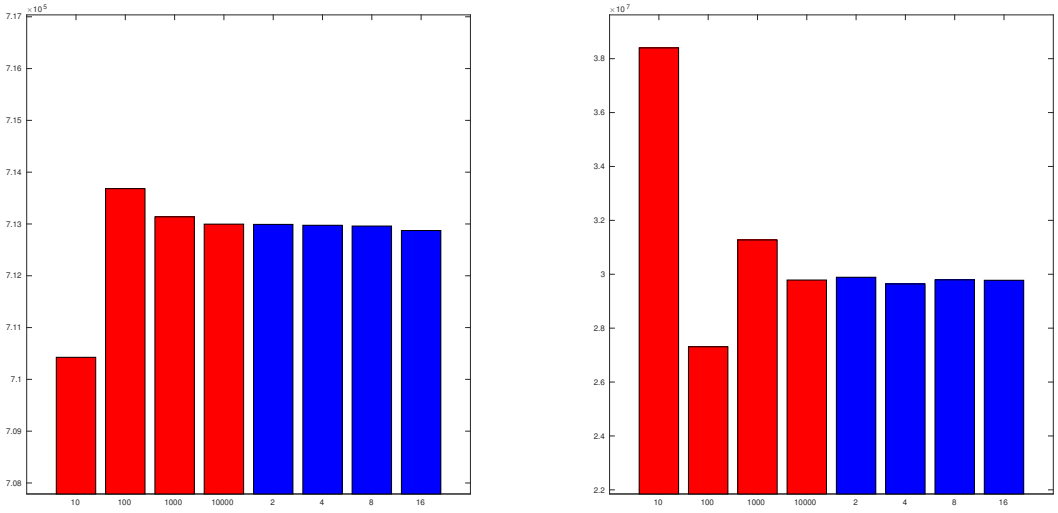
with scaling parameter $A$ and a shape parameter $k$.
Data provided by large power company claims, that the standard deviation $\sigma$ is between 2.2% and 8.6%. As it is unclear how these estimates were derived or measured, the mean value of 5.4% will be considered in this work. In other publications, such as Murcia et al. [9], lower values in the magnitude of the lower bound have been accounted. In order to determine both parameters $A$ and $k$, the value of the each measurement point would be the mean value of the Weibull distribution. To recreate the mean and standard deviation, an iterative solver would then be needed in order to approximate the parameter. Based on the equations for the mean:

$$\mathbb{E}(X) = A\Gamma(1 + \frac{1}{k}) \tag{60}$$

and the variance:

$$\text{var}(X) = A^2[\Gamma(1 + \frac{2}{k}) - (\Gamma(1 + \frac{1}{k}))^2], \tag{61}$$

the goal would be to minimize the error through a root-mean-squared error (RMSE) method.
However this did not work out well as the resulting distribution mostly had a strong asymptotic behavior at zero and also the measurement points, if proceeded like this,




(a) Expectation (b) Variance

Figure 12: Wind speed: statistics for different sample sizes/orders; MC(red), SC(blue)

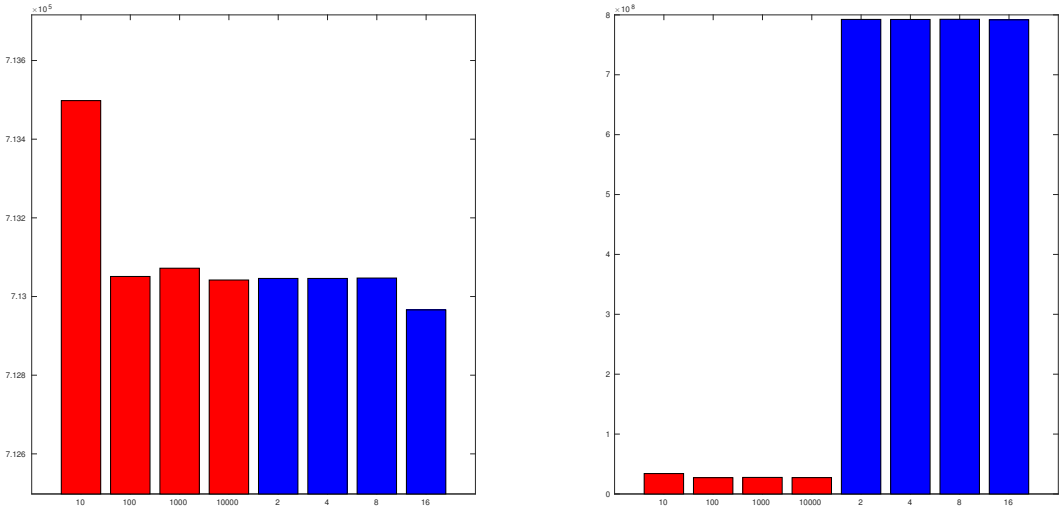
would open a huge multidimensional probability space. To work around this problem, the premise of a Weibull distributed RV was dropped and a normal distribution was assumed with $\mathcal{N}(0, 0.054)$. In the MC case this distribution is then uniformly applied to the complete velocity vector, simply shifting all velocities. The obtained results for mean and variance can be seen in figure 12. Mean and variance both converge to the same value, even tough the SC methods needs fewer points to achieve this.

### 6.1.2. Power curve

The second uncertain parameter to examine is the power curve (see figure 2a). For the power curve it is assumed, that the uncertainties are normally distributed and that the provided data is 99% accurate, thus $\mathcal{N}(0, 0.01)$. The data was also given by the industry, but is confirmed by Murcia et al. [9].
For the Park model the results can be seen in figure 13. The mean values are almost equal, but the variance differ by a complete order of magnitude. It is unclear why, but increasing the SC order did not improve the solution. The same Hermite polynomials have been used as for any other normally distributed uncertainty in this section.

### 6.1.3. Thrust curve

The last uncertain parameter is the thrust curve. Little to no data was found to model this uncertainty, but as the thrust curve is directly linked to the velocity deficit (equation 7), it was modeled likewise. Again by the provided data from the industry and the assumption of a normal distribution, it is $\mathcal{N}(0, 0.06)$. The plots for mean and variance can be seen in figure 14. Here mean and variance are in the same order of




(a) Expectation

(b) Variance

Figure 13: Power curve: statistics for different sample sizes/orders; MC(red), SC(blue)

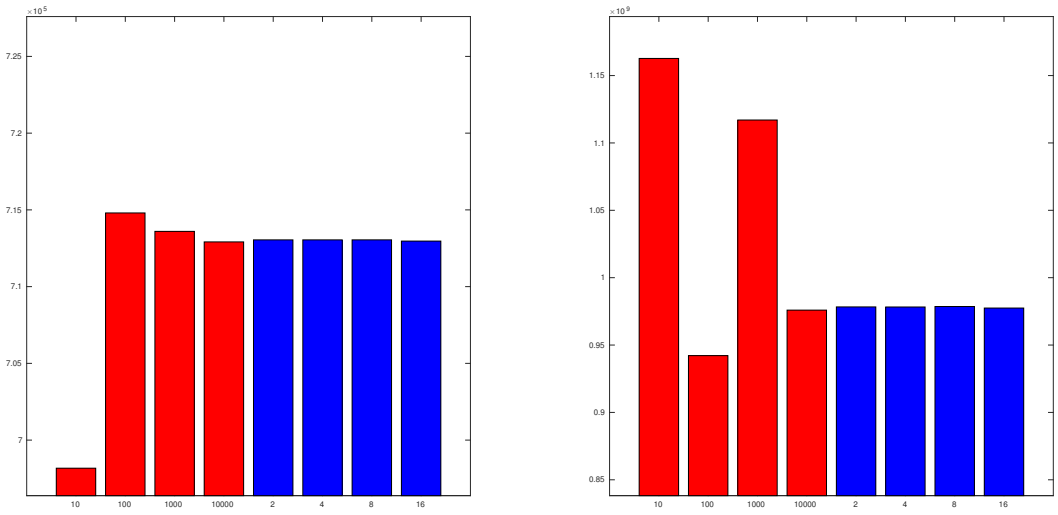
| samples | $t_{MC}$ [s] | $\mu_{SC}$ [GWh] | | order | $t_{SC}$ [s] | $\mu_{SC}$ [GWh] |
|---|---|---|---|---|---|---|
| 10 | 0.558 | 713498 | | 2 | 0.198 | 712992 |
| 100 | 4.32 | 713051 | | 4 | 0.350 | 712975 |
| 1000 | 40.947 | 713072 | | 8 | 0.373 | 712960 |
| 10000 | 408.97 | 713042 | | 16 | 0.715 | 712874 |

Table 6: Runtime comparison between MC and SC

magnitude and both seem to converge towards the same value.

## 6.2. Runtime analysis

Lastly the runtimes of both presented methods will be compared. All calculations were executed on the same machine, featuring an x86-architecture processor with 4x4.0 GHz and 8 available threads. Both methods scale well with the processors core count, as evaluations of each deterministic solve can be handled separately. Table 6 lists all runtimes for different numbers of sampling points/orders for the wind speed evaluation on the PARK model. It can clearly be determined, that for a single probability dimension, the convergence rate of the SC method is far superior for low dimensions.




(a) Expectation (b) Variance

Figure 14: $C_t$ curve: statistics for different sample sizes/orders; MC(red), SC(blue)

# 7. Conclusion

## 7.1. Summary

To sum up the thesis, two methods for quantifying uncertainties have been introduced. Both methods have then been applied to the univariate Burgers equations and each advantages and disadvantages have been highlighted for this suitable test case.
The Monte-Carlo method was shown to be able to capture more properties of a given problem, but compared to the stochastic collocation having a slow convergence for small dimensions. Vice versa, the collocation method is computationally much cheaper, but ultimately suffers from the curse of dimensionality.
In order to also be suitable for medium to high dimensions, Smolyak sparse grid were introduced. These sparse grids were also implemented and verified to be working correctly, but due to problems later on, they were not tested on the WFLOP problem, where they would have been needed.
In the end, problems with the deterministic solver and also with the implementation of the multivariate problem, left the work unfinished and lead to a foundation for future work.

## 7.2. Outlook

In a future project, fixing the eddy-viscosity wake simulation and implementing the multivariate cases of the WFLOP problem would be the next step. Additionally, other uncertainties, such as the surface roughness, air density and the hub height are interesting candidates to investigate. Especially the hub height is of great interest as it is modeled as constant, like it is only viable for onshore wind farms.

# A. Simulation code input parameters

In the following table the input parameters that are required for the simulation code are listed.

| Section | Parameter | Default value | Description |
|---|---|---|---|
| SITE | surface_roughness | 0.0002 | |
| SITE | turbulence_intensity | 0.1 | |
| WINDDATA | winddata_file | winddata.csv | |
| TURBINE | turbine_parameter_source | 0 | 0 for ini and 1 for xml |
| TURBINE | diameter | 80 | if source is json |
| TURBINE | hub_height | 80 | if source is json |
| TURBINE | cut_in_speed | 3 | if source is json |
| TURBINE | cut_out_speed | 25 | if source is json |
| TURBINE | ct_curve | 0.5 | if source is json. List with speed and $C_t$ values |
| TURBINE | power_curve | 0.5 | if source is json. List with speed and power values (in MW) |
| TURBINE | rpm_curve | 0.5 | if source is json. List with speed and RPM values |
| TURBINE | openwind_xml_file | turbine.owtg | if source is xml |
| SIMULATION | objective_function | 0 | 0 for AEP, 1 for efficiency, 2 for LCOE |
| SIMULATION | wake_model | 2 | 0 for no wake model, 1 for PARK, 2 for modified PARK, 3 for eddy viscosity |
| PARK | calculate_wake_decay | false | if PARK or modified PARK |
| PARK | wake_decay_constant | 0.07 | if PARK or modified PARK and not calculate wake decay |
| PARK | openwind_intersection | true | if PARK or modified PARK. Intersection of wakes is calculated with the maximum instead of the sum of squares, see eq. (**??**) |
| EDDY_VISCOSITY | discretization_scheme | 1 | if eddy viscosity. 0 for backward Euler, 1 for Crank-Nicolson |
| ... | | | |

| ... | | | |
|---|---|---|---|
| **Section** | **Parameter** | **Default value** | **Description** |
| EDDY_VISCOSITY | solution_algorithm | 1 | if eddy viscosity. 0 for Gaussian elimination, 1 for Thomas algorithm (tridiagonal solver) |
| EDDY_VISCOSITY | near_wake_filter | 1 | if eddy viscosity. 0 for no filter, 1 for original filter, 2 for linear, 3 for exponential |
| EDDY_VISCOSITY | max_wake_length | 50 | if eddy viscosity. In turbine diameters |
| EDDY_VISCOSITY | ignore_deficit_below | 0.0002 | if eddy viscosity |
| EDDY_VISCOSITY | x_stepsize | 0.1 | if eddy viscosity. $\Delta x$ in $D$ |
| EDDY_VISCOSITY | r_stepsize | 0.1 | if eddy viscosity. $\Delta r$ in $D$ |
| EDDY_VISCOSITY | calculate_wake_turbulence_increase | true | if eddy viscosity |
| EDDY_VISCOSITY | rotor_quadrature_points_d | 6 | if eddy viscosity. Number of quadrature points $N_r$; must be even |
| EDDY_VISCOSITY | rotor_quadrature_points_phi | 6 | if eddy viscosity. Number of quadrature points $N_\varphi$ |

Table 7: Input parameters for the simulation code

# References


[1] John F Ainslie. Calculating the flowfield in the wake of wind turbines. *Journal of Wind Engineering and Industrial Aerodynamics*, 27(1):213–224, 1988.

[2] John Burkardt and Clayton Webster. Slow exponential growth for clenshaw curtis sparse grids, 2012.

[3] Jungchul Choi and Martin Shan. Advancement of jensen (park) wake model. In *Proceedings of the European Wind Energy Conference and Exhibition*, pages 1–8, 2013.

[4] Gregor Heiming. Modeling and simulation of offshore wind farms. 2015.

[5] Niels Otto Jensen. *A note on wind generator interaction*. Number 2411 in Risø-M. 1983.

[6] Emil Brandt Kærgaard. Spectral methods for uncertainty quantification.

[7] I Katic, J Højstrup, and NO Jensen. A simple model for cluster efficiency. In *European Wind Energy Association Conference and Exhibition*, pages 407–410, 1986.

[8] Matthew A Lackner, Anthony L Rogers, and James F Manwell. Uncertainty analysis in wind resource assessment and wind energy production estimation. In *45th AIAA Aerospace Sciences Meeting and Exhibit, AIAA-2007-1222*, 2007.

[9] Juan Pablo Murcia, Pierre-Elouan Réthoré, Kurt Schaldemose Hansen, Anand Natarajan, and John Dalsgaard Sørensen. A new method to estimate the uncertainty of aep of offshore wind power plants applied to horns rev 1. In *EWEA Annual Conference and Exhibition 2015*, pages 161–165, 2015.

[10] A Peña and O Rathmann. The atmospheric stability dependent infinite wind farm and wake decay coefficient. *Wind Energy, in review*, 2011.

[11] Janet L Sawin, Freyr Sverrisson, Kristin Seyboth, Rana Adib, Hannah E Murdock, Christine Lins, Adam Brown, Stefanie E Di Domenico, Daniele Kielmanowicz, Laura E Williamson, et al. Renewables 2016 global status report. key findings. a record breaking year for renewable energy: New installations, policy targets, investment and jobs. mainstreaming renewables: guidance for policy makers.

[12] Ralph C Smith. *Uncertainty quantification: theory, implementation, and applications*, volume 12. SIAM, 2013.

[13] AWS Truepower. Openwind theoretical basis and validation. *Albany, NY, Technical Report*, 2010.