A POD-Based Reduced Order Model for Offshore Wind Applications, using LiDAR Measurements and WRF-PALM Simulations

Martine Rønning

Master's thesis in Energy

DEPARTMENT OF GEOPHYSICAL INSTITUTE
UNIVERSITY OF BERGEN

June 2, 2022

Acknowledgments

Firstly, I would like to acknowledge my supervisor Mostafa Bakhoday Paskyabi for providing me with an interesting thesis topic that is highly relevant to the industry, which I am about to enter. Thank you for enabling the study by providing me with cup anemometer and wind vane data. I greatly appreciate the motivational and insightful meetings, and all ideas and support during the work. I also wish to thank my co-supervisor Maria Krutova for providing me with WRF-LES simulations and retrieved LiDAR data, as well as insight regarding the data sets. Thank you for your patience and guidance with codes in python for LiDAR processing. Thank you Joachim Reuder, for profound discussions regarding the atmospheric boundary layer, and for helping me gain a deeper understanding of wake and turbulence.

Writing a master thesis for one year has been one of the greatest study-related challenges of my life. I would like to thank my family and friends for being supportive throughout the year. Finally, I would like to express my sincere thanks to my fellow students, and especially Tore, Adrian, Tiril, and Christina, for daily coffee breaks, study-related discussions, and for taking my mind off the thesis with good conversations and quizzes. You have been a great motivation for me.

The work is a part of the HIghly advanced Probabilistic design and Enhanced Reliability methods for the high-value, cost-efficient offshore WIND (HIPERWIND) project, which has received funding from the European Unions Horizon 2020 Research and Innovation Programme under Grant Agreement No. 101006689. The simulations were performed on resources provided by UNINETT Sigma2 - the National Infrastructure for High Performance Computing and Data Storage in Norway. Some data used in this study were gathered as part of the OBLEX-F1 field campaign that has been performed under the Norwegian Centre for Offshore Wind Energy (NORCOWE) funded by the Research Council of Norway (RCN 193821), the Offshore Boundary Layer Observatory (OBLO) project (project no. RCN 227777) and the Norwegian e-infrastructure NorStore (project no. RCN: NS9060K). OBLEX-F1 was coordinated in collaboration between the University of Bergen (Geophysical Institute) and NORCE Norwegian Research Centre (project executing organization). The Federal Maritime and Hydrographic Agency of Germany (BSH) is acknowledged for providing the FINO1 reference data through the FINO database at http://fino.bsh.de/. The FINO project (research platforms in the North Sea and Baltic Sea) is funded by the BMU, the German Federal Ministry for the Environment, Nature Conservation, Building and Nuclear Safety in collaboration with Project Management Jülich GmbH (project no. 0325321). The FINO1 meteorological reference data were provided by Deutsches Windenergi In-

stitute (DEWI). We also thank the FINO1 platform operator Forschungs- und Entwicklungszentrum Fachhochschule Kiel GmbH (FuE Kiel GmbH) for their support (project no. 0329905E).

Martine Rønning
Bergen, June 2, 2022

Acknowledgments

Abstract

Wind flow fields within a wind farm, including wake dynamics, are a complex system of large degrees of freedom. The wake shape and velocity deficit downstream of a turbine may be calculated using numerical modeling, such as Large Eddy Simulations (LES). However, these estimations are computationally expensive and time-consuming. On the other hand, analytical models may provide efficient computations but generally exclude several features of the wake dynamics. Reducing the complexity of the numerical models and improving the details of the analytical models are highly relevant issues for today's wind farm controlling and layout designing.

This work proposes a methodology applicable to the industry to conduct wind farm flow field calculations by investigating the wake dynamics based on applying Proper Orthogonal Decomposition (POD) to Light Detection And Ranging (LiDAR) measurements and Weather Research and Forecasting (WRF) data using a Parallelized Large Eddy Simulation Model (PALM). Both data sets comprise a complex, high-dimensional system consisting of an area of Germany's first offshore wind farm, *Alpha Ventus*, located in the North Sea. 10 days of wind speed and direction data are retrieved from the radial velocity measured by the LiDAR at the FINO1 platform, located in close proximity to the wind farm Alpha Ventus, between September and October 2016, during varying atmospheric forcing conditions. The WRF-PALM data were simulated for one hour on the 21^{st} of September 2015, using ERA5 data as input, during unstable conditions.

Reduced Order Models (ROMs) are built separately for both the LiDAR and WRF-PALM data, by decomposing them into a number of time-dependent, truncated weighting coefficients and spatial orthogonal basis functions. Proper Orthogonal Decomposition (POD) is shown to reconstruct selected wind fields in a reduced manner while preserving the global patterns of the wind fields for both LiDAR and WRF-LES data.

The study has further investigated the ability of the Gaussian Process (GP) to incorporate unresolved small-scale wake structures in the reconstruction that are excluded by the truncated ROM. A sensitivity study for a variety of kernels accompanied by hyperparameters is conducted. By replacing the temporal weighting coefficients obtained for the POD with stochastic weights obtained from the GP, the study has shown that the reconstruction is sensitive to kernel selections. By reconstructing the field using both the weighting coefficients from the standard POD and those obtained using GP, both for the retrieved and WRF-PALM data, the performance of the methods has been evaluated based on visual inspection, energetic contribution, and Root Mean Square Error

(RMSE).

Abstract

Contents

Acronyms and Abbreviations

ABL	Atmospheric boundary Layer
AEP	Annual Energy Production
ALR	Adiabatic Lapse Rate
ASL	Above Sea Level (a.s.l.)
BL	Boundary Layer
CFD	Computational Fluid Dynamics
FINO1	Forschungsplattformen in Nord- und Ostsee Nr.1
GP	Gaussian Process
GPR	Gaussian Process Regression
IEA	International Energy Agency
L(A)BL	Land (Atmospheric) Boundary Layer
LES	Large Eddy Simulation
LiDAR	Light Detection And Radar
M(A)BL	Marine (Atmospheric) Boundary Layer
ML	Machine Learning
NaN	Not a Number
OCC	Open Cell Convection
POD	Proper Orthogonal Decomposition
RAVE	the Research at Alpha Ventus
ROM	Reduced Order Model
SCADA	Supervisor Control and Data Acquisition
VVP	Volume Velocity Preprocessing
WRF-PALM	Weather Research and Forecasting - PArallelized LES Model

Acronyms and Abbreviations

Chapter 1

Introduction

1.1 Background

Through a series of climate conferences, various agreements and goals have been developed to reduce human impact on the climate. One, and perhaps the most important goal, was developed in Paris in 2015, stating that the average global temperature rise should be kept well below $2°C$, ideally at $1.5°C$, compared to pre-industrial levels [1]. Moreover, International Energy Agency (IEA) stated that in order to keep global warming to no more than $1.5°C$, emissions must reach net zero by 2050 [2, 3].

Today we are at $1.1°C$ [3]. Several measures must be taken to reach such goals, and wind energy is expected to play a major role in achieving them. One fundamental explanation for this is that wind is abundant, renewable, and has the potential to deliver relatively large amounts of clean energy. The wind energy sector has already faced tremendous growth during the past years. In 2001, the global cumulative installed wind energy capacity was approximately 24MW versus 743MW in 2020 [4]. Still, this only accounted for 6% of the world's energy production in 2020, where well above half came from China and the US alone. IEA has estimated that an average increase in wind energy generation of 18% every year from 2021 to 2030 is required to meet the net-zero power generation level [5].

Wind energy is often generated by clustering several turbines together in a wind farm. For such arrangements, the aerodynamic wake effect is known as one of the major power losses. The concept of wake has been known for many years, and in line with installed wind energy capacity, the amount of research on wake has increased as well (e.g. [6–12]). Wake induces turbulence and suppresses momentum, and therefore impact loading on the downwind turbines and power output of the wind farm. It has been showed that the wake effect may account for up to 20% energy loss from a wind farm [13–15]. Hence, wake is crucial to account for in wind farm layout optimization and controlling [16–19]. Wake is, however, not a straightforward system to analyze but rather a complex, non-linear system of interactions between atmospheric conditions, turbine characteristics, and control mechanisms of the turbine [16, 20]. These features

make it challenging to conclude its behaviour[1].

Through years of research, methods to study the wakefield within a wind farm have evolved [9, 13, 21–25]. These can be divided into analytical, experimental, and numerical techniques [8]. However, they are all somewhat deficient in providing efficient and precise estimations. The low-fidelity analytical methods provide efficient predictions but tend to underperform when it comes to accuracy due to simplifications and assumptions. On the other hand, high-fidelity models provide accurate estimations, including wake dynamics and relatively small and large-scale turbulence structures. However, real-life problems of this kind seldom propose solutions within a satisfactory time frame, using numerical techniques such as those under consideration. As optimization processes rely on numerous computational iterations for a potentially extensive range of parameters in e.g. design and operation calculations, today's industry heavily relies on analytical models to solve these issues [26]. In an attempt to bridge between the methods mentioned for wake prediction, Proper Orthogonal Decomposition[2] (POD) is proposed to build a Reduced Order Model (ROM) with the aim to simplify the dynamics while preserving the global behavior of the real, high-fidelity system in a wind farm.

There are several ways to create a ROM, and POD is a widely used technique in fluid dynamics, e.g. [10, 12, 27–35]. Building a POD-based ROM is achieved by utilizing a number of spatial modes and temporal weighting coefficients obtained through POD to construct a ROM of truncated order. In this process, dominant features of the wind field are isolated. Two aspects make ROM interesting to investigate further. Firstly, most POD-based research for wind energy applications is combined with high-fidelity simulations such as LES (e.g. [10, 12, 31, 33, 35, 36]). Common for these studies is that in order to determine the global patterns of the system, high-fidelity data ("snapshots") is necessary as input to the POD which is typically obtained by numerical modeling [10, 12, 31, 33, 36]. Being able to avoid the initial numerical simulation, one is not equipment dependent[3], and one can reduce the costs and potentially obtain more efficient results while preserving a relative precision. Secondly, by excluding some of the ROM modes describing the wind field, small-scale dynamics representing turbulent structures are often excluded from the ROM. Hence, the ROM reconstructs a smooth version of the original wind field. It is interesting to investigate the opportunities of including small-scale dynamics in the ROM while preserving the efficiency of the model. In this work, the widely studied Gaussian Process (GP) is investigated for this purpose [37–40].

Therefore, the main question is how the widely-used POD approach performers when applied to LiDAR scans and whether it is a procedure enabling the industry to study wake dynamics and velocity deficits more practically and efficiently than the widely used analytical models currently do. Moreover, is the ability to include small-scale dynamics into the reconstruction fields obtained from the ROM using GP studied. In the present work, the performance of POD both on high-fidelity WRF-PALM and high-

[1] Behaviour in this context includes, e.g. wake meandering, evolution, recovery, velocity deficit, etc.

[2] Also known as Principle Component Analysis (PCA) and Karhunen-Lo'eve decomposition [26].

[3] If the simulations are not of the very simplest manner, supercomputers etc. are often required to conduct the simulation.

resolution LiDAR data will be investigated.

1.2 Motivation and Contribution

The motivation for initiating this work is the desire to further improve today's methods for wake calculations, with particular regard to the industry. Enabling efficient and relatively precise[4] wake calculations are studied by investigating whether it is possible to bridge between the complex numerical methods and the simplified analytical and low-fidelity, experimental methods. By improving the available methods in the industry, one can further improve the wind farm layout optimization and controlling processes, which is beneficial for the wind energy industry.

POD is an extensively used method in the wind energy discipline and has proven to reduce the complexity of the complex wake dynamics [28, 31, 33, 36, 41–45]. However, the majority of the work suggests the research of POD applied to LES simulations. The LES research enables one to find key features of the wake meandering, evolution, and velocity recovery in general. The method is, however, only to a limited extent valuable to industry, where efficient analyses are required. To our knowledge, the technique had not been widely applied to LiDAR measurements when this thesis was initiated and was proposed as a future study by Bastine et al. [31]. The LiDAR makes it possible to remotely measure the wind field at relatively high accuracy and spatial resolution [46]. In order to drive down costs of the (offshore) wind farm and meet the net-zero targets, (offshore) turbines are upscaled at a rapid rate [47–49]. This further complicates measuring in-situ, and thus LiDAR measurements are continuously becoming more relevant to the investigation.

Therefore, the novelty of this work is the study of POD on LiDAR data. By applying POD directly on processed LiDAR data, without carrying out a simulation, one can save time and money, which generally is more applicable in the wind energy industry. Thus, it is interesting to investigate how the well-documented POD technique performs applied to LiDAR measurements. As POD applied to LES data is an already proven methodology, WRF-PALM data is used in this work for the sake of comparison. The findings obtained from the WRF-PALM data can be extended to the LiDAR data. The results provide further insight into the widely utilized POD techniques and results applied to measurements obtained from the emerging LiDAR instrument. The works suggest a methodology for the industry to gain further insight into the wake dynamics in the specific wind farm investigated, further enabling optimization processes to be improved.

As the truncated standard POD excludes small-scale motions of the real-life wake flow [33], it is further investigated ways to include these features to improve the reconstructed wind field by the standard POD. The work contributes to a method for fast and efficient wake predictions using LiDAR measurements from an offshore wind farm. ROMs are developed by reducing the complexity of the wind field using POD. ROM can further be utilized to optimize wind farm layout and operation, which again could

[4] i.e. including wake dynamics such as wake meandering, which are not present in analytical methods of today.

result in reduced loading on downstream turbines and increased power output.

1.3 Objectives

This work aims to investigate the performance of the well-established POD technique on LiDAR measurements obtained from an offshore platform in the North Sea and compare it with its performance on high-fidelity WRF-PALM data. The method aims to efficiently identify spatial modes, which yields a low order description of the full order model without compromising vital (e.g. meandering) features of the wake behavior. Reducing the dimensions of the wind field and incorporating the ROM into current wind farm optimization processes could enhance the understanding of the wake effect. Hence, this could result in improved power production and turbine loading estimations. It appears to be the ultimate method for gaining relatively accurate and fast wake predictions. However, the extent to which it applies to LiDAR data snapshots and whether it improves the accuracy of the simple analytical techniques, and the computational time for the complex methods, remains to investigate. Based on the above considerations, the main objective is to investigate how a ROM performs for wake estimations using LiDAR measurements from the Forschungsplattformen in Nord- und Ostsee Nr.1 (FINO1) offshore wind site.

Secondary objectives are:

- Conduct quality control and preprocess wind speed and direction data, which are retrieved from radial velocity obtained with a LiDAR located at the FINO1 platform in the North Sea.
- Build ROMs by applying POD to both retrieved and WRF-PALM data measured at the same offshore wind farm, Alpha Ventus (different time and subdomain).
- Improving the standard POD by altering the time-dependent weighting coefficients in the reconstruction of the wind fields. This is done by including stochastic features by utilizing the GP.

1.4 Limitations

The limitations of this work are listed below.

- Geographically limited to the North sea, and the conditions of the location (wind/wave flows, ambient atmospheric stability, cells of the earth, northern hemisphere, etc.).
- Temporal constrained to 24.09.2016 - 03.10.2016[5] for the LiDAR, cup anemometer, and wind vane data, and one single hour at 22.11.2015 for the WRF-PALM

[5]Temporal constrained to 02.10.2016 for SCADA.

data. Therefore, also limited to climatic events/conditions prominent in this period.

- Assumption that the cup anemometer and wind vane are not distorted
- Temporal and spatial resolution of all data used.
- Errors or downtime of the LiDAR instrument causing incorrect or insufficient data. E.g. is every 25^{th} scan removed due to a 40 minute gap in the data for every 20 minutes of scanning. Moreover, is the LiDAR directed in a way causing many incidents of cross-wind, resulting in a high number of bad-quality data.
- LiDAR and WRF-PALM snapshots in the period where the turbines are not operating (4m/s for Adwen AD 5-116 and 3.5m/s for Senvion 5M) are not removed, which may impact the outcome of the ROMs.
- Unprocessed SCADA, cup anemometer, and wind vane data (Met-mast effects of the cup anemometer and/or wind vane used for retrieval and validation. Rotor effects and neighboring turbine wake effects on the SCADA data).
- The displacement of the various measurements in relation to each other, resulting in the assumption that the conditions at one location hold for the other locations. SCADA data is obtained from turbine 7, and LiDAR data is measured over an area from FINO1, cup anemometer, and wind vane data is acquired at FINO1. E.g. the conditions at FINO1 are not necessarily representative of the conditions in the wind farm.
- Errors in the preprocessing of the LiDAR data. E.g. bad-quality scans are kept because the algorithms developed do not identify them as bad.
- Error in python codes.
- The results for the ROM are restricted to the snapshots utilized in this study. The possibility that the performance of POD deviates from one snapshot to another is present.
- The sensitivity study for kernels is restricted to the kernels tested and the hyper-parameters chosen. The results obtained for the stochastic POD are restricted to the kernels and hyper-parameters as well.

1.5 Thesis Outline

The thesis is distributed over five chapters. In chapter 2, the theoretical framework of the thesis is presented. It covers relevant background about wind energy, environmental conditions in a wind farm, including wake flow evolution and meandering, atmospheric boundary layer, stability regimes, transient event characteristics, and differences in offshore and onshore boundary layer conditions. At the end of the chapter, wake prediction methods of today are presented, and fields within the wind industry that would benefit from improvements in these methods are explained briefly. Chapter 3, firstly, describes

the site investigated in the work, before elaborating on the various data sets used. Finally, the methods used are ascribed, including standard POD and the stochastic model utilized. The results are presented in the following chapter 4, which is compared and discussed in chapter 5. The discussion forms the basis for the conclusion of the thesis, which is given in chapter 6, and future work is suggested.

Chapter 2

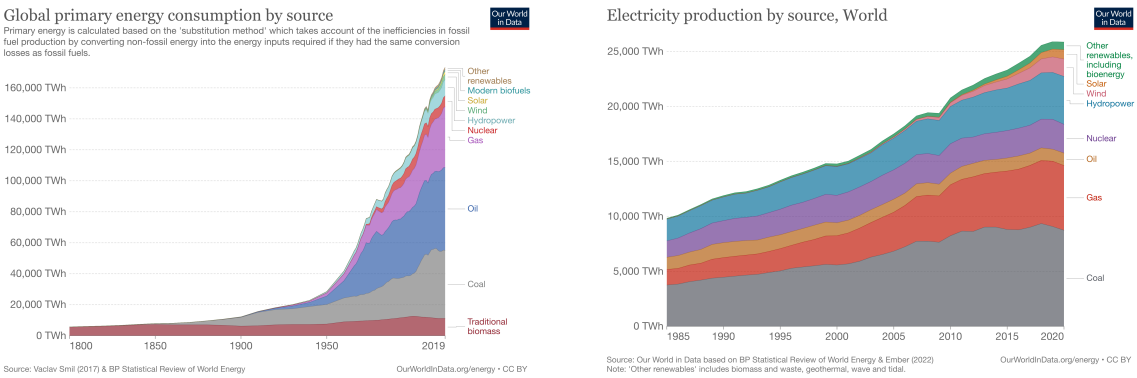
Background

As detailed in this chapter, wind speed is a key factor in wind farm power production. Hence, it is beneficial to gain a thorough understanding of the factors that influence it and how it interacts with turbines and moves within the wind farm.

The first section of the chapter briefly elaborates on the current role of wind energy in the world before a brief introduction to wind energy is presented. Subsequently, we delver into environmental conditions, consisting of seven subsections. We start by elaborating on wake and turbulence, two critical features within the wind farm of interest to this study. In order to gain a comprehensive understanding of these features and how they behave, we describe the wind profile, the atmospheric boundary layer and -stability, transient event, and finally, offshore in relation to onshore conditions next. These topics all affect wake propagation and recovery in some way. We complete the chapter by briefly describing the methods for calculating wake today and the areas where the current study may contribute with valuable insights.

2.1 Global Energy Consumption

While the importance of utilizing renewable versus non-renewable energy sources is constantly growing [50], so is the energy consumption, as illustrated in figure 2.1a. In order to meet the net-zero goals developed as a result of the Paris agreement [1] in 2015, an energy transition is required, and various road maps have been established to achieve it [51]. However, as IEA points out, reaching netzero emissions by 2050 is a monumental task [2], with numerous challenges. One of these challenges is the fact that the widely used non-renewable energy sources of today have a high energy density, often hard to match with renewable sources [52, 53]. Thus, for renewable energy sources to replace the current production of oil, gas, and coal, more power plants are required to provide the same amount of energy.



(a) Global energy consumption per year *(b) Global electricity production per year*

Figure 2.1: Global energy consumption and electricity production, figures created by [54].

When considering the numbers from the figure 2.1 [54], one can find that renewable energy sources accounted for about 27% of the total energy consumption in 2019. In contrast, the remaining 73% were covered by coal, gas, and oil. According to IEA's road map to net-zero emissions by 2050, nearly 90% of global electricity must come from renewable energy sources in 2050, with solar and wind energy accounting for almost 70% [2]. This necessitates comprehensive and costly infrastructure changes [2]. Furthermore, several of the emerging renewable sources are not yet well-established and mass-produced, making them more expensive than the well-established non-renewable energy sources. Another consideration is that solar and wind energy, which are expected to account for nearly 70% of the total energy in 2050, are intermittent resources. For this reason, it is not only necessary with renewable power plants, but also energy storage facilities to meet the energy demand at all times [2]. In conclusion, one can say that it is not a trivial task we are facing. However, one thing that is for certain is that wind energy will play an essential role in the energy transition.

2.2 Wind Energy

Wind energy is a renewable energy resource exploited by utilizing a turbine to generate electricity from the kinetic energy in the wind. The wind is an abundant resource, free to use worldwide. Delivering emission-free electricity is crucial given the current climate situation of the earth previously noted [55, 56]. Although the use of wind energy seems trivial, it poses several challenges in terms of being done in an economically, socially, and environmentally friendly way [55].

The scale on which wind energy power plants are developed increases, and for this reason, turbines are often clustered together in wind farms. The wind farm can be regarded as a system, with the entire farm situated in an atmosphere with an associated stability. As wind enters the system, the goal is to extract energy from the turbines. When the wind evolves through the system, its natural flow is affected by the ambient stability and turbines in the system. The turbines act as blockages for the wind, and energy extraction induces wake. The wake causes a velocity deficit and increased turbulence intensity in the wind flow downstream of the turbines. The concept has been known for

decades, but as can be seen in section 2.1, it has yet to be implemented on a large-scale globally [56, 57].

Thus, the energy production in a wind farm is directly related to the amount of energy in the wind, partly determined by the velocity of the wind. It is known from physics that kinetic energy is given by $E_k = 1/2mU^2$, where m represents mass and U the velocity (of the wind in this case). By combining this knowledge with the formula for power, $P = dE_k/dt$, the power, P, produced by the wind turbine can be calculated as [56, 58]:

$$P = \frac{1}{2}\rho A U^3 C_p, \tag{2.1}$$

where ρ is the air density, A is the swept area by the rotor, U is the wind speed (velocity) perpendicular to the swept area, and C_p is the power coefficient[1]. As indicated by the equation, the power output is proportional to the cubed wind speed[2]. Hence, any change in wind speed has a significant impact on total power production. The wind has great freedom to vary independently of the surrounding wind flows, and is therefore chaotic and frequently fluctuating, both on a temporal and spatial scale. It is affected both by local and global wind systems, ranging from winds at a planetary scale, through synoptic and mesoscale, to microscale [55]. As mentioned, one such factor impacting the wind velocity in a wind farm is the wake effect. It is crucial to understand the characteristics of the wind resources and the drivers behind the wind flow, in any phase of wind energy development [55, 57].

2.3 Environmental Conditions in the Wind Farm

The wake of interest in this work is the one that occurs as the wind hits a turbine. On this basis, it is intuitive that the turbine itself and the incoming wind are key components for understanding wake evolution. Like most renewable energy sources, wind energy is driven by the sun. Thermal gradients are formed due to the uneven solar radiation of the earth's surface, generally causing air masses to move from cold toward warmer areas. Thus, the incoming wind field is greatly affected by the atmospheric conditions[3] [59] and the topography of the site [60, 61]. The increasing rotor size of the modern turbines makes it even more important to investigate the atmospheric conditions in which they will operate, as these may vary significantly from one rotor tip to the other [20, 62].

2.3.1 Wake

The downstream wind flow of a turbine experiences two physical phenomena of interest; increasing turbulence and decreasing velocity. The phenomenon is known as the wake effect. In other words, as the wind passes the turbine, some of the wind flow changes state to become turbulent (see chapter 2.3.2), and a velocity deficit occurs when the turbines extract some of the wind energy. In a wind farm, turbines are

[1] The power coefficient measures how efficiently a turbine converts kinetic energy in the wind into electricity.

[2] Energy output is proportional to the wind speed squared.

[3] A case detecting 20% power production difference for various stability classes is found in [59].

often clustered together with less distance than what is required for the wind to recover to free stream (see more in section 2.4.2), and they are therefore greatly exposed to the wake effect [21]. The velocity deficit and turbulence affect the power output and fatigue load, respectively, on the downstream turbines [7, 11, 18, 20, 36, 63].

As seen in section 2.2, the power extracted from the turbine is greatly affected by the incoming wind speed. As the wake reduces wind speed, it impacts the performance of the downwind turbine when operated in the wake. Moreover, as seen in equation (2.1), a relatively small reduction in wind speed caused by wake, can reduce the power output significantly due to the cubed relationship. Wake is considered one of the major energy losses in a wind farm and has been found to account for up to 20% of the total energy loss [14, 15, 64]. Even power losses of as much as 28% have been detected [21]. By considering the wake effect in wind farm layout optimization and controlling, one is able to optimize and somewhat reduce the total power loss and loading it exerts on the downstream turbines. However, this depends on reliable wake models and calculations.

Wakes are complex, and the wake evolution depends on various factors, such as turbine size and design, wind inflow conditions, and turbine operation [65]. The area behind the rotor where the wake unfolds is typically divided into two regions, the near- and far-wake region[4] [66–68]. The near wake region is just behind the rotor, and the characteristics of the wind flow in this area are dominated by the rotor geometry[5]. The length of the region has been debated through previous studies but can be approximated to fall in the range from $1D$ to $5D$ downstream from the turbine [8, 9, 66, 67, 70–72]. In this area, phenomena such as decreased velocity, wake expansion, pressure gradient, and blade tip vortices dominate. The wake occurs within the free stream wind flow, and the result is a strongly sheared layer[6] separating them [69, 70]. Behind the rotor, will a combination of velocity reduction[7] and added wake turbulence[8] increase entertainment[9], causing the wake to gradually expand and transfer momentum into the wake [69]. In the far-wake region, which is seldom given a specific distance in literature [9, 67, 70], the wind flow is in a self-similarity state. Besides the wake added turbulence and velocity deficit, the flow is mainly dominated by wake meandering[10], whereas the turbine characteristics are limited. With increased distance from rotor, the energy-containing eddies[11] generated in the sheared layer is transferred towards smaller eddy scales (see figure 2.2), and with additional help from the ambient turbulence is the velocity deficit reduced, and eventually, free wind flow is recovered [9, 67, 70, 74].

As already touched upon, the size and evolution of the wake structure are defined by factors such as turbine design (e.g. thrust coefficient and therefore the ambient wind speed), turbine control settings, and ambient atmospheric conditions [25, 71]. The thrust coefficient of a turbine affects the amount of momentum possible for the turbine

[4] The wake region can be divided into finer regions, such as including an intermediate region.

[5] Geometries such as the number of blades, blade aerodynamics, pitch mechanisms, 3D effects, and tip vortices [69].

[6] Shear: describes the change of wind speed and direction from one point in the atmosphere to another.

[7] From the conservation of mass, we have that the area expands due to reduced velocity [7].

[8] Mechanical turbulence is also known as forced convection [73].

[9] wind flow from outside the wake mixes with the wake flow.

[10] Wake meandering: large-scale movements of the entire wake due to atmospheric/ambient eddies [42].

[11] An eddy denotes a turbulent vortex-like structure.

to extract, which again outlines the initial velocity deficit. In general, thrust coefficients are inversely proportional to wind speed. The higher the thrust force applied on the rotor, the lower the wake velocity, and the larger the shear [75]. Thus, wake velocity deficits are maximal at high thrust levels (low wind speeds) and in low ambient turbulence intensities (stable atmospheric conditions) [71]. Another feature possibly contributing to velocity deficit recovery is the wake meandering [70]. However, wake meandering might also considerably increase the loading on the downstream turbines [9, 70].

2.3.2 Turbulence

Turbulence is a well-known phenomenon discovered several decades ago and has been studied frequently since then, e.g. [62, 63, 66, 67, 76–78]. In this work, turbulence is of interest because it has shown that fatigue loading may increase significantly on the downwind turbine if operated in the wake of another turbine [70, 79]. Moreover, as seen in chapter 2.3.1, turbulence is crucial in understanding the evolution and dynamics of the wake effect. Thus, turbulence is key in wind energy, as it has the potential to increase loading on turbines and decrease wake recovery time [80]. Even though turbulence has been studied comprehensively throughout the past decades, it has proven to be a rather complex system to investigate. Thus, understanding and to some extent predict turbulence are still highly relevant topics [73, 81–83].

One can imagine turbulence as a superposition of various sized eddies interacting nonlinearly, resulting in chaotic, random motions [73, 78]. Turbulence occurs naturally as a response to instabilities in a flow[12], tending to reduce the instability and returning to equilibrium (neutral conditions). When the instability is no longer present, turbulence ceases [73]. The main characteristic of turbulence is that it is stochastic and consists of vorticial structures (eddies) in all three dimensions, and it has strong diffusivity (efficient mixing). Moreover, it is known for having a broad energy spectrum, transferring energy from large-scaled toward small-scaled eddies[13], where energy eventually is dissipated by viscosity (see figure 2.2). Lewis Fry Richardson put this nicely in 1922 as [62, 73]:

"Big whirls have little whirls that feed on their velocity, and little whirls have lesser whirls and so on to viscosity" [84, 85]

Turbulent describes the state of a fluid as it reaches a specific Reynolds number[14]. Very viscous or slow-moving flows are considered *laminar* and are characterized as calm, smooth flows. At some point, when the velocity increases in the laminar flow, the viscosity becomes relatively low, and the movements become chaotic and fluctuating, in other words, turbulent [62]. Turbulence is present in various ways in a wind

[12] Instabilities such as thermal gradient or shear.

[13] The size of an eddy may range from the size of an air molecule ($<$ 2mm) up to the planetary scale of 20 000km [73].

[14] Reynold found through an experiment in 1883 a critical Reynolds number for turbulence, $Re_{critical} > 2000$, meaning that turbulence cannot exist in a flow with Re $<$ 2000 [78](Can also be defined using the Richardson number [73]).

farm, which can be divided into wake added turbulence and ambient (atmospheric) turbulence. Wake added turbulence is e.g. turbulent vortices shed by the rotor blades and tips, mechanical induced turbulence by the tower and nacelle, and turbulence from the mean velocity shear of the wake [67]. Ambient turbulence is elaborated on later in section 2.3.

Hence, (mechanical) turbulence is induced as the wind passes through the turbine. With increasing distance from rotor, the kinetic energy in the turbulent flow is transferred towards smaller spatial scales, like L.F. Richardson explained it in 1922. The internal energy of the wake will increase steadily towards its initial state, as it was before it interacted with the turbine. In 1941, Kolmogorov presented a theory on this energy distribution with various eddy length scales in a turbulent flow [86]. The energy spectrum is frequently separated into three ranges of eddy scales referred to as the energy-containing, inertial, and dissipation ranges [87], as seen in the illustration 2.2. The eddies in the energy-containing range are largest in size, and turbulent kinetic energy (TKE) is generated at these scales. As the TKE is transferred to the smaller-scaled eddies, they eventually enter the inertial subrange. In this range, the turbulent eddies are affected by neither the effects of viscosity nor the generation of TKE [78]. Kolmogorov found that the energy of the eddies in this range would cascade in a characteristic $-5/3$ slope. The energy cascade continues into the range for the smallest eddies in the dissipative range, where the viscosity comes into matter and the energy eventually dissipates into heat [62, 88].

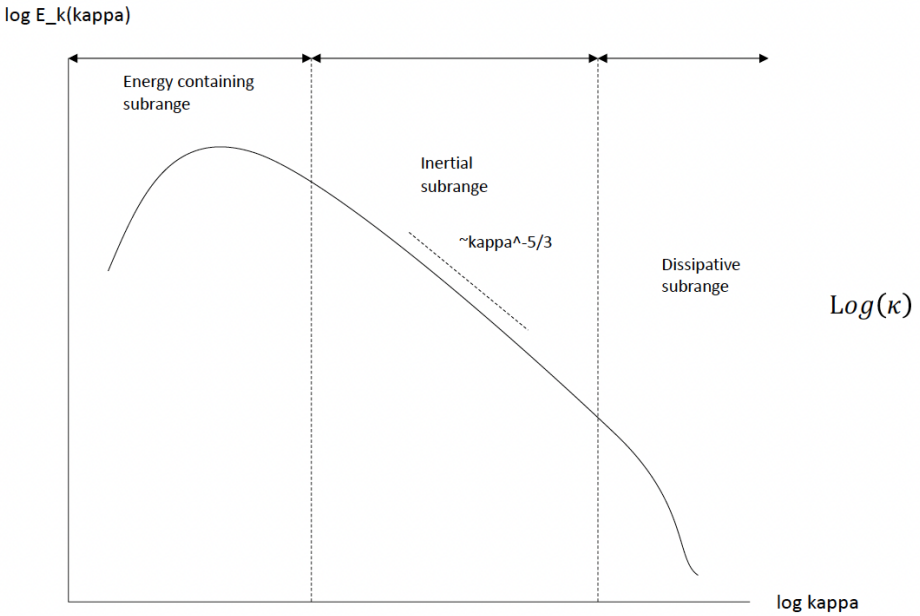


Figure 2.2: Kolmogorov energy spectrum, figure inspired by [78].

From fluid dynamics, it has been exhibited that the mathematical equations of Navier-Stokes, based on the conservation of momentum for an incompressible fluid, can be utilized to describe a general turbulent flow. These equations can be discretized and solved numerically. However, in most cases, this process is computationally demand-

ing. The reason is mainly the broad range of eddy sizes that need to be solved, and the process of transferring energy from large to smaller eddy scales is described through nonlinearities. The higher the Reynolds number of the turbulent flow, the smaller scales are the eddies where the energy is dissipated. As a result, the model may become very complex to solve for some specific flows. Today there exist methods to approximate solutions, such as Large Eddy Simulation (LES) and Reynolds Averaged Navier-Stokes (RANS) [62], or creating models based on simplified equations [11, 23, 89]. Even though turbulence has been investigated for several decades, no theoretical model describing the exact behavior of the phenomenon exists yet. Thus, it can be concluded that there is no coherent *theory of turbulence*, but rather various problems and theories [62, 73].

2.3.3 Wind Profile

Some parameters will have a so-called vertical profile in the atmosphere. This is the case for wind, temperature, and humidity as will be explained later in section 2.3.5 and 2.3.7, respectively. The wind profile[15] describes the variation of the mean wind speed with elevation. Due to friction on the surface, the wind experiences a drag force closer to the ground. Thus, the wind velocity is zero at the surface and gradually increasing with altitude in the ABL, as seen in figure 2.3. The wind profile is important in wind resource assessment and for turbine design [62, 68].

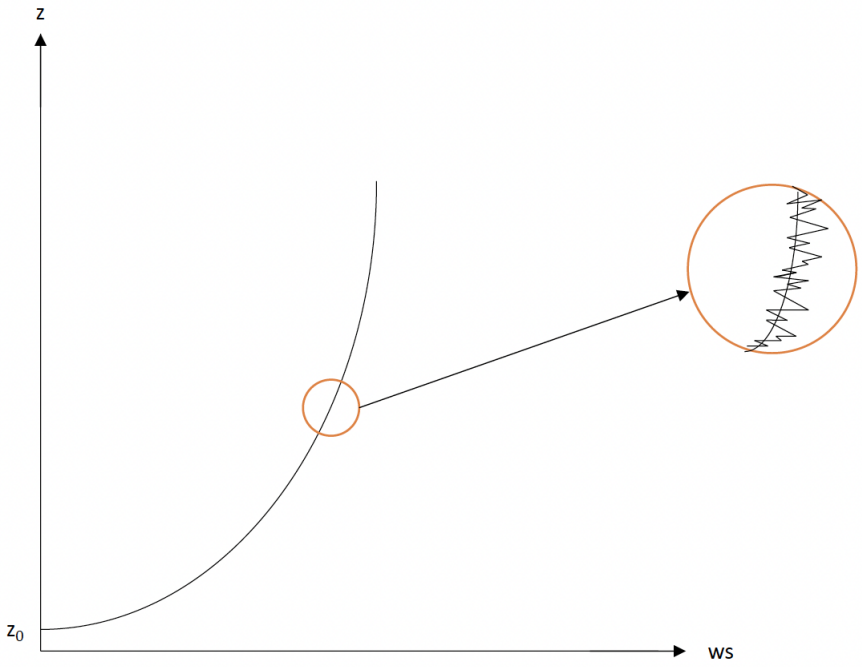


Figure 2.3: Illustration of a typical logarithmic wind profile in the surface layer. z_0 is the surface roughness length.

There are two common models for determining the vertical wind profile over flat terrain in the surface layer: the logarithmic law and the power law. The logarithmic law is

[15] The wind profile can also be referred to as wind shear [56].

based on fluid mechanic principles and can be described as in the following equation[16]:

$$u(z) = \frac{u_*}{\kappa} \ln\left(\frac{z}{z_0}\right), \tag{2.2}$$

where $u(z)$ is the velocity at height z, κ represents the von Karman constant, u_* is the friction velocity, and z_0 is the roughness length[17](see section 2.3.7). The equation (2.2) is valid for neutral conditions; however, it is possible to incorporate atmospheric stability into the model, resulting in the following equation:

$$u(z) = \frac{u_*}{\kappa}\left[\ln\left(\frac{z}{z_0}\right) + \Psi_s\left(\frac{z}{L}\right)\right], \quad z \gg z_o, \tag{2.3}$$

where Ψ_s is an atmospheric stability function dependent on z/L[18], and L is the Obukhov length [56, 68].

On the contrary, the power law is empirically derived and can be described as:

$$u(z) = u_r(z_r)\left(\frac{z}{z_r}\right)^{\alpha}, \tag{2.4}$$

where u_r is the reference wind speed, z_r is the reference elevation, and α is an empirical shear exponent, dependent on parameters such as time of day and year, terrain, and wind speed[19]. The model is a simple, more practical, and direct model compared with the logarithmic law. Thus, it is widely used by engineers to describe the vertical wind profile [56].

It has been demonstrated for flat terrains, such as an ocean surface, that these laws provide results with small differences. However, when the vertical wind profile in the surface layer is extrapolated into the Ekman layer, the power law is more suitable to utilize, as the shear exponent can be defined as a function of height and atmospheric stability [68, 90].

2.3.4 The Atmospheric Boundary Layer (ABL)

Wind turbines operate in a Boundary Layer (BL), which is where two boundaries with distinctly different physical characteristics interact [78]. The BL relevant for this work is between the atmosphere and the surface, the so-called Atmospheric Boundary Layer (ABL)[20]. The ABL is the lower region of the atmosphere that is directly influenced by the surface and responds to forcings applied by the surface within an hour [78]. The depth of the ABL may vary greatly[21], but a typical depth is in the order of 1km [73, 78]. A BL is characterized by the generation of turbulence in the form of thermals[22] due to solar heating of the surface, and therefore frequent mixing of matter and energy

[16] In this work, the displacement height d is neglected because of the flat terrain offshore.
[17] The roughness length defines the elevation at which the wind speed becomes zero [78].
[18] The reader is referred to [56, 57] for more reading on this.
[19] The reader is referred to [56] for further information.
[20] The ABL is also known as the planetary boundary layer.
[21] Typically in the range from 1km to 2km [78] but can range from tens of meters to 4km and more [73].
[22] Thermals are large eddies in this case [78].

[78]. As turbulence redistributes matter, it also regulates the wind strength and patterns [78, 91].

One can separate the ABL into two main layers in the vertical of interest for wind energy applications; the lower 10% can be referred to as the inner layer, surface layer, constant-flux layer, or Prandtl layer [68], whereas the upper 90% of the ABL is be referred to as the outer layer or Ekman layer [68]. As explained in section, the wind flow in the surface layer is highly affected by the surface friction [62, 68]n 2.3.3. Moreover, the layer is characterized by relative constant fluxes, hence the name *constant-flux layer* [78]. In the Ekman layer, the Coriolis and pressure forces[23] dominate [62, 68, 78]. The region outside the outer layer, thus; outside of the BL and not responding directly to surface changes, is naturally called *the free atmosphere* [78]. The wind in the free atmosphere is said to be geostrophic[24]. Between the BL and the free atmosphere is a capping inversion layer. This layer is characterized by high static stability and limits the efficient mixing caused by turbulence to stay within the BL. Thus, the layer marks a relatively large change in the vertical profiles in the atmosphere [73].

2.3.5 Atmospheric Stability

Like most renewable energy sources, wind energy is mainly generated by the sun. Due to the uneven solar radiation, the temperature is distributed in a vertical profile in the atmosphere, just like the wind explained in section 2.3.3. This profile varies diurnally and/or annually, depending on the surface underneath. An air parcel[25] in the atmosphere has a temperature and corresponding density, relative to the ambient temperature at the same height. According to this relationship, the buoyancy force in the atmosphere will either suppress or enhance the vertical motions of the air parcel [78]. Based on this knowledge, one may define three[26] (static) atmospheric stability conditions, referred to as stable-, neutral-, and unstable. The dry Adiabatic Lapse Rate (ALR) is a useful tool to identify and separate the stability regimes. It describes the theoretical temperature change in the atmosphere with altitude, $\Gamma_{ALR} = -\left(\frac{dT}{dz}\right)_{ALR}$, assuming that the air parcel moves adiabatically[27]. In general, the dry ALR implies that an unsaturated air parcel will cool at a rate of 9.8K km^{-1} [73] until the parcel becomes saturated. A saturated air parcel will cool at a slower rate due to the heat provided by the humidity. The ALR, denoting the actual temperature change with height, can be described as $\Gamma = -\left(\frac{dT}{dz}\right)_{air}$ [73]. All stability regimes are illustrated in figure 2.4 by considering a rising air parcel.

[23] The frictional force is also present in the outer layer, but decreasing with height [78].

[24] Geostrophic wind is when the Coriolis force is in balance with the pressure gradient force, due to the absence of the frictional force [56, 73, 77].

[25] An air parcel can be thought of as a "box" of air.

[26] The atmospheric stability can be further divided into finer regimes [92].

[27] An adiabatic system means that there is no heat exchange between the system(air parcel) and the surroundings(atmosphere). In this case, the temperature change happens because of the air parcels' expansion or compression.

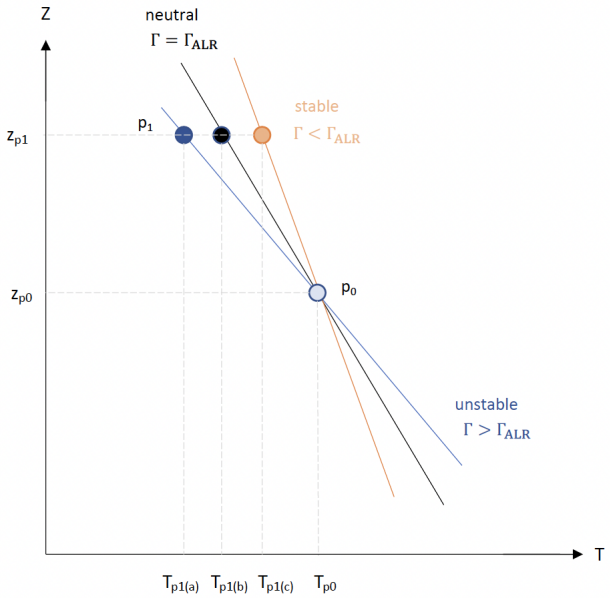


Figure 2.4: Illustration of how stability is defined by following a particle moving from the initial position p_0 upwards in altitude, to position p_1. If the air parcel follows the ALR as it raises from z_{p0} to z_{p1}, the parcel will have the temperature T_{p1b} and correspondingly neutral conditions. However, if the adiabatic cooling effect causes the rising air to become colder (T_{p1b}) than its surroundings (T_{p1c}), its vertical motion will be suppressed, and the conditions are stable. On the other hand, if the adiabatic cooling effect causes the rising air to become warmer (T_{p1b}) then the surroundings (T_{p1a}), the conditions are unstable, and the vertical motion is enhanced. Figure inspired by from [73].

Neutral stability means that the air parcel's temperature is in equilibrium with the surrounding air ($\Gamma_{ALR} = \Gamma$). Hence, the parcel is not exposed to any forces in the vertical direction (negligible buoyancy forces). A neutral BL is in the state of forced convection[28] [73], often characterized by strong winds interacting with the surface roughness, causing sufficient mixing of the BL [57].

The conditions are unstable when the air parcel has a higher temperature than the surroundings ($\Gamma_{ALR} < \Gamma$). In this case, the parcel has a lower density than the surroundings (positively buoyant). While considering an upward displacement of the parcel, the atmosphere enhances the upward movement to further ascent. This happens e.g. when the surface is warm, heating the air close to the ground so that it is warmer than the overlying air. An unstable BL is in the state of free convection[29] and has efficient mixing due to increased turbulence and a weaker vertical wind speed gradient (less stratified BL). The mixing results in smaller wake regions and increased loading on the downstream turbines if operated in the wake [70, 73].

The conditions are denoted as stable when the air parcel has a lower temperature than the surroundings ($\Gamma_{ALR} > \Gamma$). Considering an upward displacement of the air parcel, it has a higher density than the surroundings due to its low temperature. Therefore, the

[28] Forced convection occurs when shear is present in the mean wind (also known as mechanical turbulence).

[29] Free convection is the movement of air due to thermal gradients.

atmosphere will suppress the upward displacement of the air parcel (negatively buoyant). This happens e.g. when a cold surface cools down the air adjacent to the ground so that it is colder than the air right above. A stable BL has a stable stratification due to suppressed turbulence intensity and mixing, and therefore a larger shear [91]. This will result in longer wake recovery distance and larger wake velocity deficit regions, but less loading on downstream turbines if operated with the same conditions[30] as mentioned in the unstable case [70, 73].

2.3.6 Transient Event

In the atmosphere, the stability regimes may provoke various transient events, which are local, short-lasting events that can affect the wind farm significantly[31]. Some transient events can e.g. be Low-Level Jets (LLJ), ramp events, or Open Cellular Convection (OCC) [93–97]. For this work, OCC is of particular interest, as some data have been measured during the OCC event (section 3.2.1).

Cellular Convection (CC) is shallow cloud formation at mesoscale in the ABL[32]. It is formed during unstable atmospheric conditions, as cold air masses approximate relatively warm air masses, in a so-called cold front. The phenomenon is often formed over the ocean, as cold air blows from the continents and lays on top of the relatively warm ocean air. The CC happens in ABLs with a capping inversion layer (see section 2.3.4). The process of CC formation can be explained through three stages. Sea-fog is initially formed due to cold air advection, which develops further downwind into *cloud streets*[33], and eventually evolves into CC [78]. According to Hubert [98], the CC could appear as open (Open Cellular Convection (OCC)) or closed. Open cells tend to occur when the temperature gradient (air-ocean (surface)) is relatively large, whereas closed cells occur during weak gradients. The OCC appears as cloud-free cells (hexagons) capped by a cloudy border due to the circulation within the cells. The hexagonal cloudy rings experience an upward motion, whereas the cloud-free center a downward motion[34]. OCCs typically have diameters of 10km to 100km and depths ranging from 1km to 3km [78, 99]. Satellite images often detect the phenomenon, and the reader is referred to [97] for more information about the OCC event [95, 98, 99].

2.3.7 Offshore and Onshore Conditions

For the wind energy discipline, it is convenient to separate the ABL further based on the surface interacting with the atmosphere. This is because the various surfaces may greatly affect the atmosphere's stability. The BL between the atmosphere and the ocean (offshore wind farm) is referred to as the Marine (Atmospheric) Boundary Layer (M(A)BL). On the contrary, when the BL is between the atmosphere and land (onshore wind farm), the ABL is referred to as Land (Atmospheric) Boundary Layer (L(A)BL)

[30] Same conditions are in this case equal spacing between turbines and turbine characteristics (i.e. same wind farm).

[31] Transient events can impact power output and structural loading on the turbines.

[32] Also known as mesoscale cellular convection.

[33] Cloud streets are linear cloud formations.

[34] On the contrary are closed CCs cloudy closed cells surrounded by clear rings.

from now on. The stability conditions in a MBL region differ from a LBL region. In this section, key drivers causing differences in the conditions in the MBL and LBL are briefly described [68].

Heat Capacity. Firstly, there is a difference in the heat capacity of the ocean and surface of the earth, resulting in different temperature gradients. The ocean's heat capacity is relatively large, allowing it to store rather much heat whiteout, increasing the ocean's temperature significantly. With few exceptions, the air-ocean temperature variations are no more than $1°C$ to $2°C$ [73], resulting in minimal temperature gradients. Hence, the stability offshore is generally driven by the movements of the low-level wind field, where the tendencies are from cold towards warm areas at microscale. The result is cold air above warm ocean; thus, an unstable MBL [73]. However, the wind patterns are also driven at a synoptic- and mesoscale, which are complex systems that are products of several factors in the atmosphere. At times, these systems cause the warm wind to advect over cold waters. For example, this can be when warm air travels with the Gulf Stream northwards to colder waters [73]. In this kind of event, the MBL is stably stratified. Generally, the MBL exhibits a weak diurnal cycle due to the minor temperature variations in the ocean. One can instead identify an annual variation, where the dominating pattern is unstable MBL in the autumn/early winter and stable MBL in the spring/early summer [68]. This is mainly because the ocean's temperature is lagging approximately one month with the variation in air temperature. So, when the air is cooled in the autumn, the ocean keeps its heat for approximately one month longer, resulting in unstable conditions. On the other hand, the ocean is cold relative to the heated air in the spring, resulting in stable conditions. Another effect of the heat capacity is the effect on the BL depth. The MBL is generally shallow compared to the LBL, and more stable and homogeneous. Moreover, the earth's surface (land) heat capacity is relatively low. Thus, it requires little energy to heat it. In response to the alternating heating and cooling of the surface, the LBL exhibits a great diurnal cycle. Thus, when the sun rises in the morning, the surface responds almost instantaneously by increasing the heat at the surface. As does the air above, creating thermal gradients resulting in unstable conditions. The low heat capacity also makes the earth's surface poor at storing heat. Therefore, it cools down relatively quickly, causing air parcels adjacent to the ground to cool, causing stable conditions. Thus, assuming clear sky, the pattern recognized is stable conditions at nighttime and unstable conditions in the daytime. The different heat capacity also results in different temperature profiles, resulting in a much larger variation in the LBL depth. In general, this makes the LBL deeper than the MBL [68, 78].

Humidity. Secondly, humidity at the site will generally affect the conditions of the site. There is also a vertical distribution of moisture in the atmosphere, as described previously for wind and temperature [73]. Onshore, this profile changes frequently according to a diurnal cycle. The vertical profile of humidity offshore, on the other hand, constantly decreases with altitude from the surface. As moist air is lighter than dry air at a specific temperature, this contributes to a continuous and slight destabilization of the MBL [68]. The humidity flux's impact may vary depending on the stability, but the flux will always be upward offshore [68]. Moreover, moisture will change the lapse

rate of the air parcel as it reaches saturation. Saturated air is cooled down slower than dry air parcels with altitude because condensation releases heat energy into the atmosphere as the water vapor turns into liquid water [78]. In addition, moist air above the ocean results in a more extensive cloud formation than onshore. This often results in strengthening the capping inversion layer, which reduces the entertainment of dry air into the capping layer. These conditions are ideal for OCC, as elaborated on in section 2.3.6 [73].

Roughness Elements. Thirdly, the surface roughness influences the ambient turbulence intensity of the site [74]. The surface roughness is indicated by the so-called surface roughness length, z_0, which is a length scale valid in the surface layer. As previously mentioned, the wind velocity is distributed in a vertical profile in the atmosphere, and the surface roughness (together with atmospheric stability) is key for determining the wind profile. Offshore, the surface roughness varies depending on wave characteristics[35] such as wave size, height, and shape. However, one can often assume that the surface roughness offshore is smaller, more homogeneous, and with the possibility of fluctuating as opposed to onshore surface roughness [68, 100]. Onshore, roughness elements range from small pebbles or trees to big mountains or buildings, increasing the drag force at higher altitudes. As seen from the (neutral) logarithmic law, equation (2.2), the velocity (theoretically) is logarithmically proportional to the ratio of surface elevation to the (aerodynamic) surface roughness length. Hence, a lower surface roughness length increases the logarithmic expression, resulting in a steeper vertical wind profile. This is because a low surface roughness length results in less drag forcing from the ground. Typical surface roughness length values range from 0.0001m to 0.001m offshore and from 0.002m[36] to 3m onshore [56]. The surface roughness length may result in a slight difference in the wind profile onshore versus offshore. However, the greatest impact is perhaps due to very closely packed roughness elements onshore, which may displace the entire wind profile to some elevation. Hence, when the wind profile begins at the top of the wave offshore (e.g. at 1m height, often neglected), the profile onshore may begin on top of a forest (e.g. 30m height). A turbine at some specified hub height will therefore experience higher wind speeds offshore in this case, as the wind speed has accelerated to higher wind speeds offshore [56].

Summary. One can summarize by saying that the MBL exhibits annual variation, whereas the LBL experiences diurnal variations. The MBL is generally unstable in the autumn/early winter and stable in the spring/early summer. On the other hand, the LBL will have stable conditions at night and unstable during the day [73, 78]. Moreover, the MBL is shallower than the LBL, mainly due to the different temperature gradients. Thus, turbines of equal size are, onshore, placed in the surface layer, whereas offshore, they are placed in the Ekman layer. Additionally, will the roughness elements be smaller in the MBL compared with the LBL. As a result, one can generally expect the wind in offshore conditions to be more uniform and reliable, less turbulent, and of higher velocity at a given height [68]. Thus, offshore turbines generally experience less shear over the rotor. On the other side, offshore wind turbines need larger free flow re-

[35] The wave formation depends on the wind speed and atmospheric conditions [68].

[36] Can also be in the order of 0.00001m for areas of ice/mudflats.

covery distances than onshore wind turbines. It is important to emphasize that these assumptions are broad generalizations, and local solar radiation and cloud formation, topography, and season significantly impact the site's stability, wake formation, and loading on turbines. Moreover, ideal stability regimes such as those mentioned here are not often found in nature, but rather occur as a combination of all regimes simultaneously. Therefore, it is necessary to exterminate the wind closely to determine stability and turbulence at a specific site [91].

2.4 Wake Predictions

Firstly, this section describes widely used wake estimation techniques. Next, two main optimization processes in the wind energy industry which may benefit greatly from POD is briefly described.

2.4.1 Methods For Predicting Wake

Modeling the wakefield and wind dynamics in a wind farm is a comprehensive and ongoing challenge. However, the wake effect is a phenomenon that has been known for several decades, and thorough research has therefore been done in this regard, e.g. [15, 25, 60, 101]. In general, one can separate the methods for estimating wake into three techniques; experimentally [81, 102], analytically [11, 23, 103], and numerically [20, 25].

Experiments

Experiments in the wind energy discipline mainly cover laboratory experiments such as wind-/water-tunnel experiments and field measurement experiments. In wind-/water tunnel experiments, one attempt to construct a small-scale model representing the true utility-scale system, in a laboratory. This approach has provided details information about the wake region throughout the years [20]. However, obtaining utility-scale results from the laboratory-scaled findings is often limited by the ability to reproduce the same conditions in a laboratory as in reality. Usually, significant differences are found in various parameters when comparing the imitated, small-scale model to the real, large-scale model (i.e. atmospheric stability) [20, 70].

On the other hand, field experiments do provide measurements at utility-scale. Utilizing several measurement instruments at the same site has provided further knowledge about the wake effect [74, 102, 104]. However, field measurements are not a trivial task and are frequently of limited spatial-temporal resolution, thus insufficient to explain the structures of the wake in a detailed manner [8]. Moreover, it is challenging to fully know the everchanging inflow conditions, resulting in uncertainties [42].

Experimental methods generally capture vital features of the flow structure in the wake region, but are often mainly used to validate the skill of numerical and analytical models [20, 69, 105].

Numerical Models

Numerical methods, or high-fidelity models [106], does in this context comprise Computational Fluid Dynamics (CFD) and is amongst the most sophisticated and precise methods of predicting wake today. CFD makes predictions of the wind flow, by utilizing high-fidelity numerical techniques to solve the partial differential equations governing turbulent fluid [24]. In general, these methods are accurate. However, the calculations are time-consuming, computationally demanding, and expensive. For complex, non-linear models of high-fidelity and degrees of freedom, one numerical simulation may take from several hours up to days to finish, which is not applicable in most practical cases in the wind energy industry. The presence of turbulence, with its broad range of energy scales, is the main reason for this. Additionally, calculations are even more complex near the walls (physical boundaries such as turbine blades). To address the complex nonlinearities in the equations, manipulation is exploited. The result is an increased number of unknowns, making the equations impossible to solve. This is the so-called *closure problem of turbulence*. Thus, there is no rigorous, statistical theory of turbulence [107]. In order to approximate results and models using these equations, the unknowns are parameterized, which introduces uncertainties [78].

One widely used method for calculating wake is the Large Eddy Simulations (LES). Simply put, is LES a filtering operation. A lower grid size is chosen, and all turbulent structures (eddies) larger than this size are resolved in the Navier-Stokes equations. All eddies smaller than this scale, mainly the eddies responsible for dissipation, are parameterized [62, 70]. LES is a widely used method, often combined with simplified turbine models such as the actuator disk/line method [36]. However, due to the broad eddy-size range in a turbulent flow, it is still a computationally demanding process, inapplicable in most practical cases [10, 108].

Another well-known method for engineering problems is the Reynolds Averaged Navier-Stokes (RANS) approach, where the velocity flow is split into a fluctuating[37] and a mean part (see section 3.3). In RANS, only the mean part is utilized, which simplifies the cascade challenge in LES. However, whereas the original equation was a closed system, the averaging process of nonlinearities includes new unknown Reynold stress terms which must be parameterized. Moreover, RANS resolves all scales of turbulence, making it inapplicable for complex turbulent flows [41, 62, 70, 76].

Finally, PArallelized LES Model (PALM), which is used in this work, means that the LES model is intended for massively parallel systems with distributed memory. Parallelization of the LES results in less computational time and thus the ability to include more features. The PALM model enables the multiscale framework to predict turbulent wind flows within a wind farm with increased precision due to the implementation of the actuator disk parameterization with rotation. However, the model is sensitive to the selection of input parameters [97, 109, 110].

[37] By a fluctuating part mean deviations from the mean.

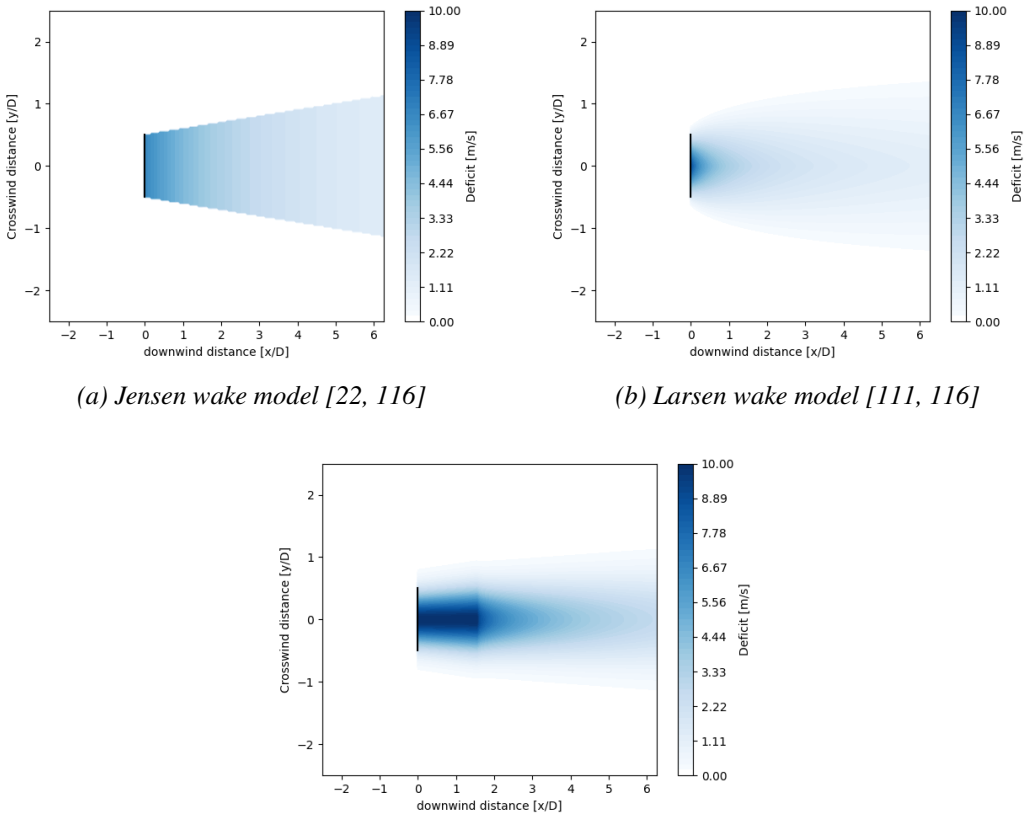
Analytical Models

In the past decades, several *analytical models*, or low-fidelity models [106] have been proposed to reproduce the wakefield. Some well-known are those developed by Jensen [22], Frandsen et al. [103], Niayifar and Porté-Agel [89], Larsen [111], and Bastankhah et al. [23]. The models are computationally efficient and are based on simplified versions of the fluid dynamical equations and/or empirical descriptions of the velocity field [33, 36]. They can be divided into simple top-hat models and Gaussian or 'bell-shaped' models [11, 22]. The analytical wake models are often not valid for the entire wake region, only parts. Analytical models are the state-of-the-art method in near real-time optimization processes such as wind farm controlling (see section 2.4.2) and wind farm layout optimization (see section 2.4.2) [20, 36] and can be justified in some cases. However, these models generally exclude key features such as turbine properties and operation, atmospheric stability, and flow characteristics. Therefore, the results are often limited when describing the wake dynamics (e.g. wake meandering), making them inapplicable to address turbine performance and loads [28, 36]. Some models will be addressed briefly in this section [11].

One of the pioneering wake models is the one proposed by Niels O. Jensen in 1983, later referred to as the *Jensen Wake model* [22]. This model falls within the top-hat models and is valid in the far-wake region only. It is a simple wake model, assuming linear wake expansion as a function of the downstream distance from rotor, with a wake spreading parameter, k, representing the growth rate (diameter) of the wake. Hence, both the wake expansion and velocity deficit are a linear function of the downstream distance from rotor. The growth rate is an empirical parameter, varying from one wind farm to another. However, the default values of 0.05 offshore and 0.075 onshore are often used [112]. The growth parameter is also frequently seen as a function of the ambient turbulence intensity [11, 112, 113], as this has shown to be one of the most influencing factors on the wake recovery [23, 112, 114]. Moreover, the model is characterized by an equal velocity deficit in the cross-section of the wake. The model is widely used due to its low computational costs and robustness [11, 112, 115].

The Gaussian wake models are characterized by the bell-shaped wake distribution, which, in general, is a more realistic view of the wake. One of the most widely known Gaussian models is the one proposed by Bastankhah and Porté-Age in 2014 [23]. Similar to the Jensen model, this model retains the assumption of linear wake expansion. The growth rate, k^* is expressed as a linear function of the ambient turbulence intensity, which can be tuned using empirical experiments. The model is valid in the far wake only and derived from the conservation of mass and momentum equations [23, 115].

PyWake PyWake is a package in python, developed for wind farm flow field calculations, amongst others, where a variety of wake models have been implemented. Some of the most widely used models are displayed in figure 2.5, using the open-source codes provided by PyWake. Figure 2.5 a represents the wake model developed by Jensen, figure 2.5 b depicts the wake model proposed by Larsen, and figure 2.5 c is the wake model from Bastankhah et al. [116].



(a) Jensen wake model [22, 116]

(b) Larsen wake model [111, 116]

(c) Bastankhah wake model [23, 116]

Figure 2.5: An overview of some of the most widely used analytical models in the industry today. Figure (a) is of the Jensen wake model, (b) is the Larsen wake model, and (d) is the Bastankhah model.

2.4.2 Wind Farm Layout Optimizing and Controlling

As already touched upon, precise wake estimations are crucial to enable wind farm design and operational improvements [28]. The gains obtained from a wind farm are superficially determined by the balance of costs versus the Annual Energy Production (AEP) of the wind farm. Being able to find the optimal balance between increasing AEP and reducing costs is a comprehensive task and a hot topic to further improve [13, 17, 117–119]. In the planning phase, one needs to establish an optimal layout for turbine operation. Next, based on the chosen layout and the conditions at the site, one must operate the turbines in harmony to obtain the most power at the least cost. An optimization problem is essentially several similar problems with small differences that need to be solved. Achieving this enables one to compare results, thus deciding the "best" solution. Therefore, efficient predictions are required in this process [120]. Today, the state-of-the-art method for such optimization processes is the analytical model, as mentioned in section 2.4.1. However, as seen previously in this work, there is potential for improvements in accuracy in these models. Including more precision and dynamics into today's models could improve layout optimization and controlling of wind farms significantly.

Wind Farm Layout Optimizing. In a wind farm, the turbines are arranged together in a given layout, separated by some spacing distance. The various ways of clustering the turbines together[38] may significantly affect power output and loading on the downstream turbines. Wind farm layout optimization aims to find the optimal layout which provides the highest power output at the lowest expenditures at the given site [121]. The wind farm layout optimization problem is complex and multidisciplinary [122] and includes aspects such as system expenditures (related to cable, mooring, foundation, turbine etc...), environment, technical constraints, area constraints, wind conditions, ecosystems, structural loading, etc. Since wake may account for up to 20% power losses, as already mentioned, it is one of the main drivers to consider when identifying the optimal wind farm layout [36]. In the optimization process, several layouts are calculated and compared. Therefore, it is beneficial to obtain fast and relatively accurate estimations [16, 106, 122]. The faster the calculations are, while being fairly precise, the more parameters one can account for in the optimization.

Wind Farm Controlling. *Wind farm controlling* becomes vital during the operation of the wind farm and attempts to coordinate all turbines in a wind farm in harmony. In other words, by controlling the wind farm, one aspires to obtain maximum energy at the lowest loading and expenditures. Depending on the wind conditions, either optimizing or limiting the power output can be achieved by wind farm control. One control mechanism is passive/active yaw-, pitch- and generator load/torque control [56].

Yaw control is the mechanism responsible for changing the direction of the entire rotor/nacelle on the horizontal axis. The yaw makes it possible to steer the rotor to face the incoming wind at the desired angle, e.g. aligned with the wind, or at some angle to steer the wake by yaw misalignment [123]. Pitch control is the mechanism altering the turbine blade, where the chord line of the blade is altered relative to the incoming wind speed, referred to as the angle of attack. Pitch control allows one to always operate at an optimal angle of attack with regards to the wind conditions and thus control the rotational speed of the blades [124, 125]. Finally, generator torque can also be utilized to regulate the tip speed ratio[39] [19, 56].

By utilizing control mechanisms such as these, one can control the aerodynamic force on the turbine blade and loading of each turbine in the wind farm. However, the process is data sensitive in the sense that the data forms the basis of the steering [58]. To obtain the optimal solution for e.g. the pitch angle, one must try and fail. Being able to conduct several calculations for various pitch angles within a short period of time, further optimization of the power output is enabled. Since the wind flow changes continuously, it is key to get fast wake predictions. Detailed and efficient wake predictions can improve the controlling process and increase the power output [33, 56].

[38] Layouts can be rectangular, triangle, semicircle, staggered, random, etc.

[39] The ratio between the tip speed of one blade and the wind speed [56].

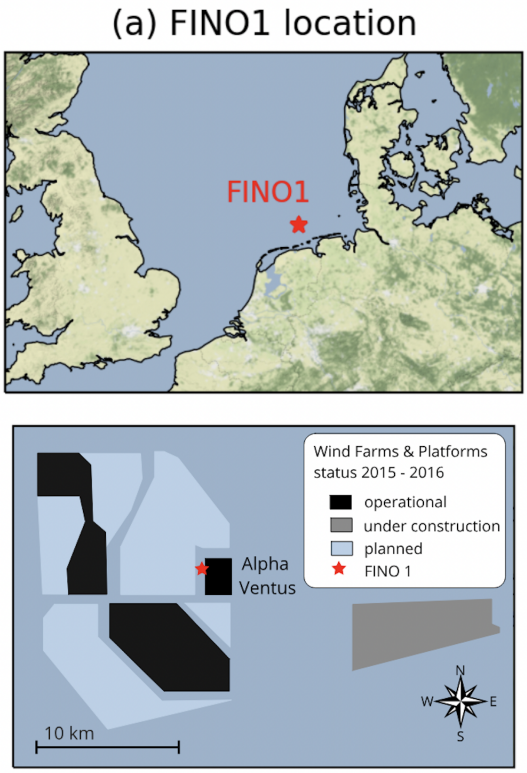
Chapter 3

Data and Methods

This chapter elaborates on the data and methods used to achieve the results provided in chapter 4. All procedures are conducted in PyCharm using Python. Firstly, is a description of the site where the data is collected from presented. Next, is the data used presented with an explanation of the processing conducted. Finally, is a brief description of ROM done before the process for obtaining the ROM through POD is elaborated. Finally, the methods for including stochasticity in the time-dependent weighting coefficient are explained.

3.1 Site

FINO1 (Forschungsplattformen In Nord- und Ostsee nr.1) is a research platform located in the North Sea, close to Germany's first commissioned offshore wind farm *Alpha Ventus*[126]. Alpha Ventus consists of 12 turbines organized in a rectangular layout. An illustration is shown in figure 3.1a. The distance between the turbines varies but is generally approximately seven rotor diameters ($7D$) [127]. Because the wind farm was commissioned to be an offshore test site, it consists of two types of turbines with different foundations. Turbines 1-6 are Senvion 5M turbines, and 7-12 are Adwen Wind M5000 (Multibrid) turbines, with jacket and tripod foundations, respectively[128, 129], see figure 3.1b. For more information about the turbines, see Table 3.1.



(a) Location of FINO1 in relation to Alpha Ventus, figure from [127].

(b) Alpha Ventus turbines and foundations: Senvion turbine with jacket foundation in front, and Adwen turbines with tripod foundations in the back, figure from [130].

Table 3.1: Information about the turbines in the Alpha Ventus wind farm [129, 131, 132].

Turbine overview		
Turbine name	Senvion 5M	Adwen AD 5-116
Manufacturer	Senvion SE	ADWEN Offshore
Formerly turbine manufacturer	REpower Systems SE	AREVA Wind GmbH
Turbines	1-6	7-12
Rated power [MW]	5	5
Rotor diameter [m]	126	116
Hub height [m]	92	90
Cut-in speed [m/s]	3.5	4
Cut-out speed [m/s]	25	25
Foundation	jacket	tripod

3.2 Data

This section aims to provide an overview of the various data used and how they were processed. In this work, LiDAR- and wind vane and cup anemometer measurements were obtained at FINO1, SCADA data from turbine 7 (AV7) in Alpha Ventus, and Weather Research and Forecasting (WRF)-PALM are data used. Two periods are investigated, hereinafter referred to as the first and second periods. They are one hour at the 22^{nd} of September 2015, and the period of 24^{th} of September to the 3^{rd} of October 2016, respectively. Data collected in the first period, WRF-PALM data, is presented in section 3.2.1. In the second period, SCADA, cup anemometer, wind vane, and LiDAR data are collected. This is mainly due to the retrieval process necessary for the LiDAR measurements, which is described in section 3.2.2. Both periods were collected during the Offshore Boundary-Layer Experiment at FINO1 (OBLEX-F1) campaign [133]. The periods are chosen based on the availability of the various data measurements.

3.2.1 WRF-PALM Data

WRF-PALM data from the first period, the first hour of 22.11.2015, is used in this work for POD. As mentioned previously, POD applied to LES data is an already proven method for reducing the order of a complex system. Hence, in this work, the WRF-PALM data is used to compare the results using the LiDAR data. However, it must be specified that WRF-PALM data is known to be less idealized than LES [97]. The data is collected in unstable conditions during the passage of an OCC event (see section 2.3.6), which commonly occurs in the North Sea [94].

WRF-PALM data denotes data obtained by utilizing the Weather Research and Forecasting (WRF) model and parallelized LES (WRF-PALM) as a multi-scale model chain. The WRF simulation is initialized with ERA5 reanalysis data at 30km resolution in this work [134]. The result of the WRF simulation is a wind field of 40m grid spacing and is conducted with the aim of reproducing the atmospheric conditions of the site. However, the resolution is too coarse to resolve the frequently occurring small-scale eddies in turbine wakes. Downscaling the original WRF domain spacing from 40m to 10m (horizontal) increases the number of grid points by 64 times ($4\times4\times4$), allowing the small scald dynamics to appear in LES[1]. However, the computational time would increase by the same order. The grid spacing is only reduced over a local area around the wind farm to reduce the computational time to the greatest extent possible. The local area is then a child domain added to the LES, and the original domain becomes the parent domain (see information about resolution in Table 3.2). The child domain is obtained through a one-way nesting approach. Hence, the results from solving the equations in the parent domain are used to obtain the nested boundaries for the child domain through interpolation, so that the same equations can be solved in the child domain. An overview of the grid transformation is presented in Table 3.2

The equations are solved using PALM, and the reader is referred to [110] for further reading regarding the nesting system utilized, and [97] for information about the WRF model.

[1] The section desribes the LES model being parallelized using PALM

Table 3.2: Table presenting information about the grid resolution in the WRF-PALM data.

	Grid spacing		Number of grid points			Nested domain lower left corner coordinates	
	Δx, m	Δz, m	Nx	Ny	Nz	x, m	y, m
WRF	40	10	576	576	96	-	-
LES, parent	40	10	576	576	96	-	-
LES, child	10	5	448	512	64	10480	8400

WRF-PALM Data Description. Figure 3.2 illustrates the domain for the WRF-PALM data used in this work, covering four turbines (AV1, AV4, AV7, and AV10). This corresponds to the spatial resolution of x =[1000, 1900] and y =[895, 4495] with [dx, dy] = 10m. The temporal resolution is ∼2 seconds for one snapshot[2] to be conducted, consisting of 32400 measurements. Therefore, the amount of measurements per second is 16200. Thus, and as expected, both the spatial and temporal resolutions are finer than for the retrieved LES data described in section 3.2.2. As the data is simulated using ERA5 reanalysis data, all snapshots are of the same amount of measurements and high-quality, and there is no need for processing. From investigating the minimum and maximum values for the mean wind speed, the values were found to be ∼6.16m/s and ∼16.8m/s, respectively. The hours simulated are during an OCC event (explained in section 2.3.6). Hence, the conditions are volatile, as explained in section 2.3.6.

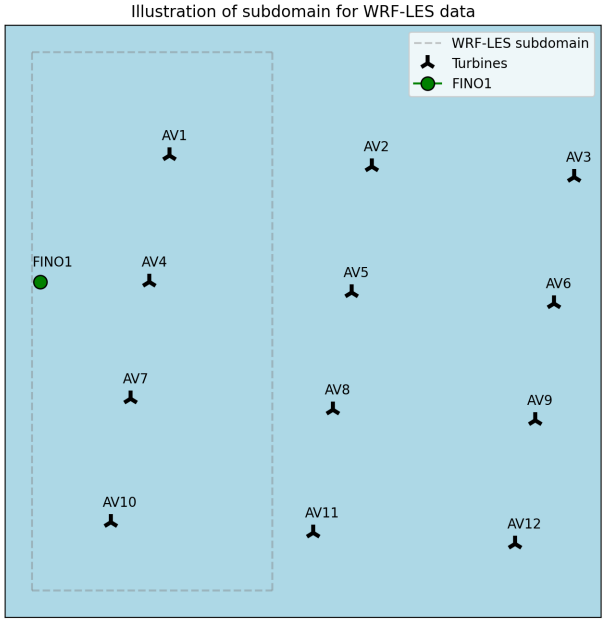


Figure 3.2: Illustration of the WRF-PALM subdomain. Showing which turbines in Alpha Ventus the simulation covers. Inspiration from [128].

[2]One snapshot in the WRF-PALM data is analogs to one scan for the LiDAR data.

3.2.2 Data Measured at FINO1

FINO1 is a platform densely packed with various measurement equipment for wind, temperature, and humidity [128, 135, 136]. These are taken from both the platform itself and the met-mast at FINO1. This work uses cup and vane anemometer measurements from the met-mast and LiDAR data recorded at the FINO1 platform. The data sets are from the second period (24.09.2016 to 03.10.2016) obtained during the OBLEX campaign. This section presents the data and briefly the instruments used.

Cup Anemometer and Wind Vane Data

The cup and vane anemometer measurements obtained during OBLEX are available in heights ranging from 30m to 100m (for every 10^{th}m) and from 30m to 90m (for every 20^{th}m) Above Sea Level (ASL, henceforth a.s.l.), respectively. Wind speed data from a cup anemometer installed at 100m a.s.l. is used in this work. The wind direction is from a wind vane mounted 90m a.s.l. Both data sets are used to retrieve wind speed and direction from the LiDAR data, which is explained in section 3.2.2. Moreover, the data is used for validating the retrieved wind speed and direction. Hereafter, the wind speed and direction is referred to as the reference wind speed and direction, respectively.

Both the cup anemometer and the wind vane are based on in-situ sensing technology, meaning that they measure directly where they are located. Thus, e.g. a mast to be mounted on is required to obtain measurements at turbine rotor height [135]. A cup anemometer consists of a number of vertical cups springing from equal distanced horizontal arms mounted on a vertical shaft. As the wind hits the collecting cups and turns them, the horizontal wind speed can be found [137, 138]. For an extensive explanation of the cup anemometer, see [137]. A wind vane is a vertical vane mounted on a rotating shaft, driven by the pressure exerted by the wind, causing the vane to align with the wind direction [138]. Generally, it is essential to process the data from anemometers and wind vanes. Typical disturbances for cup anemometer is calibration errors and wear and tear, whereas the errors for wind vanes are most related to installation [138]. Generally, since the measurements are mounted at the met-mast on FINO1, the so-called *geographical exclusion zone* is one potential error. This is distortion in the measurements caused by the wind hitting the mast before the anemometer/vane [135, 138].

Cup Anemometer and Wind Vane Data Description. The wind speed and direction are measured at a single point from the met-mast at FINO1, at the coordinates, 6°35'15.58"E, 54°00'53.94"N [128]. The measurements are mean values for every $\sim 50^{th}$ second, corresponding to one LiDAR scan (see section 3.3). Hence, the reference data also contains the irregularities explained later in section 3.2.2[3]. The original temporal resolution of the data is unknown, as this was not provided.

As mentioned, the data should be processed in some way. Due to lack of information about the distortion ranges[4], the cup anemometer and wind vane data are not treated in any way. However, a peak removal with threshold values for wind speed of 30m/s and wind direction for > 360 and < 360 was conducted. The findings were that the wind speed in the period is in

[3] A subset of 25 LiDAR scans represents the first 20 minutes of every hour. The last 40 minutes are not captured by the LiDAR.

[4] [135] states that the cup anemometer at 100m a.s.l. at FINO1 is in a height where the wind flow is not distorted due to the met-mast.

the range from 0.78m/s to 19.42m/s, and the wind direction in the range from 3.83° to 359.42°, at corresponding 100m and 90m, in which no values are considered av unrealistic. The mean wind speed is ∼9.2m/s, and the dominant wind direction is from the southwest, as illustrated by a wind rose in figure 3.3. As the distribution of the wind direction is relatively concentrated from the southwest region, and no unrealistic values in the wind speed or direction are detected, it is subsequently assumed that the distortions range is not in this region. However, some wind speed is detected from the northeast region, and the quality of that data remains unknown.

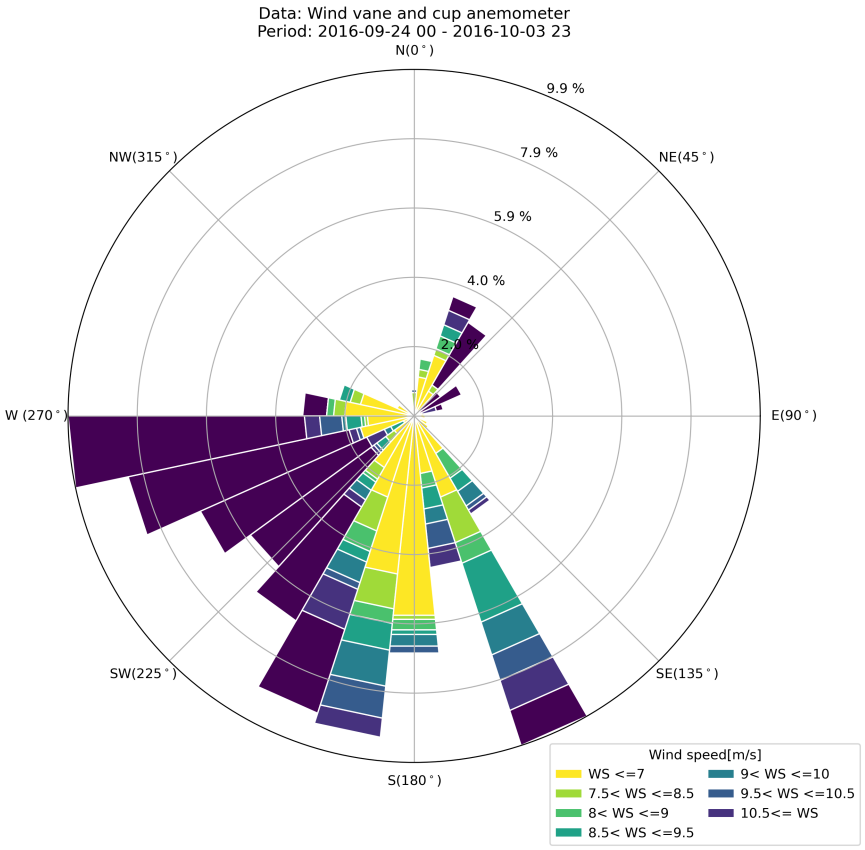


Figure 3.3: Wind speed distribution in the second period, using cup anemometer and wind vane data obtained at FINO1.

Light Detection And Ranging (LiDAR

The LiDAR data used in this work is radial velocity obtained by a single scanning Doppler wind LiDAR Leosphere WindCube 100S, mounted 23.5m a.s.l. at FINO1. Wind speed and direction are first retrieved from the radial velocity. Next, a subdomain is chosen before the processed data is used in the POD, see sections 3.2.2, 3.2.2, and 3.3, respectively.

LiDAR is a remote sensing technology providing information about the radial velocity, i.e. a component of the velocity in the line-of-sight [46], as seen in figure 3.8. Laser is emitted from the LiDAR and then reflected by the aerosol particles in the atmosphere. However, not all laser pulses are reflected well by the atmosphere[5], causing gaps in the data. Some signals are returned, and due to the motions of the aerosols, a Doppler frequency shift occurs

[5]This is heavily dependent on aerosol content in the atmosphere and weather conditions, e.g. rain can weaken the backscatter significantly [139].

in the returned signal, providing information about the radial velocity. The LiDAR provides data in a specific grid resolution (beam range) and returns a volumetric average of the radial velocities within this range [135, 138, 140]. The location can be derived from the line of sight and time of flight. In this way, the wind field over a considerable distance can be measured. LiDAR measurements are advantageous in the wind energy industry as they make it possible to measure the wind before the turbine possibly alters it. Moreover, they can provide data over the entire rotor instead of in-situ measurements such as anemometers [141]. The laser at FINO1 is emitted at a more or less constant elevation angle, $\phi = 4.62°$. The elevation angle causes the LiDAR to scan the wind field at varying heights depending on the turbine position in relation to the LiDAR, as shown in the figure 3.4.

LiDAR Data Description. Moreover, figure 3.4 visualizes the scanning range of the LiDAR, covering the three[6] bottom left turbines of Alpha Ventus (AV7, AV10, and AV11). We find that the LiDAR was directed in a south-east (from north-west) direction towards the wind farm during the OBLEX campaign. The LiDAR data was measured in a Plan Position Indicator (PPI) scanning mode, meaning that the LiDAR scans an area in 2D, explained by the beam range and azimuth angle.

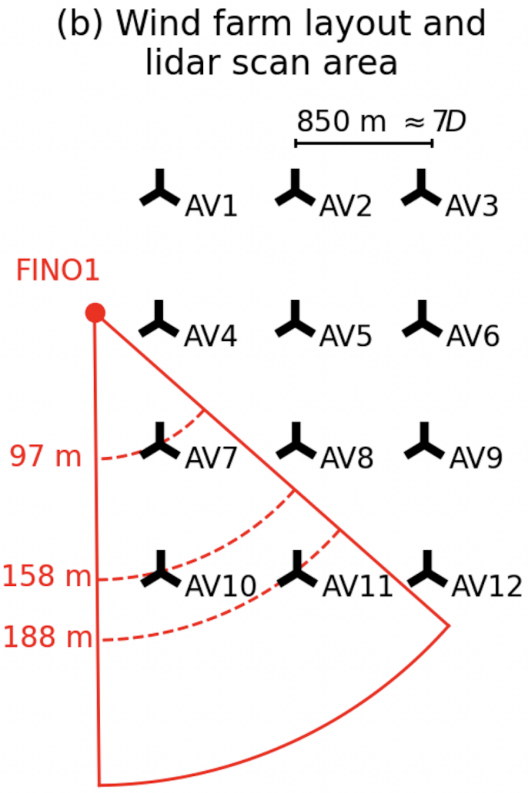


Figure 3.4: Illustration of the LiDAR scan area during the OBLEX campaign, covering AV7, AV10, and AV11 in Alpha Ventus. The numbers indicate the scanning height at each turbine captured by the LiDAR scanning area. Figure from [127].

In this work, 6000 scans[7], 600 irregular scans per day for ten days, are investigated. The LiDAR produces one scan approximately every ∼50 seconds, which is required to finish one scan. The 600 scans per day are irregular in the sense that they are separated into 24 batches of 25 scans. As a result, the is a ∼40 minute data gap every 25^{th} scan. It is clear from this that the data represents the first ∼20 minutes of each hour [142]. The beam range is from 50m to 3300m[8] (seen in figure 3.7) in a 25m range resolution, and the azimuth angle, θ, from 132° to 180°[9] in a ∼1° range resolution. Dependent on the backscatter, is the radial velocity measured in a range from 8 to 5849 measurements per scan. The wind speed components, u and v, used in this work are explained briefly in section 3.2.2. A summary of the raw LiDAR data can be seen in Table 3.3.

[6] Some of the scans only contain AV7 and AV10.

[7] Scans are equivalent to snapshots.

[8] The range varies from scan to scan, depending on the weather conditions and the atmospheric aerosol content [138].

[9] The directions are provided in meteorological convention

Table 3.3: Overview LiDAR data ranges.

LiDAR data overview					
Data	**Nomenclature**	**Min**	**Max**	**Resolution**	**Unit**
Date period	-	24.09.2016	03.10.2016	600	$\frac{scans}{day}$
Beam range	-	50	3300	25	m
Azimuth angle	θ	132	180	1	$^\circ$
Elevation angle	ϕ	4.608	4.632	irregular	$^\circ$
Seconds per scan	-	44	60	irregular	$\frac{s}{scan}$
Measurements per scan	-	8	5849	irregular	$\frac{nr}{scan}$

Data Processing.

The following paragraphs will explain how the LiDAR data is processed in order to obtain high-quality data[10] for the POD. The process is visualized in figure 3.5. Firstly, a subdomain is chosen. Subsequently, insufficient scans are detected and removed. After that, the wind speed components u and v are retrieved from the raw radial velocity. The next step involves identifying and treating outliers, and finally, the data is interpolated and prepared for the POD.

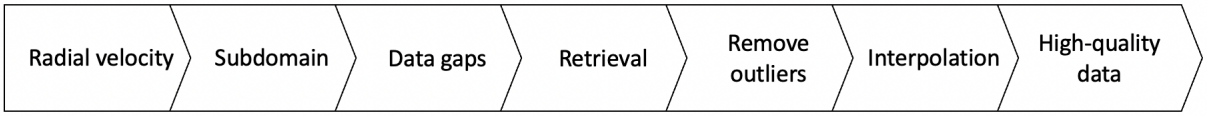


Figure 3.5: Overview of the processing steps of the LiDAR data.

Subdomain. When the retrieved wind speed (and direction) is obtained, a subdomain is chosen to investigate closer. This is to ensure high-quality data[11] and make the data applicable for the POD. The area chosen is in the range of x = [20, 800] and y = [-700, -1800] , with a resolution of [dx, dy] = 25m, as can be seen in figure 3.6. By choosing this subdomain, the POD captures at least one wake regardless of the wind direction.

Data Gaps. Significant data gaps detected as Not A Number (NaN) values frequently occur in the LiDAR scans. As previously stated, some scans contain as few as eight measurements. In this work, 95% of the data must be available in each scan to be quantified as high-quality. Hence, scans of measurements below 95% of the total scan are considered insufficient, i.e. not providing enough information about the wind field to retrieve or reconstruct it, and are consequently discarded. An overview of where the data is considered insufficient is shown in figure 4.2 in the results. Reasons for this can either be the weakened signal-to-noise ratio (SNR),

[10] high-quality is a relative term in this case.

[11] By choosing a subdomain, the NaN values caused by turbine shadow and SNR are reduced significantly.

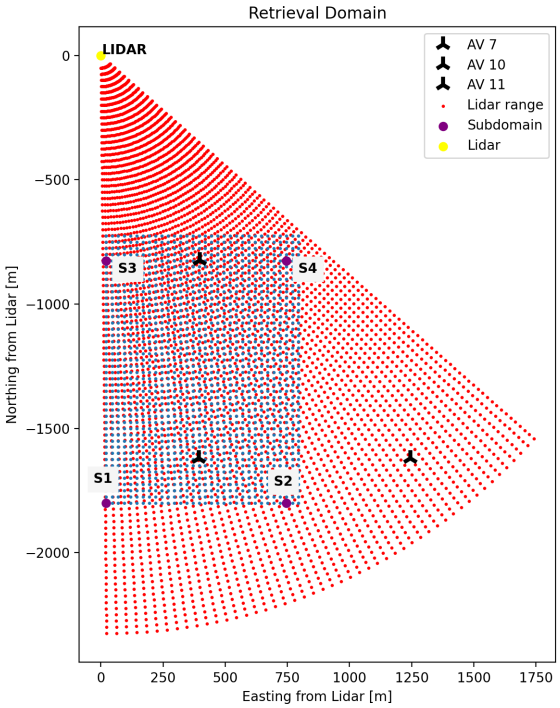


Figure 3.6: Illustration of the subdomain used in the POD algorithm.

the turbines acting as obstacles for the laser beam, or low aerosol content (as described earlier in section 3.2.2). Firstly, the SNR says something about the strength of the signal relative to the background noise and will therefore decrease with increased distance from the LiDAR [143]. Secondly, in the so-called *turbine shadow*, we also see clear gaps in data measurements [127]. Both the SNR (the ruffled edge of the scans) and the turbine shadow (present as one long white line) are illustrated in figure 3.7. As insufficient data is removed, the retrieval can be conducted.

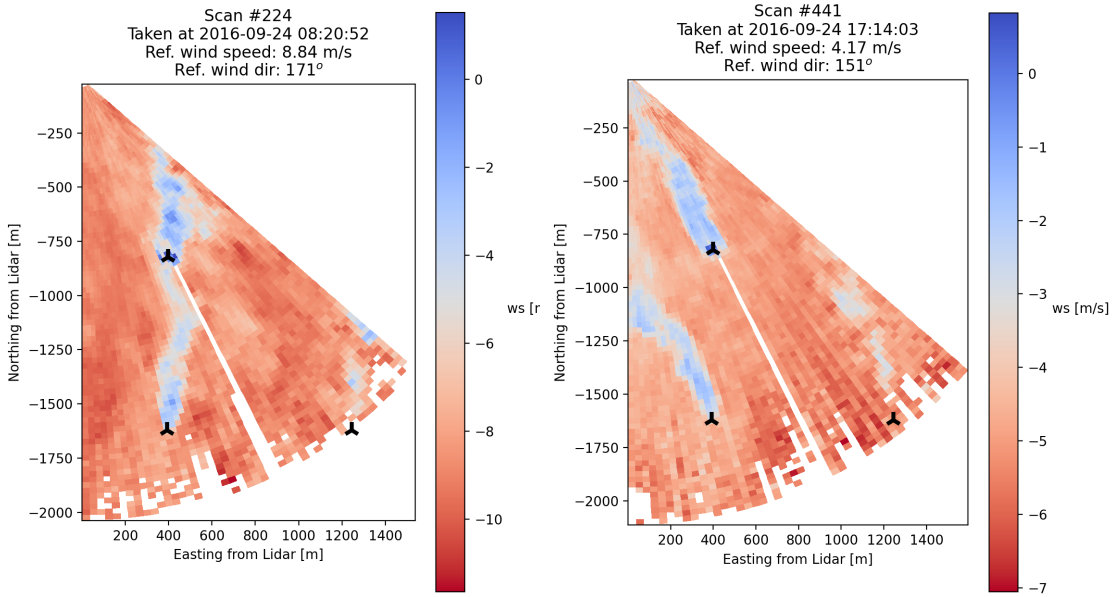


Figure 3.7: Turbine shadow in LiDAR measurements. The figure also provides an indication of the various beam ranges measured depending on the backscatter of the lasar from the LiDAR.

Retrieval. As already mentioned, the LiDAR measurements are provided as radial velocity measurements, which is not a complete picture of the total wind speed (see figure 3.8). Therefore, the wind speed retrieval (u and v) is required before the POD is conducted. In addition to the radial velocity, wind speed and direction are needed in this process, which in this work are the reference wind speed and direction (section 3.2.2). Regarding wind speed retrieval, wind direction is essential to keep in mind, as it may result in crosswind effects. This will be elaborated on further in section 3.11.

The retrieval algorithm is a study of its own; thus, numerous methods for retrieving wind speed and direction exist [46, 80, 143, 144]. In this work, the so-called Volume Velocity Preprocessing (VVP) was used before the retrieval algorithm was conducted in the manner of the article by Nihanth W. et al. [46].

In short, is the VVP based on the governing equation for radial velocity, V_r, describing the radial velocity in 3D for (u, v, w), along the line of sight given by the elevation angle, ϕ and the azimuth angle, θ:

$$V_r = u\cos\theta\cos\phi + v\sin\theta\cos\phi + w\sin\theta. \tag{3.1}$$

Moreover, can the tangential velocity in the horizontal, V_{th} and vertical, V_{tv} plane be described as:

$$V_{th} = -u\sin\theta + v\cos\theta, \tag{3.2}$$

$$V_{tv} = -u\cos\theta\sin\phi - v\sin\theta\sin\phi + w\sin\phi. \tag{3.3}$$

By assuming that the elevation height can be neglected, what remains from equation (3.1) is the 2D field where the wind speed can be described as:

$$V_r u\cos\theta + v\sin\theta. \tag{3.4}$$

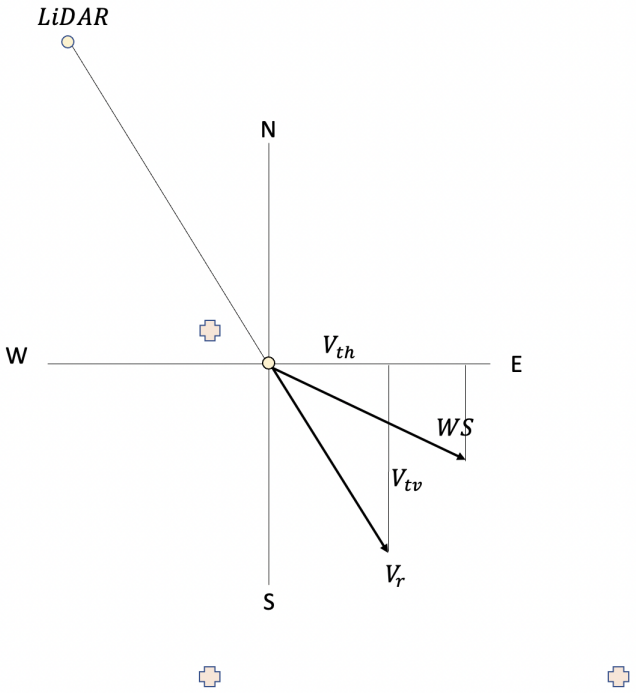


Figure 3.8: Illustration of the radial velocity, V_r, compared to the actual wind speed. Radial velocity can be decomposed into tangential velocity in the horizontal, V_{th}, and vertical, V_{tv}, plane.

At this stage, equation (3.4) contains two unknowns, u and v. With only one equation to solve with, it becomes impossible to solve without making any assumptions. In the VVP method, we exploit two consecutive scans (i) and $(i+1)$, and attempt to fit a solution using the two equations:

$$V^{(i)} = u\sin\theta + v\sin\theta, \tag{3.5}$$

$$V^{(i+1)} = u\sin\theta + v\sin\theta. \tag{3.6}$$

The method is based on the assumption of a uniform flow, meaning that it generally does not recognize the wake. The result is that the wake's shape is not adequately resolved in some scans. This, however, is used as input to the 2D VAR retrieval algorithm [145]. A cost function is derived based on the 2D field. The reader is referred to [46] for more details on the derivation. By minimizing the cost function, we achieve the wind speed's u and v components. In this work, u is retrieved as east and v as North. Due to the gap every 25^{th} scan of $\sim$40 min, and the pairwise retrieval method just explained, every 25^{th} scan is not retrieved to avoid errors due to very different flows. Finally, u and v are utilized to compute the wind speed, ws, and direction, wd are computed as follows:

$$ws = \sqrt{u^2 + v^2}, \tag{3.7}$$

$$wd = 270^\circ - \left(\arctan\left(\frac{v}{u}\right) * \left(\frac{180^\circ}{\pi}\right)\right), \tag{3.8}$$

where 270° is the angle required to obtain the wind direction in mathematical convention[12], as illustrated in figure 3.9. Since wind speed is a scalar, it is not affected by the convention of the wind direction. This data is hereafter referred to as the retrieved data.

[12] In meteorology is 0° defined as from North, whereas 0° in mathematics is from East.

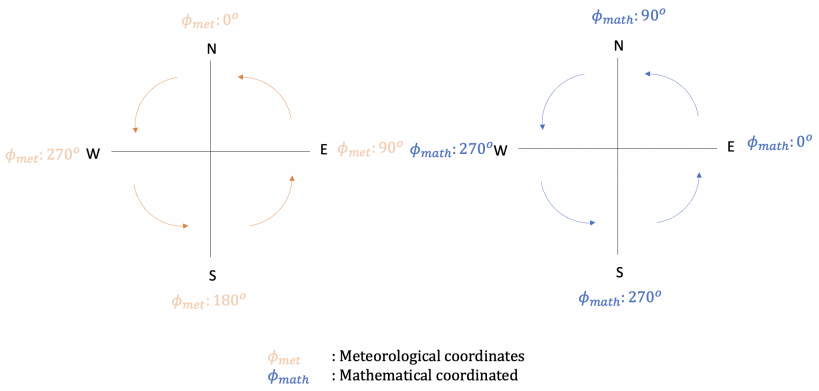


Figure 3.9: Illustration of mathematical versus meteorological convention.

Outliers. Outliers are unrealistic values common for most types of data sets, and frequently seen in LiDAR measurements [135]. For this kind of data, a correlation between unrealistic wind speed values and the corresponding wind direction is often detected (the wind direction for this period can be seen in figure 3.3). This is explained by considering equation (3.4)[13] for the 2D radial velocity field in section 3.2.2. If the wind blows towards the LiDAR, the conditions are favorable for the LiDAR to detect the Doppler shift in the frequency of the returned laser signals. In the event of crosswind, when the wind blows perpendicular to the line-of-sight of the LiDAR (see figure 3.8), the radial velocity is mainly the horizontal wind speed. Because the radial velocity component converges towards zero in these cases, the data is exposed to bias [127, 146].

Based on the azimuth angle, wind directions where bad-quality LiDAR scans are expected to occur due to crosswind events have been identified. The azimuth angle (line of sight from the LiDAR) range is from 132° to 180°. Hence we expect bias when the radial velocity angle is 42° to 90° and 222° to 270°. However, when investigating the data, we find that the range of unrealistic data starts at $\sim 210°$ [127], and the ranges 30° to 80° and 210° to 260° seem to fit the unrealistic radial velocities quite well, as seen in figure 3.10. Thus, the scans in these regions are discarded.

[13] Assuming that the elevation angle is negligible.

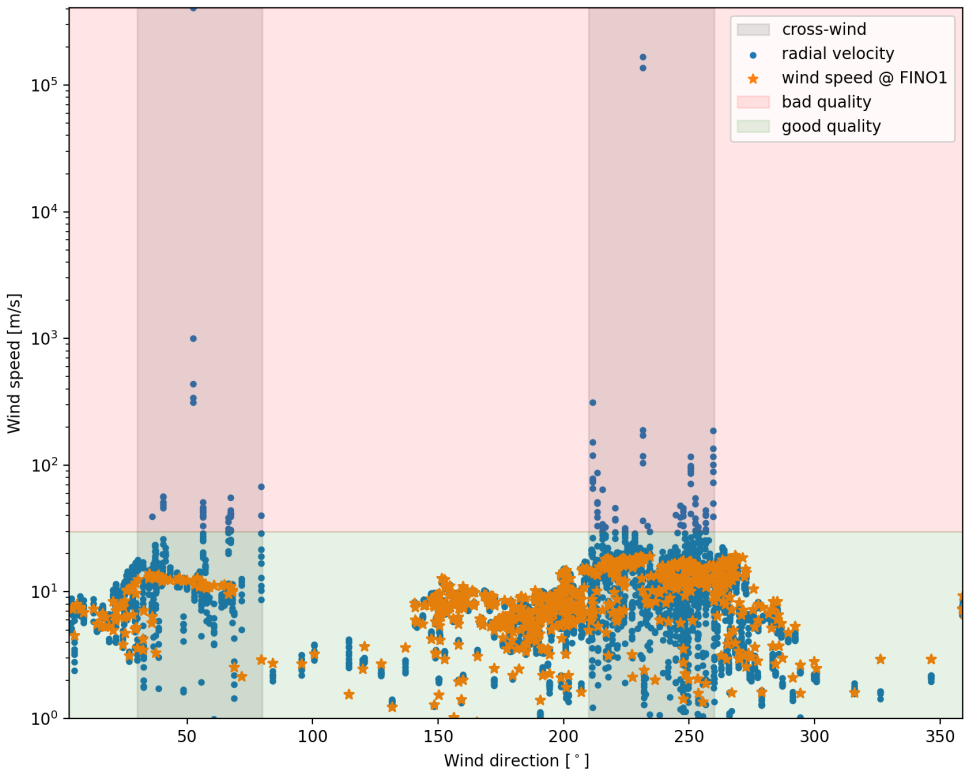
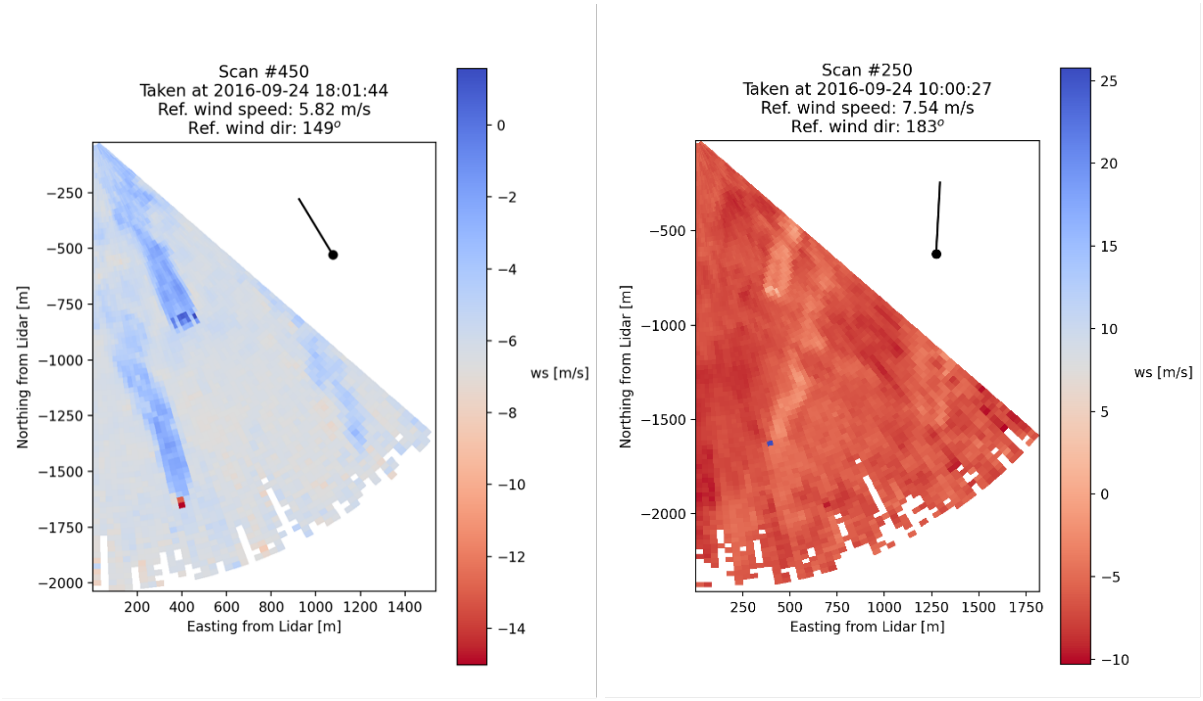


Figure 3.10: Comparing reference wind speed from FINO1 with wind speed (simply) retrieved from measurements by the LiDAR, the red area indicated values $\geq$ 30m/s, green below. The vertical lines are ranging from 30° *to* 80° *and from* 210° *to* 260°. *Figure inspired by [142].*

In general, are all values above 30m/s considered outliers and were removed before the retrieval was performed. In addition to these, outliers are detected in the turbine positions, as seen in figure 3.11. One reasonable explanation for this could be that the rotating turbine blade interferes with the reflection of the laser beam sent from the LiDAR. In LiDAR data, local peaks can be challenging to detect. This is because the area of the local peaks and the values vary significantly from one scan to the next. A local peak removal algorithm is, based on separating the data into quantiles, developed in an attempt to remove the majority of these types of outliers. The local peak removal algorithm was developed based on visual inspection for a sample of the total data. The detected local peaks are replaced with NaN values.



(a) High outlier *(b) Low outlier*

Figure 3.11: Outliers due to turbine reflection.

Interpolation In order to reconstruct the entire wind field, the final processing step before the data can go into the POD is an interpolation. In this process, the values that have been replaced by NaN values in the former stages, mainly due to the local peak removal algorithm (the other scans are removed entirely), are filled with interpolated values. As the local peak removal algorithm is limited to a few peaks per scan, the values missing may be interpolated without removing any of the physical meaning of the scans. In this work, this is done by leveraging linear interpolation. After the data processing procedure is conducted, the final data set, originally 6000 scans, is reduced to 3292.

3.2.3 Supervisory Control and Data Acquisition (SCADA) of turbine 7

The SCADA data is collected from AV7 in Alpha Ventus, which is of the type Adwen AD 5-116 (see Table 3.1). The turbine is located 919m (7.92D) away from AV7 in Alpha Ventus [127]. While turbine 4 (AV4) is the closest turbine to FINO1 (where cup anemometer and wind vane measurements have been measured), AV7 is the first turbine detected by the LiDAR scans, as seen in figure 3.4 [128, 142]. The SCADA data is used for validation of the retrieved wind speed and direction from the LiDAR data. The data is obtained from the Research at Alpha Ventus (RAVE) portal by applying for access to the Federal Maritime and Hydrographic Agency. When accessed, the SCADA data from AV7 in the second period (24.09.2016 to 02.10.2016) was downloaded through their online platform [128].

SCADA systems are computer-based systems collecting data directly from the wind turbines in the wind farm, which in turn can be used in continuously analyzing, monitoring, and control of the primary turbine functions in the wind farm [21, 68]. Through communication with

Fraunhofer Institut für Windenergiesysteme IWES, it is established that the SCADA data is unprocessed, and wind speed and direction measurements are located on top of the nacelle of AV7, right behind the rotor. Hence, we can assume that the SCADA data is, to some degree, continuously distorted by the blades of AV7 during operation. We expect the wind speed to be lower behind the rotor from knowledge obtained about power, equation (2.1), and wake (specified in section 2.3.1). Moreover, the measurements are always at risk of being distorted by wakes from neighboring turbines[14].

SCADA Data Description. The SCADA data is obtained at a temporal resolution of every 60 seconds. The data is collected at AV7 with coordinates 6°35′38.970”E, 54°00′27.018”N [128]. Hence, the data is measured at >90 a.s.l.[15] (from Table 3.1). As for the reference wind speed and direction data, the SCADA data is kept raw, and only outliers (spikes) are filtered out. In this work, the threshold values are 30m/s and in the range 0° to 360° for wind speed and direction, respectively.

Total Data Overview

For an overview of the entire data used in this work, see Table 3.4.

[14] E.g. is A7 operating in the wakes of AV4 in ∼0°, by AV10 in ∼180°, and AV5 in the range from ∼20° to ∼50°.

[15] The height of the tripod foundation is unknown.

Table 3.4: Overview of data used in this work.

Data overview					
Data	Cup anemometer	Wind vane	SCADA	Retrieved (LiDAR)	WRF-PALM
Parameter	Wind speed	Wind direction	Wind speed and direction	Wind speed and direction	Wind speed
Unit	m/s	°	m/s , °	m/s	m/s
Measurement height [m]	100	90	∼90	23.5 - ∼290	87.5
Measurement location	FINO1	FINO1	AV7	AV7, AV10, AV11 (from FINO1)	AV1, AV4, AV7, AV10 (from ERA5)
Purpose	Retrieval and validation	Retrieval and validation	Validation	POD	POD
Period	24.09.2016-03.10.2016	24.09.2016-03.10.2016	24.09.2016-02.10.2016	24.09.2016-03.10.2016	22.11.2015 00:00:00 - 01:00:00
Time interval	10 days	10 days	10 days	10 days (3292 snapshots)	1 hour (1800 snapshots)
Grid points	single point	single point	single point	1408 (780m ×1100m)	32400 (890m ×3590m)
Subdomain in x-range, y-range, (dx,dy)	single point	single point	single point	[20m, 800m], [-700m, -1800m], 25m	[1000m, 1900m], [895m, 4495m], 10m
Distorted data [%]	0	0	0.116 (both)	45.1	0

3.3 Reduced Order Model (ROM)

As perhaps realized by now, there exists a gap in the current methods for calculating and predicting wake in a wind farm. It seems, we can either obtain significantly simplified wake models in a short period (see section 2.4.1) or considerably time-consuming and detailed models (see section 2.4.1) [36]. Thus, there is still potential for improvement on the wake prediction methods of today in order to enable relatively precise and efficient estimations, which in turn can be utilized in wind farm layout optimization and controlling (see sections 2.4.2 and 2.4.2).

ROM is a well-established technique for reducing the computational complexity of a computer model without it going to the detriment of the global behavior of the full order model [28, 29, 34, 35, 43, 147, 148]. In other words, we want to draw conclusions for the full order model based on calculations with the ROM. As the ROM is based on data snapshots, it is beneficial to develop a data-driven ROM [33]. This enables the ROM to be "trained" by feeding it with so-called *training data* in machine learning terminology, enabling the ROM to produce estimations for periods not available (e.g. in the future) [37, 40]. ROM is a broad term with contributions from many research branches and can be obtained in various ways, but will in this work be accomplished using POD. POD has been demonstrated numerous times to be efficient within multiple engineering fields [147]. One of these fields is fluid dynamics. POD

is a proven method to significantly reduce the order of the original model while preserving a reliable approximated model [28, 31, 33, 36, 41–44]. This is mainly done through investigating and identifying the coherent structures[16] and extract these features from a random, chaotic process such as wake [28, 41]. A lower rank of the data inserted can be represented on a lower-dimensional, linear subspace, including the modes representing the most energy [28, 33]. This is accomplished by transforming the data linearly to a new orthogonal coordinate system, so that the first mode (coordinate) covers the greatest variance by any projection of the data. The second mode then covers the second largest variance by any projection of the data and so on [34]. One concern about the POD-based ROM is that it is limited to the boundaries of the data used as the basis for the model. This means that when the real wind flow is outside these boundaries, the ROM will no longer be applicable. Hence, the ROM is utterly dependent on the data input [33]. It is important to carefully consider and be aware of the boundaries of the data used for building the ROM to acquire reliable results. The boundaries are defined by the data input or the so-called 'snapshots' of the system.

3.3.1 Standard Proper Orthogonal Decomposition (POD)

In this work, the method of snapshots will be used [12, 149]. Since the main interest is to investigate the wakefield, i.e. the variations in the wind field, POD is applied only on the fluctuations[17]. If the two-dimensional preprocessed, retrieved wind velocity vector from the LiDAR data is written as $\mathbf{u} = (u, v)$, the mean is then annotated as $\bar{\mathbf{u}}$, then the fluctuations are written as $\mathbf{u}' = \mathbf{u} - \bar{\mathbf{u}}$ [33, 78]. The main idea of the POD is to utilize a modal decomposition of the dynamic wind field to obtain a lower-dimensional model of the system, as explained in section 3.3 [36]. This is done by decomposing the fluctuating flow field, $\mathbf{u}'(\mathbf{x},t)$ into deterministic spatial (POD) modes and time-dependent weighting coefficients, respectively $\phi_i(\mathbf{x})$ and $a_i(t)$ [10]. The assumes that the position vector $\mathbf{x}$ can be written as $\mathbf{x} = (x, y)$. The fluctuating flow field is then described by the 2D LiDAR data as follows [10, 32, 33, 45]:

$$\mathbf{u}'(\mathbf{x},t) = \sum_{i=1}^{M} a_i(t)\phi_i(\mathbf{x}), \tag{3.9}$$

where, t represents time, and M is the number of snapshots. The wind field, $\mathbf{u}(\mathbf{x},t)$, is obtained by adding the mean $\bar{\mathbf{u}}(\mathbf{x},t)$ to $\mathbf{u}'(\mathbf{x},t)$. Despite that the decomposition procedure is linear, it does not neglect the non-linearities of the original dynamical system [38]. The spatial modes, $\phi_i(\mathbf{x})$, representing the spatial dependence, are referred to as the orthogonal POD modes. The weighting coefficients, $a_i(t)$, characterizes all time dependence and can be understood as a variable capturing the energy of each POD mode. These variables are approximated by calculating the covariance matrix of the fluctuating vector field (mean removed).

In order to obtain the spatial and temporal weighting coefficients required for the ROM, the eigenvalue problem must be solved. In this work, it is approximated by calculating the covariance matrix [10]. Firstly, the time dependent fluctuating velocity field (mean removed) is represented in a snapshot matrix $\mathbf{U}$:

[16] Coherent structures are repeatedly appearing spatial features that undergo a characteristic temporal life cycle. Coherent structures often appear in flows that are dominated by local shear [41].

[17] The mean is removed because it is considered as a disturbance.

$$\mathbf{U} = \begin{bmatrix} \vdots & \vdots & & \vdots \\ \mathbf{u}_1(\mathbf{x},t) & \mathbf{u}_2(\mathbf{x},t) & \cdots & \mathbf{u}_m(\mathbf{x},t) \\ \vdots & \vdots & & \vdots \end{bmatrix}_{N\times M}, \tag{3.10}$$

where N represents the length of both the u and v components. The matrix is, in this work, a composition of the (processed) u and v data retrieved from the LiDAR data or WRF-PALM data. Next, the covariance matrix, $\mathbf{C}$ is created as follows:

$$\mathbf{C} = \mathbf{U}^T\mathbf{U}, \tag{3.11}$$

where $\mathbf{C}$ is a $M \times M$ matrix. By exploiting tools from linear algebra, we obtain the eigenvectors and eigenvalues, λ_i, from the covariance matrix [10]. The eigenvectors represent spatial patterns/modes of the wind field and are obtained as an orthogonal matrix, $\boldsymbol{\psi}$. The eigenvalues are non-negative values storing the energetic contribution (variance) in the various POD modes. The eigenvalues and corresponding eigenvectors are hierarchically arranged by the energy of the eigenvalues, such that $\lambda_1 > \lambda_2 > \ldots \lambda_M$ [41]. Once the eigenvalues and eigenvectors are determined, the spatial modes and temporal dynamics are, respectively, determined by [149–152]:

$$\phi_i(\mathbf{x}) = \mathbf{U}\sum_{i=0}^{M} \boldsymbol{\psi}_i \frac{1}{\sqrt{\boldsymbol{\zeta}_i}}, \tag{3.12}$$

and:

$$a_i(t) = \sum_{i=0}^{M} \boldsymbol{\zeta}_i \boldsymbol{\psi}_i^T, \tag{3.13}$$

where $\boldsymbol{\psi}$ is a square matrix of eigenvectors and $\boldsymbol{\zeta}$ is a diagonal matrix of eigenvalues.

Based on the energy-containing eigenvalues, we chose a reasonable number of modes to include in the ROM, the so-called rank, r. The total wind field seen in equation (3.9) is then truncated so that the following equation holds:

$$\mathbf{u}'^{(r)}(\mathbf{x},t) = \sum_{i=1}^{r} a_i(t)\phi_i(\mathbf{x}) \approx \mathbf{u}'(\mathbf{x},t). \tag{3.14}$$

This is possible since the eigenvalues are arranged in descending order or in nature. The accuracy of this model thus depends on the number of POD modes included in the ROM [10, 33].

3.3.2 Stochastic Model

As seen previously, from the equation (3.14) of the standard POD, the wind field is described through deterministic temporal and spatial weighting coefficients, $a_i(t)$ and $\phi_i(\mathbf{x})$, respectively. However, do the temporal weighting coefficients not include any modeling of the stochasticity of the system, but are rather deterministic. Hence, the POD can be considered a spatial filter of some kind [10]. As the wind field behind a turbine is a stochastic phenomenon, the POD can be improved by including these features [33]. This could perhaps reduce the number of modes

required in the reconstruction. One way of achieving this is by replacing the deterministic time-dependent weighting coefficients, $a_i(t)$ from the standard POD with the stochastic time-dependent weighting coefficients, $\tilde{a}_i(t)$, yielding [10, 33]:

$$\tilde{\mathbf{u}}'^{(r)}(\mathbf{x},t) = \sum_{i=1}^{r} \tilde{a}_i(t)\phi_i(\mathbf{x}). \tag{3.15}$$

Including stochasticity by altering the temporal weighting coefficients as demonstrated above, we can include more turbulence, which is crucial when investigating power output and turbine loading in a wind farm. The result is more precision in the reconstructed wind field. Furthermore, the approach described in the next section 3.3.2 allows us to make predictions for $\tilde{a}_i(t)$ at unknown times, which are used to produce (forecast) the corresponding wind field, $\tilde{\mathbf{u}}'^{(r)}(\mathbf{x},t)$. There exist several ways of doing this. In this work will, the standard *Gaussian Process* be investigated closer [33].

Gaussian Process (GP)

In order to predict something unknown, i.e. a real-life process such as wake, the approach is often to leverage statistical tools. The Gaussian distribution, also known as the normal distribution, is one of the most important probability distributions for continuous variables in statistics [39]. A single random variable from the Gaussian distribution is fully characterized by the sample mean and variance, making it trivial to work with.

This is the foundation of the so-called Gaussian Process. GP can be understood as a generalization of the multivariate Gaussian distribution, like an infinite-dimensional multivariate Gaussian distribution[18] [37]. It is a probabilistic, stochastic process applicable for predicting stochastic, real-life systems such as a wake area. GP is a process (function) collecting random variables at some observed, known inputs. This is the training data mentioned previously. The inputs corresponding to the variables collected have some correlation value, depending on the distance between the input variables. As the probability distribution used to collect the random variables is a GP, is the random process characterized by the Gaussian distribution. Hence, a finite number of this collection (a finite vector) will have a joint Gaussian distribution [40, 153].

A Gaussian Process Regression (GPR) [19] can be conducted by assuming a GP prior over functions. The GP prior assumption, based on the so-called *Bayes' theorem*[20], means that one makes some assumption of what type of function one would expect to observe, before seeing the data [37]. The GPR can be considered a classical (probabilistic) regression, seeking to determine continuous quantities of interest by constructing a model based on the training data [30]. A GPR is generally the process of finding the function describing unknown data in the best manner, based on known input-output data sets [33, 37]. The GPR is categorized as a supervised learning technique[21], which is a type of Machine Learning (ML). ML is a general term for allowing the computer to learn[22] from algorithms [154]. ML is often separated into

[18] The multivariate Gaussian distribution is a generalization of the Gaussian distribution [40].

[19] GP is also known as the kriging process.

[20] Bayesian approach involves the process of converting a prior probability into a posterior probability, based on the observed data [39].

[21] Supervised learning is generally divided into regression and classification problems [37].

[22] 'Learn' as in finding patterns in the data [154].

supervised and unsupervised ML, where supervised learning simply means that the data needs to be labeled. A function generates a label that maps the label to the input data. Supervised learning is often in need of human supervision/interpretation in order to do the labeling process [154].

The GPR allows us to build a data-driven ROM. A data-driven process is driven by the aim of obtaining estimations for new, unknown test data, based on the known training data. The process of building the data-driven ROM can generally be separated into two phases, an offline and an online phase [40, 120, 151]. In the offline phase, the high-fidelity (full order) data (LiDAR and WRF-PALM snapshots) are obtained, enabling the building of an approximated model at some prescribed accuracy. This is the most time-consuming and costly process of building the ROM. The mapping between the time parameters and coefficients is reached through the GPR, trained with the snapshot data by supervised learning. New solutions are computed for the test data during the online phase, which is generally an inexpensive process [30, 40].

In short, we aim to develop a GP posterior function to compute a predictive distribution, y^* given the unknown input, t^*. The GPR conducted in this work can be formulated as follows: A finite training data set at the i^{th} mode is collected, S={(t_i^j, y_i^j), j= 1,2,...n}, of n pairs of input values, t_i^j, and noisy output values, y_i^j (label). By assuming that the relationship between the observed, noisy output, y_i^j and the corresponding input, t_j (at some mode i) follows some regression function $\hat{a}_i$ - the GP prior, we assume that $\hat{a}(t_i)$ is a stochastic (random) time series at any time t_i. Hence, as analogue to the Gaussian distribution, the GP is specified by a the mean μ_i, and a covariance function (kernel function) $k_i(t,t')$. This is displayed in the following equations [30, 33, 37, 153]:

$$\hat{a}_i(t) \sim GP(\mu_i, k_i(t,t')), \quad y_i^j = \hat{a}_i(t_j) + \varepsilon_i, \quad \varepsilon_i \sim \mathcal{N}(0, \sigma_\varepsilon^2), \tag{3.16}$$

where ε_i is an independent error term, also known as the Gaussian-noise, corrupting the GP prior. It is defined by an associated mean of 0 and variance σ_ε^2. In equation (3.15), $\tilde{a}_i(t) = y_i^j$.

Posterior Function. As previously mentioned, the aim is to utilize the training data set, S to obtain prediction values y_i^* at corresponding unknown time values t^*, such that the test data set can be written as S^*={(t^{j*}_i, y^{j*}_i), j^*= 1,2,...n^*}, of n^* pairs of test points. As follows of the Bayes approach (the GP prior assumption), the conditional probability of observing y_i at the a single input t_i, follows a Gaussian distribution such that $y_i|t = \mathcal{N}(0, k_i + \sigma_\varepsilon^2 I)$. By incorporating this knowledge provided by the training data into the prior and combining it with the predictions for the test data, we obtain a posterior function restricted by this knowledge [37]. Thus, the training and test data sets follow a multivariate Gaussian distribution, which is described as [33, 37, 40]:

$$\begin{bmatrix} y_i \\ y_i^* \end{bmatrix} \sim \mathcal{N}\Big(0, \begin{bmatrix} k_i(t,t) + \sigma_\varepsilon^2 I & k_i(t,t^*) \\ k_i(t^*,t) & k_i(t^*,t^*) \end{bmatrix}\Big), \tag{3.17}$$

where the mean function, μ is zero for both the training and predicted sets in this work, $k(t,t')$ is the prescribed kernel, $k_i(t,t^*) = k_i(t^*,t)$, and $\sigma_\varepsilon^2 I$ represents the diagonal matrix containing the assumed Gaussian noise corrupting the training data, of size $M \times M$. The GP models an

underlying true function (labels) $y_i(t)$ of the training data as one random draw from this joint distribution. Hence, can the GP model be trained by inferring predicted values for the mean function, $\mu_i(t)$, and kernel, $k_i(t,t')$ which affects the generated observed values, $y_i(t)$.

On this foundation, the new outputs can be expressed in a probabilistic manner [37]. The new predictions will also follow a normal distribution, and the predictions at the ith mode are obtained as [37, 40]:

$$y_i^*|t_i, y_i, t_i^* \sim \mathcal{N}(\mu^*, c_i(t^*,t^*)) \tag{3.18}$$

with the predicted mean function, μ^*, and kernel function, σ^* [30, 37, 40]:

$$\mu_i^* = k_i(t^*,t)k_i(t,t)^{-1}y_i^*, \tag{3.19}$$

$$c_i(t^*,t^*) = k_i(t^*,t^*) - k_i(t^*,t)k_i(t,t)^{-1}k_i(t,t^*). \tag{3.20}$$

The posterior mean function can be understood as a correction to the prior function [37]. Time-dependent coefficients at the training data labels, t^*, and/or in the future, are obtained directly from the regression model and used in combination with the standard POD modes of some chosen range to create the stochastic ROM.

Model Selection. The kernel function (determining the covariance matrix) in the GP is predefined and selected, and tuned to get the best results for the given data. The GP is based on the assumption that similar input values should have similar outputs. This foundation is included through the covariance function [37]. The covariance function measures the covariance between the outputs, $y_i(t)$ and $y^*(t)$, as a function of the corresponding inputs, t_i^j and t_i^* [37]. A high kernel function value indicates that the outputs are alike, whereas a small kernel function value implies small similarity. There are several types of kernels, which are further specified by several hyper-parameters[23]. As long as the kernel function has the properties of being symmetrical and positive definite[24], it may be used as a kernel. The flexibility induced by model selection makes the GP model flexible to fit many data types. However, finding the best model is at the same time one of the key challenges when exploiting the GPR in variable predictions. Several kernels are investigated in this work with the objective of identifying the best fit.

Every kernel has various hyper-parameter(s) that determine the kernel's shape. One recurring hyper-parameter is the length scale, l, which controls the width of the kernel [33, 155]. In other words, it reflects the input distance required for the function output to change significantly [37]. As described in the following chapter, a shorter length-scale result in more fluctuating function and visa-versa. Other hyper-parameters present in this work are γ, α, and χ. All kernels also have a σ^2 term, which is a normalization factor regulating the normalization (mean).

The first kernel investigated is the widely used kernel named the Squared Exponential (SE) covariance kernel [33, 37, 38], also known as the Radial Basis Function (RBF) [155], which follows:

$$k_i(t,t') = \sigma^2 \exp\Big(-\frac{(t-t')^2}{2l^2}\Big) \tag{3.21}$$

[23] They are defined as hyper-parameters because they do not directly specify the function, but rather the distribution over function [155].

[24] Positive definite means that the eigenvalues of the matrix are non-negative [37].

The next kernel explored, the exponential covariance function, follows:

$$k_i(t,t') = \sigma^2 \exp\Big(-\frac{t-t'}{l}\Big). \tag{3.22}$$

The γ-exponential covariance function is also investigated, and can be found in the following manner [37]:

$$k_i(t,t') = \sigma^2 \exp\Big(-\Big(\frac{t-t'}{l}\Big)^{\gamma}\Big) \quad \text{for} \quad 0 < \gamma \leq 2. \tag{3.23}$$

Next, the Rational Quadratic (RQ) covariance function is presented [37]:

$$k_i(t,t') = \sigma^2 \Big(1 + \frac{(t-t')^2}{2\alpha l^2}\Big)^{-\alpha}. \tag{3.24}$$

Finally, the periodic random covariance function is investigated [37, 155]:

$$k_i(t,t') = \sigma^2 \exp\Big(-\frac{2\sin^2(\frac{t-t'}{\chi})}{l^2}\Big), \tag{3.25}$$

where the period, χ, determines the distance between the repetitions of the kernel function [155].

3.3.3 Error Analysis of POD

In order to statistically interpret the POD performance, error calculation may be utilized. There are several ways to estimate the relative error between the reconstructed and original wind fields [152, 156, 157]. In this work, the Root Mean Square Error (RMSE) is utilized to obtain a statistical indication of the POD performance. The calculations are conducted for a small cross-section within the snapshots, for both the reconstructed vector field, $\tilde{\mathbf{u}}$ and the original fields, u. The calculations are performed as follows [156, 158]:

$$\sqrt{\frac{1}{n}\sum_{p=1}^{n}(\mathbf{u}_p - \tilde{\mathbf{u}}_p)^2}. \tag{3.26}$$

Chapter 4

Results and Interpretation

In the following chapter, the overall results from conducting the method are presented and interpreted. Firstly, results from the retrieval and processing of the LiDAR data are displayed[1]. Next, the visual and RMSE calculations of the standard POD on the retrieved and WRF-PALM data are elaborated on. Finally, the results from the stochastic POD are described for both LiDAR and WRF-PALM data.

4.1 Data Processing of LiDAR Data - Second Period

4.1.1 Data Retrieval

As the wind speed and direction must be retrieved from the radial velocity by utilizing reference wind speed and direction, it is desired to validate the data. In this case, the wind speed and direction are validated using cup anemometer data at 100m a.s.l. and wind vane at 90m a.s.l. obtained at FINO1, as elaborated on in section 3.2.2, and SCADA data from AV7 at $\sim$90m a.s.l. found in section 3.2.3. Figure 4.1 displays normalized wind speed (upper) and direction (lower) data for the entire second period investigated[2]. The data is resampled in a 10 minute resolution for both SCADA (green), reference (yellow), and retrieved (blue and red) data, accompanied by one standard deviation for the reference data displayed in grey.

By investigating figure 4.1, we see that the 10 minute, normalized retrieved wind speed and direction appear to follow the respective reference and SCADA data reasonably well. We detect a limited number of red dots for the wind speed, which are considered outliers concerning one standard deviation from the reference value. For the wind direction, it appears as if there are more deviations present. Small clusters are gathered at the 26^{th}, 27^{th}, 30^{th} of September, and 3^{rd} of October. Otherwise, only a few data points are marked red.

However, we must emphasize that the values presented in this figure are resampled in 10 minute values. Hence, 40 minutes every hour for the LiDAR data, which have not been recorded as explained in the section 3.2.2, are present as linearly interpolated values. Therefore, the figure indicates the data quality where outliers are detected but provides no infor-

[1] As the other data is barley/not processed, it is no presented in this chapter.

[2] The 03.10.2016 is excluded from the figure due to the SCADA data. All results are, however, including this day as well.

mation on data quality where the data appears to be following the validation data. Since the wind speed and direction are derived from the retrieved u and v data, see equations (3.7-3.8), it is reasonable to link outliers detected in the wind speed plot with the wind direction and visa-versa. Thus, it is conceivable that the outliers detected for wind direction at the 26^{th}, 27^{th}, 30^{th} of September, and 3^{rd} of October also apply at the respective time for the wind speed. In summary, outlier detects areas needing closer inspection.

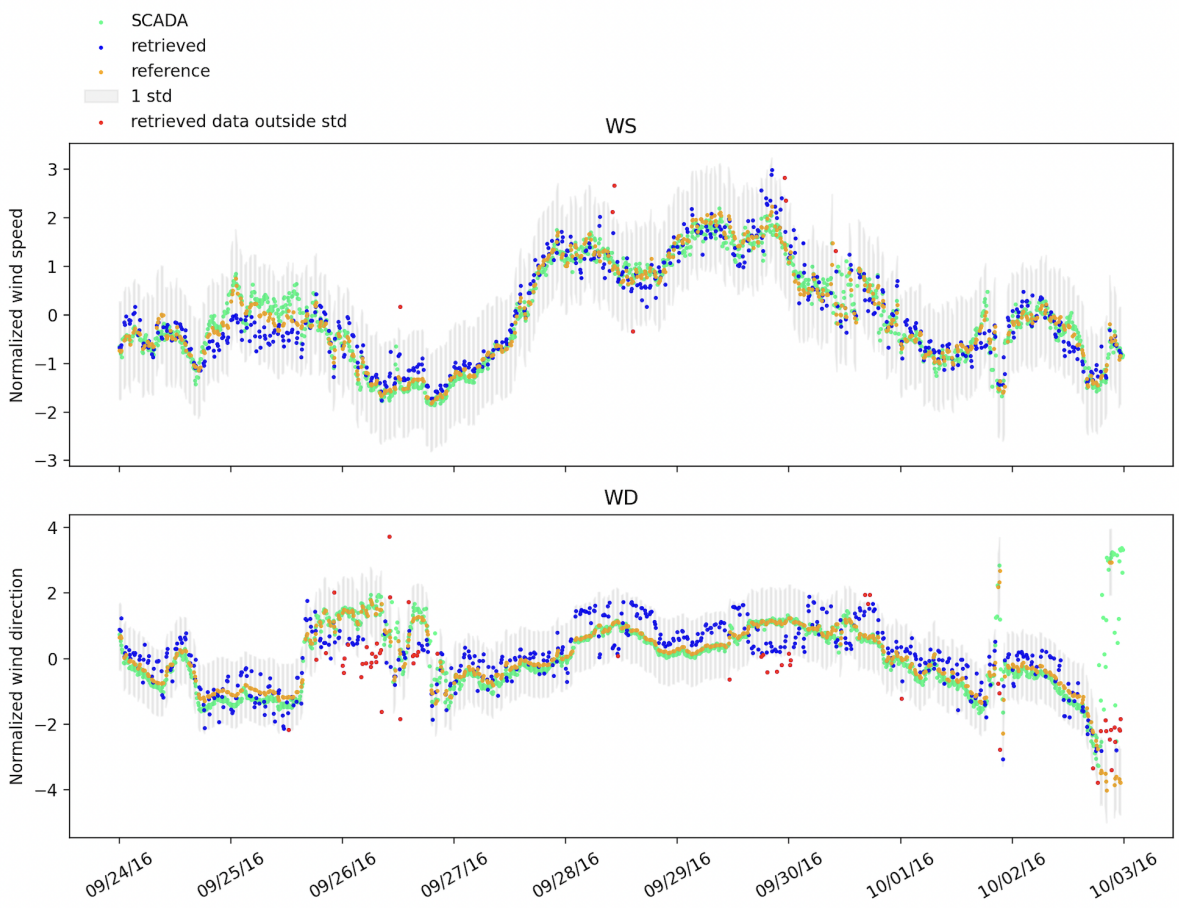


Figure 4.1: Figure illustrating the normalized wind speed (upper) and direction (lower) for SCADA, retrieved (LiDAR), and reference (cup and vane) data. The period investigated is from 24.09.2016 to and including the 02.09.2016. For comparison, the data sets are resampled in a 10 minute temporal incline.

4.1.2 Insufficient Data

Figure 4.1 implies that the retrieval algorithm performed well throughout most of the investigated period. By examining insufficient data (i.e. available measurements per scan), we gain a more comprehensive picture of the retrieval performance. To obtain meaningful wind speed and direction values from the radial velocity recorded by the LiDAR, a threshold of 95%[3] measurements available per scan were established in section 3.2.2. With this threshold in mind, figure 4.2 presents the insufficient data scans (snapshots) as red dots, with reference wind speed in the upper image and reference wind direction in the lower[4] These snapshots,

[3]The test was conducted before outliers were removed.

[4]Every 25^{th} bad-quality scan (described in section 3.2.2) is excluded from the illustration.

marked in red, are categorized as inadequate for the retrieval algorithm to perform well and are therefore removed entirely from the retrieved data set.

It is noticeable that the 28^{th}, 29^{th}, and 30^{th} of September and 3^{rd} of October are left with few snapshots. For a complete overview of distorted data per day, see Table 4.1. On the respective days, the distorted data was reported to be 96.7%, 92%, 88.3%, and 77.8%.

The wind speed figure (upper) shows no evident pattern generating the insufficient snapshots. However, one possible explanation emerges when looking at the wind direction (lower). A clear relationship arises between the insufficient snapshots and the blue bands, indicating crosswind effect (ranging from 30° to 80° and from 210° to 260°), also seen in figure 3.11. One can therefore imagine that the majority of the bad-quality snapshots (besides every 25^{th} scan) are caused by the crosswind effect. Due to the similarities in outlier clusters found in figure 4.1, snapshots measured in the crosswind directions were examined and found to be of bad-quality. Examples of this are depicted in appendix 6.

As previously mentioned, the turbines are scanned by the LiDAR of the type AREVA M5000-116, which has a cut-in wind speed of 4m/s and a cut-out wind speed of 25m/s (see Table 3.1). The lower grey band found in the retrieved wind speed image indicates cut-in wind speed, i.e. when the turbine is not operating. We expect the turbine to induce minimal wake and turbulence in these periods. The graphic does not include the region indicating cut-out wind speed because no data values were detected above ∼25m/s. From visual inspection, it appears that the non-operating snapshots are of high-quality, and are not removed in this work due to the large amount of data discarded.

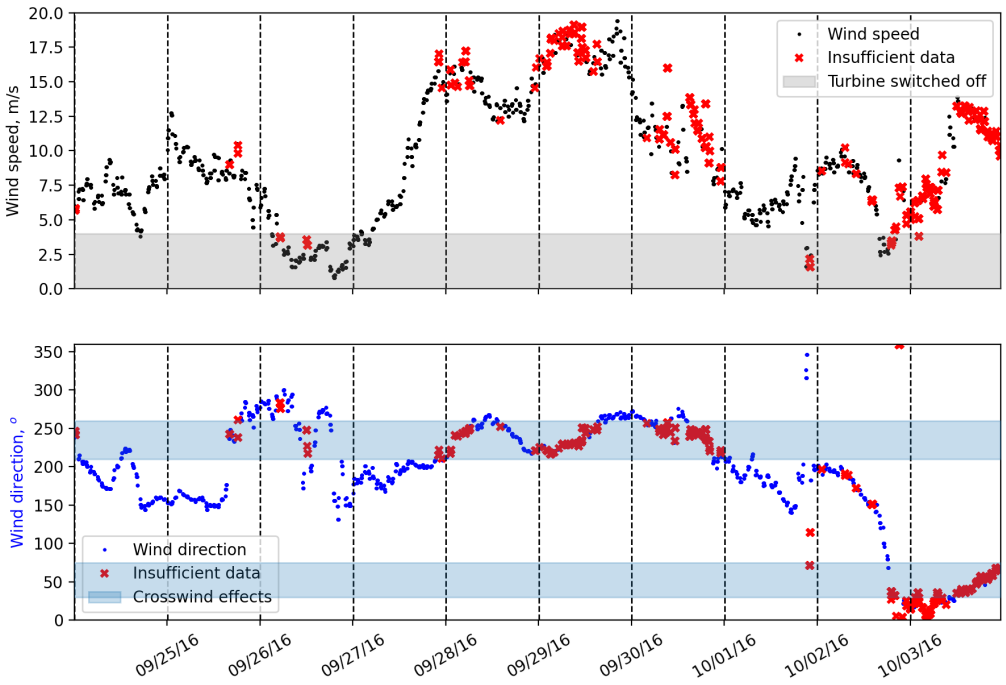


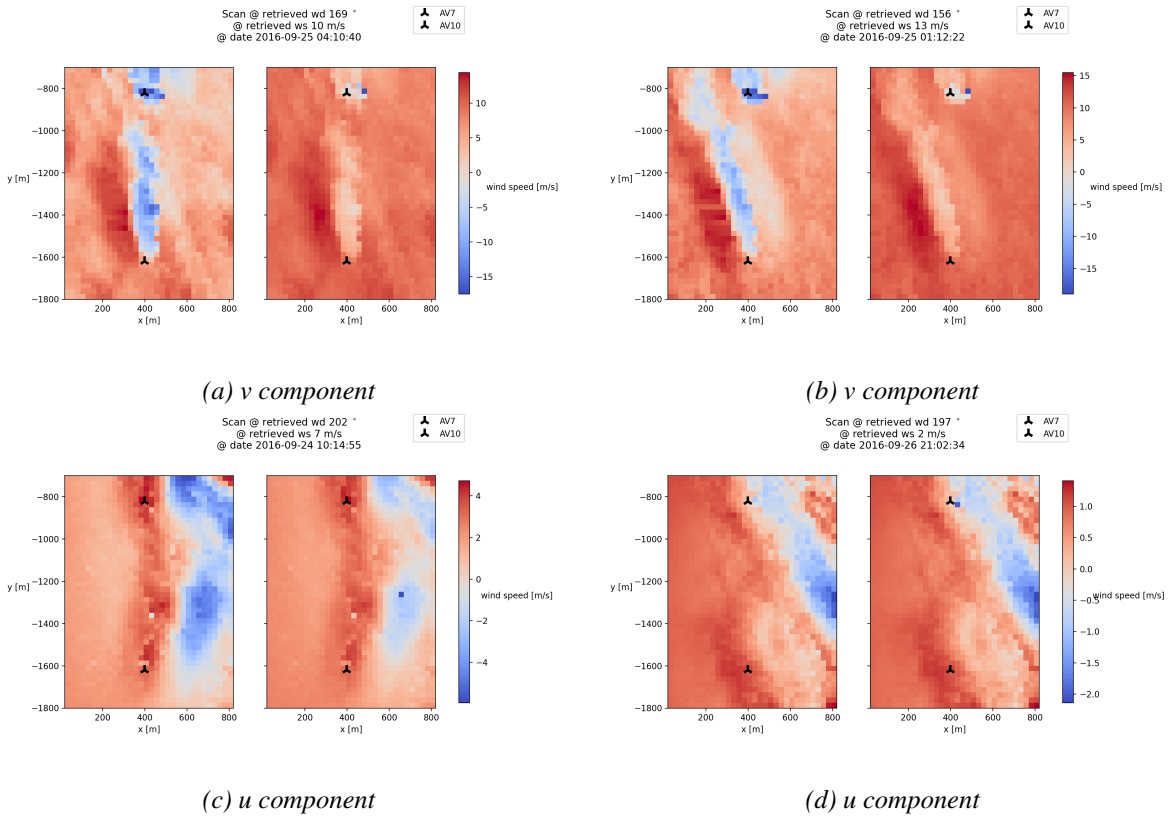
Figure 4.2: Overview of insufficient data as a function of time for both wind speed (top) and direction (bottom). The grey band, indicating wind speed below cut-in wind speed, is below 4m/s. The blue band's indication crosswind effect in the lower image ranges from 30° to 80° and from 210° to 260°.

Table 4.1: Table overview of removed snapshots per day from the final processed retrieved data used for POD.

Distorted data										
Date (2016)	24.09	25.09	26.09	27.09	28.09	29.09	30.09	01.10	02.10	03.10
%	24.2	32.3	53.5	9.3	96.7	92	88.3	15.3	20.5	77.8

4.1.3 Local Peak Removal

The two final steps in the data processing of the LiDAR data presented in section 3.2.2, are local peak removal and interpolation. In figure 4.3, four visually improved snapshots that have undergone local peak removal and interpolation are presented alongside the original velocity field. The figures 4.3 a-d have the processed image on the left and the original image on the right. Moreover, are figure 4.3 a and b of the u velocity component, whereas figure 4.3 c and d are of the v component. We see clear visual improvements for the u component, in the sense that the wake dynamics become more prominent when some local peaks around AV7 are removed. The fields that initially were monotonic at 0m/s become more dynamic. We find at least one clear peak in the lowermost images, resulting in increased wind speed variability in the processed snapshots. Hence, the illustrations suggest that the local peak removal algorithm combined with linear interpolation increases the resulting image at the time incidents are presented. However, snapshots not affected by the algorithm were detected. Moreover, new local peaks occurred as local peaks were removed in some cases.



(a) v component *(b) v component*

(c) u component *(d) u component*

Figure 4.3: Figure illustrating the improvement in data quality due to local peaks removal, by comparing original and processed snapshots. The leftmost snapshot in each subfigure is of the processed data (local peak remover and interpolation), whereas the image on the right is the original snapshot (in each figure). The upper row (a) and (b) reflects the v component of the velocity field, and the lower row (c) and (d) are the u field.

4.2 Reconstruction of the Velocity Field

By varying the number of POD modes (rank), r, to include in the approximation of the original field, as described by equation (3.14), one can visually compare the reconstructed wind fields with their corresponding real wind fields. By increasing the rank, we find that adding more modes increases the accuracy of the reconstructed wake area, as expected. Even though the snapshots provide some visual indication of how much energy the modes cover, it is beneficial to exterminate the eigenvalues following the mode examined. This section presents the results for both the standard and stochastic POD.

4.2.1 Standard POD

The following section presents the results from the standard POD of wind speed (u and v component), obtained for both retrieved and WRF-PALM data, to facilitate comparison. The results are separated into three contributions; the energetic coverage, visual inspection, and the RMSE calculations. Firstly, are the variance and cumulative variance (energetic coverage) visualized, thereafter are the outcomes of the POD (visual inspection) presented, and finally are RMSE calculations displayed.

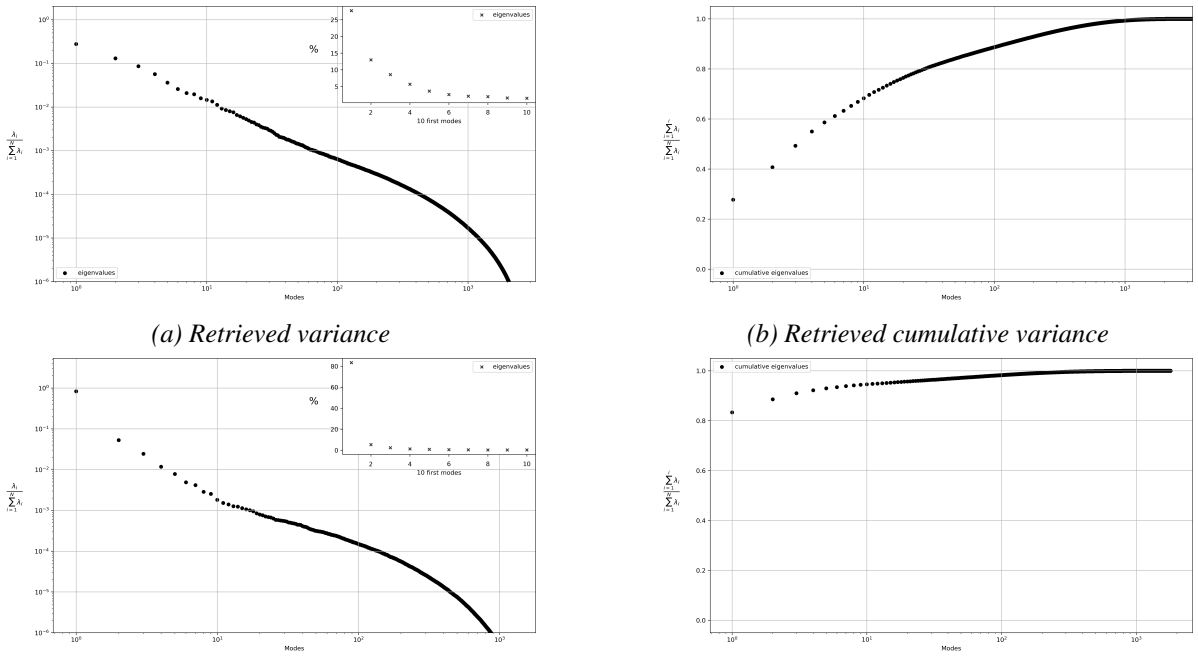
Figure 4.4 a and c exhibits the variance of the POD modes, and figure 4.4 b and d the respective

cumulative variance increasing with POD modes for the retrieved and WRF-PALM data, respectively. By choosing a rank to reconstruct the wind field according to, the reproduced wind field contains some percentage of the total energy captured by the entire covariance matrix. The optimal rank can be indicated by investigating the eigenvalues of this matrix. Hence, figure 4.4 indicates the minor rank possible to reconstruct the wind fields efficiently, by providing information about the energy of each POD mode.

Figure 4.4 b reports that the first mode captures $\sim$30% of the total kinetic energy captured by the entire retrieved data matrix displayed equation (3.17). From the subplot in figure 4.4 a, we get the impression that the first 6-7 modes represent most of the energy, and thereafter the energetic contribution decreases significantly. From inspecting figure 4.4 b, we find that the energetic coverage is >60% at 7 modes. Figure 4.4 a exhibits that close to $\sim$70 modes are required before the energy contribution is in the order of 10^{-3}.

Figure 4.4 c shows that the first mode captures $>$80% of the total kinetic energy for the WRF-PALM data. Hence, the remaining modes are responsible for the residual of $\sim$20%. The shape of both figure 4.4 c and d suggests that the increase in energy by adding more modes is modest compared with what was for the LiDAR data (see figure 4.4). Figure 4.4 c indicates that only 2 to 5 modes effectively capture enough energy to reconstruct the wind fields. From the fifth mode on, the energetic contribution by increasing the number of modes is minimal. From inspecting figure 4.4 d, we find that the energetic coverage is >90% at 5 modes. Additionally, information found in figure 4.4 c is that only $\sim$20 modes are required before the energetic contribution by adding more modes is in the order of 10^{-3}.

Hence, the variances indicates $\sim$7 and $\sim$5 modes for the retrieved and WRF-PALM data, respectively. However, we notice that the difference in energetic contribution at the respective modes for the two data sets are relatively great.



(a) Retrieved variance

(b) Retrieved cumulative variance

(c) WRF-PALM variance

(d) WRF-PALM cumulative variance

Figure 4.4: Eigenvalues (a) and cumulative eigenvalues (b) are obtained from the covariance matrix of retrieved snapshots, whereas (c) and (d) is obtained from the WRF-PALM covariance matrix.

By combining variance and modal plots, we obtain a more comprehensive overview of the energetic contribution related to each POD mode. With the information provided by figure 4.4 in mind, we can now explore the modes with their respective variance. The overall results of the two data sets are represented at a chosen time and date. The retrieved data used to represent the results is 01.10.2016 at ∼16:17, and the WRF-PALM data is from 22.11.2015 at ∼00:33. Firstly, the results for the retrieved data are depicted in figure 4.5, and then the results for the WRF-PALM data are presented in figure 4.6.

In figure 4.5, the results from applying POD to the retrieved data are depicted. The reconstructed fields are found on the left, and the original wind fields on the right. The u field is found in the upper figures, and the v field in the lower figures. The modes presented are within the range from 1 to 501 POD modes included in the reconstruction.

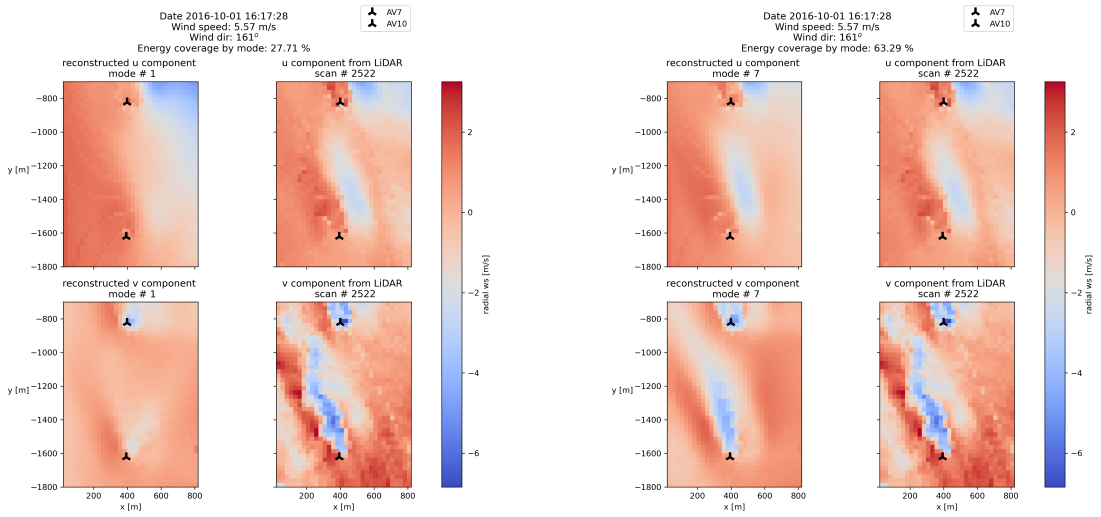
Probing figure 4.5 a, illustrating the reconstructed field of one single mode, we see that the reconstructed u and v fields mirror the respective initial snapshots rather poorly. These reconstructions provide an indication of the turbine locations and the wake orientation for the v field but cover no dynamics beyond this. As stated by Basitne et al., the first mode simply reflects the large-scale horizontal motion of the wake [36]. Given that the respective energy captured for this mode is ∼28%, this is as could be expected.

Assuming that the ∼7 first modes, which were found to capture the majority of the energy for the retrieved data in figure 4.4, were enough to reconstruct the wind fields efficiently, the energy captured by the reconstructions would be ∼63%. From a visual inspection of 7 modes seen in figure 4.5 b, we find elements of the large-scale dynamics of the wakes in the original wind fields. Since the u field at this time does not represent that many fluctuations, it appears as if the u field is better described by this reconstruction. However, regarding the v field, we can see clear discrepancies between the reconstructed field and the original field. It appears as if the meandering and velocity recovery distance (see section 2.3.1) is not present. Moreover, the centerline appears to be somewhat skewed (clockwise) compared with the original wake centerline. These features are, as mentioned previously, of great importance for downstream turbine loading and power calculations.

By inspecting figure 4.5 c including 21 modes, we see that 77% of the total energy is captured in the reconstruction. The reconstructions appear to be significantly improved from the 7 modal reconstructions. Moreover, the reconstruction for the u field appears to be, more or less, fully described by 7 modes. However, clear deviations are still detected for the v field.

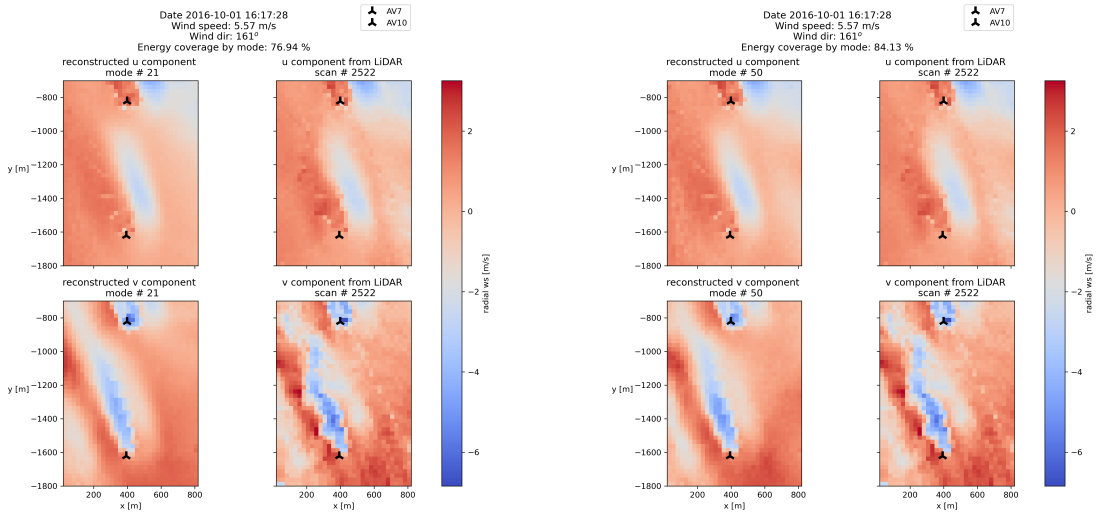
As 50 modes are included, seen in figure 4.5 d, the reconstructed v field is starting to take shape. Figure 4.5 e, depicting the reconstructions utilizing 101 modes, the meandering dynamics in the original v field is clearly present. It is, however, noticed that by including 501 modes, presented in figure 4.5 f, the improvements are even greater compared to the reconstruction using 101 modes.

From the visual inspection, we get the impression that at least 50 modes are required in order to recover the features seen in the original snapshot to some degree. The energy contribution is above ∼84% for 50 modes.



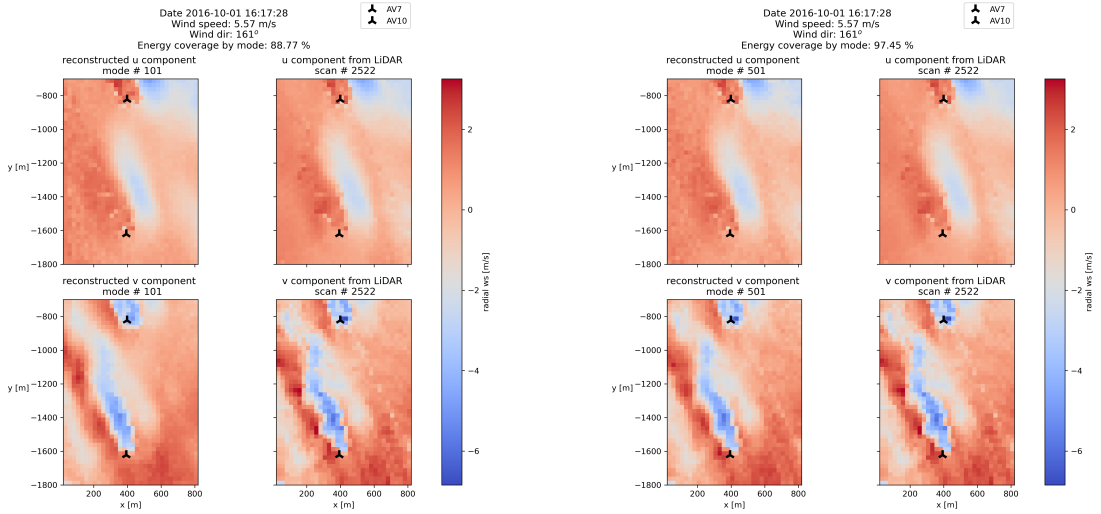
(a) 1 POD mode

(b) 7 POD modes



(c) 21 POD modes

(d) 50 POD modes



(e) 101 POD modes

(f) 501 POD modes

Figure 4.5: Figures of the accumulated POD modes (i) used to reconstruct the retrieved data. In every subfigure, the two upper figures are of the u field, whereas the lower ones show the v field. The reconstructed fields are found in the left figures, and the original snapshots on the right. The number of modes used in the reconstructions are (a) i = 1 (b) i = 7 (c) i = 21 (d) i = 50 (e) i = 101 (f) i = 501.

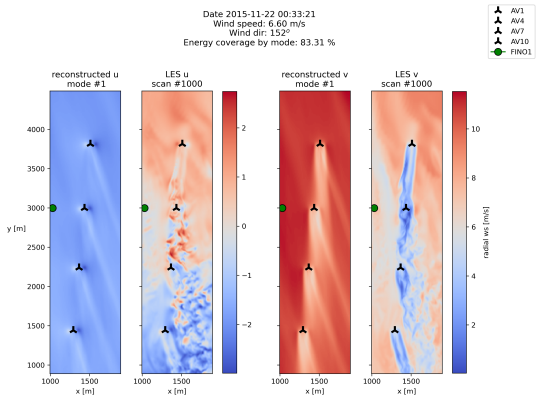
Figure 4.6 is arranged with the reconstructed and original u field in the two images on the far left before the reconstructed and original v field is presented in the two images furthest to the right. The same number of modes seen in figure 4.5 is used for the WRF-PALM data.

Starting by exploring figure 4.6 a visually, which illustrates the reconstructed u and v fields of the first mode, the original wind fields are not well recovered. It rather indicates the direction of the wakes, as it did for the retrieved data. Moreover, as opposed to the retrieved data, it appears to overestimate the fields' values greatly. This is the case, despite an energy coverage of $\sim$83.3%. As indicated by figure 4.4 c and d, two modes are rejected as sufficient for the reconstructions of the original snapshots.

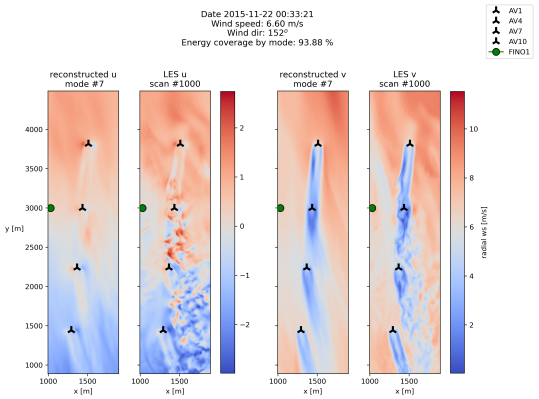
By investigating the reconstruction of 7 modes (figure 4.6 b), we see significant improvements from the one modal reconstructions. The large-scale structures of the wake appear to be present in both wind fields. However, the small-scale structures seems to be neglected entirely.

The reconstruction from 21 modes, figure 4.6 c, appears to perform very much the same. Moving on to figure 4.6 d, the reconstruction of 50 modes is displayed. These indicated that significantly more small-scale features of the wakes are present for the u field. However, the improvements in the v field are somewhat modest in comparison. From figure 4.6 e, we get the impression that by adding 101 modes in the reconstructions, small-scale dynamics are present in both wind fields. It is however a significant change seen in the reconstructions utilizing 501 modes (figure 4.6 f), where both wind fields appear to be more or less perfectly described.

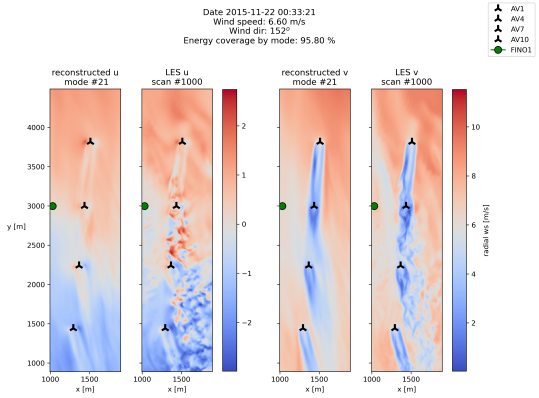
From the visual inspection of the reconstructed fields for the WRF-PALM data, we find that including a minimum of 101 modes in the reconstructions captures the small-scale turbulent structures relatively well. The energy contribution for 101 modes is $\sim$98%. Recovering only large-scale features appears to be achieved with only 7 modes.



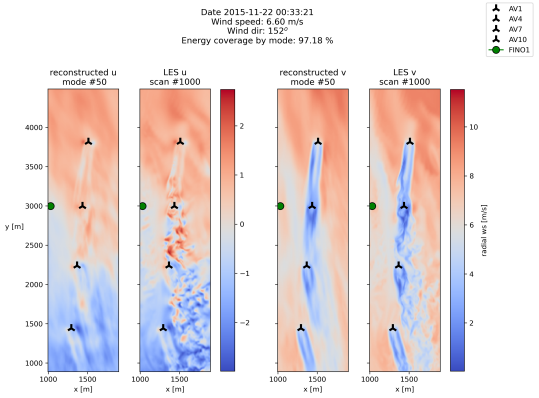
(a) 1 POD mode



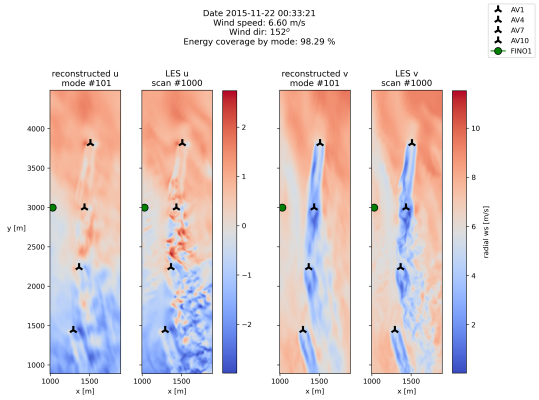
(b) 7 POD modes



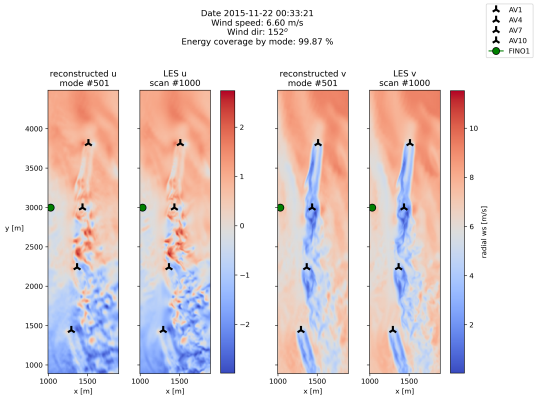
(c) 21 POD modes



(d) 50 POD modes



(e) 101 POD modes



(f) 501 POD modes

Figure 4.6: Figures of the reconstructed and original u and v components of the wind field obtained from the WRF-PALM data. In every subfigure, the reconstruction at mode i of the u and v fields is presented, from left, in the first and third snapshot. Their respective original snapshots are found in the second and forth image, from left. The reconstructions are done with modes (a) i = 1 (b) i = 7 (c) i = 21 (d) i = 50 (e) i = 101 (f) i = 501.

RMSE. The Root Mean Square Error (RMSE) is calculated for a small area in front (without wake) of and behind the rotor (with wake) and an area of interest. Henceforth, these are referred to as area 1, area 2, and area 3, respectively. The areas for the retrieved data are illustrated in figure 4.7 a and b, and figure 4.7 c and d for the WRF-PALM data. The relative error for the respective areas are displayed in Table 4.2 and 4.3, for the retrieved and WRF-PALM data, respectively.

By inspecting the areas depicted in figure 4.7 a and b in the Table 4.2, an expected disproportional relationship is detected between the number of modes included in the reconstruction and the RMSE for the entire snapshot; an exponential increase in the number of modes results in a flattening of the RMSE. One interesting observation is made for 101 modes in the u field, where an exception from this trend is observed. Moreover, we notice that the overall RMSE for the v field in area 2 is relatively much higher than for the other areas, which is also somewhat unexpected.

From Table 4.2, we find that 21 modes appear to perform well for area 1, in front of the rotor, for both u and v. As area 1 covers the free-flow area, it may be an indication of that 21 modes performs well for this flow. However, the significant reduction is somewhat suspicious. Moreover, for 21 modes, Table 4.2 reports a relatively small error in area 3 for the v field, which mainly includes the meandering dynamics of the field. This is however not supported by the visual inspection conducted in section 4.2.1. One reason could be that area 3 is placed in the area where the tail of the wake is altering its direction (meandering). Hence, in the original snapshot, this area is generally in the value of the free flow. In section 4.2.1, we found that by including a limited number of modes, the wake tail was both shortened and the centerline appeared to be skewed some degrees clockwise compared with the original snapshot. Thus, the unexpected low RMSE value for area 3 could potentially be caused by that the area falls in a free-flow area for the 21 modal reconstructions, without them accounting for the meandering dynamics.

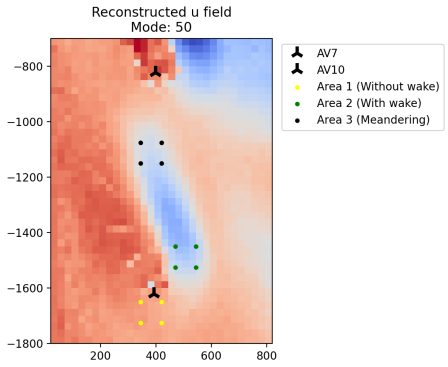
Comparing the RMSE values for 50 and 101 modes (Table 4.2), they appear to perform relatively similarly. From the RMSE values, the u field is reconstructed better with 50 modes. For the v field, on the other hand, the greatest difference is found in area 3, which mainly covers the wake meandering features. Hence, are the values provided by Table 4.2 further substantiating the findings from section 4.2.1, of that 100 modes are required in the reconstruction to capture the meandering features. By further investigating 501 modes, a significant improvement in areas 1 and 2 is observed for the v field.

Table 4.3, displaying the RMSE for the WRF-PALM data, suggests that the relative error of adding one single mode is relatively much more prominent than adding 5 modes and more. This seems to be in line with the observations made from the energetic contribution and the visual inspection, presented in section 4.2.1. As was seen for Table 4.2, the total RMSE is continuously decreasing with an increasing number of modes for the WRF-PALM data as well, with no exceptions. The RMSEs for the first mode are, however, detected as rather high compared with the RMSE calculated for the retrieved data. This is probably due to the overestimation of the wind field flow found in section 4.2.1. Therefore, neither the high RMSE values nor the significant reduction from the first mode is unexpected observations.

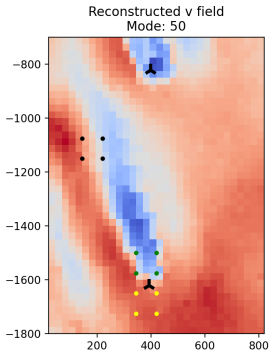
Further inspection of Table 4.3 informs that the decreasing RMSE trend is not as prominent for the WRF-PALM data as for the retrieved, and the RMSE calculations more frequently alternate. Hence, the table can be understood in various ways and is somewhat hard to interpret. However, this observation is also in line with the eigenvalues inspected in section 4.2.1. It is somewhat unexpected that the visual improvements observed in section 4.2.1 were experienced as significant, when both the RMSE calculations and energy added by increasing the number of modes appear to be relatively small. One observation in line with this observation is, however, the increase in the 501 modal reconstructions for the v field, which is reduced relatively much from the 101 modal reconstructions.

Comparing the u field calculations for 21 and 50 modes, the main differences are a significant decrease, a slight increase, and a slight decrease for areas 1, 2, and 3, respectively. A visual comparison is very much in line with the reported RMSE values. We see relatively much more structures in the free-flow area in front of the turbine (area 1), substantiating the significant RMSE decrease found, whereas, for area 2, the changes are modest. Perhaps somewhat surprisingly is the decrease observed for area 3, as the visual improvements detected for this area are rather great. By comparing the same modal reconstruction for the v field, we see that area 1 decreases, area 2 remains the same, and area 3 increases somewhat. As for the u field, are the visual improvements in the free-flow area rather great, whereas area 2 remains similar. The improvements in area 3 are not as prominent for the v field, and the results are therefore very much in line with the RMSE values.

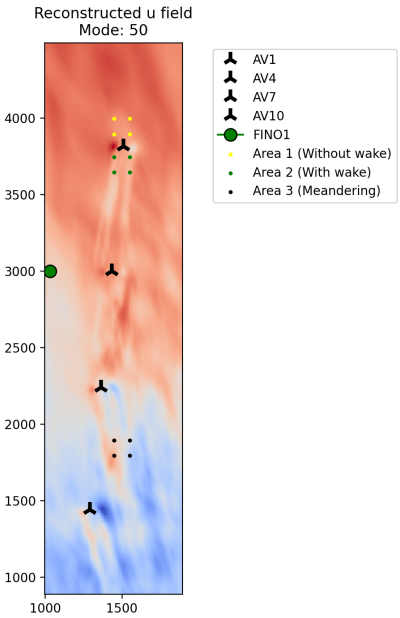
The reconstructions including 101 modes exhibit a decrease in all areas for both wind fields. The values are reflected in the findings in section 4.2.1, and 101 modes appear to partly include small-scale turbulent features. It is however inevitable that both the RMSE calculations and the visual improvement for 501 modes are significantly improved.



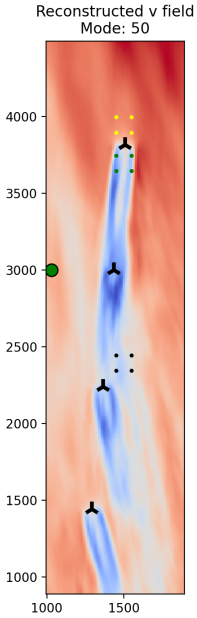
(a)



(b)



(c)



(d)

Figure 4.7: Illustration of the cross-section areas used for Root Mean Square Error (RMSE) calculation. Area 1 is defined as the yellow scatter (square) in front of the rotor, area 2 as the blue scatter right behind the rotor, and finally, area 3, indicated by black scatter, is an area of interest. We see the retrieved data field (a) u and (b) v, and the WRF-PALM (c) u and (d) v velocity components.

Table 4.2: RMSE and cumulative eigenvalues with mode, for POD on the retrieved data. RMSE is calculated for the entire snapshot, area 1, 2, and 3. For further information about the areas, see figure 4.7.

RMSE for retrieved data									
		RMSE							
		u field				*v* field			
N modes	**Cumulative energy [%]**	**Total**	**Area 1**	**Area 2**	**Area 3**	**Total**	**Area 1**	**Area 2**	**Area 3**
1	27.713	0.963	0.537	0.879	0.598	1.405	0.556	2.659	1.482
5	58.609	0.513	0.092	0.777	0.478	1.05	0.529	2.492	0.754
7	63.291	0.416	0.319	0.316	0.416	0.9308	0.91	1.709	1.058
21	76.941	0.350	0.197	0.201	0.659	0.697	0.185	1.613	0.308
50	84.130	0.254	0.225	0.181	0.166	0.572	0.628	1.283	0.378
101	88.771	0.257	0.224	0.162	0.181	0.480	0.613	1.274	0.197
501	97.453	0.188	0.109	0.177	0.2	0.272	0.219	0.700	0.157

Table 4.3: RMSE and cumulative eigenvalues with mode, for POD on the WRF-PALM data. RMSE is calculated for the entire snapshot, areas 1, 2, and 3. For further information about the areas, see figure 4.7.

RMSE for WRF-PALM data									
		RMSE							
		u field				*v* field			
N modes	**Cumulative energy [%]**	**Total**	**Area 1**	**Area 2**	**Area 3**	**Total**	**Area 1**	**Area 2**	**Area 3**
1	83.309	2.043	2.510	2.061	1.606	3.885	3.681	3.807	4.196
5	92.977	0.530	0.363	0.165	0.861	0.883	0.849	0.760	0.884
7	93.881	0.463	0.26	0.198	0.792	0.689	0.929	0.386	0.726
21	95.797	0.448	0.155	0.107	0.821	0.583	0.939	0.204	0.650
50	97.177	0.389	0.079	0.143	0.816	0.447	0.591	0.209	0.711
101	98.292	0.336	0.049	0.106	0.783	0.373	0.294	0.130	0.687
501	99.868	0.108	0.019	0.019	0.302	0.101	0.024	0.031	0.201

4.2.2 Stochastic POD

As we have seen from figures 4.5 and 4.6, the standard POD seems to provide reduced order models of the wind fields for both data sets. However, the wake dynamics appears to be significantly more smooth than the original wakes when considering only the first couple of modes ($\sim 60\%$ for the LiDAR and $\sim 90\%$ for the LES). In this section, the reconstructions of changing the standard- with the stochastic temporal weighting coefficients, (a_i with $\tilde{a}_i$) is presented. The stochastic temporal weighting coefficients are constructed utilizing GPR.

Firstly, a sensitivity study is proposed for various parameters in the kernels explained in section 3.3.2, which is required to obtain the stochastic temporal weighting coefficients. Finally, are the parameters considered as best fitted used in a stochastic reconstruction for both the retrieved and WRF-PALM data. The process is displayed in figure 4.8.

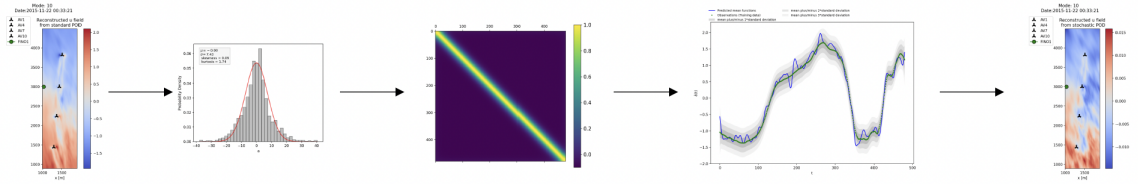
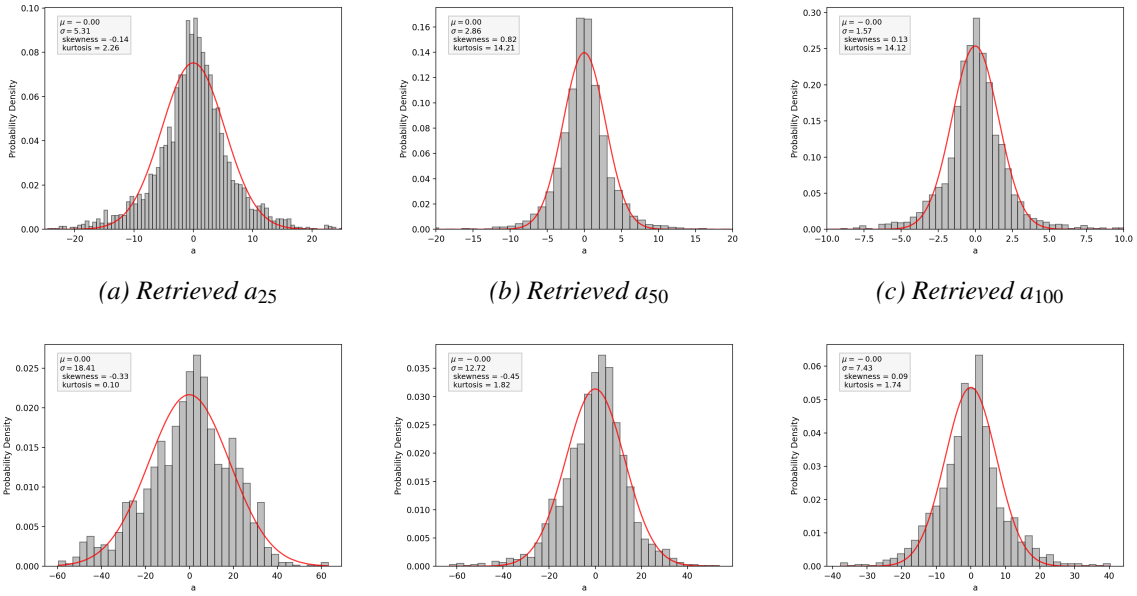


Figure 4.8: Illustrations of the process (five steps) from standard POD to stochastic POD (the figure uses WRF-PALM data). The temporal weighting coefficients obtained from the standard POD are the starting point. Next, check whether the temporal weighting coefficients follow the Gaussian distribution. The third step involves finding the best kernel and hyper-parameters for the data investigated. Thereafter, use the GPR to reproduce new, stochastic temporal weighting coefficients. We obtain the stochastic field by replacing the ones produced in the standard POD.

Gaussian Distribution. Figure 4.9 illustrates histograms representing the probability density distribution of the temporal weighting coefficient, a_i at the i^{th} POD mode, for the retrieved data (upper) and WRF-PALM data (lower). Their corresponding mean, μ, standard deviation, σ, skewness, and kurtosis are found in the figures. As implied by the fitted Gaussian distribution (red curve), the temporal weighting coefficients are relatively well described by the distribution at the presented modes, both for the retrieved and WRF-PALM data. These characteristics for a_i allow us to apply the GP to the data.



(a) Retrieved a_{25} *(b) Retrieved a_{50}* *(c) Retrieved a_{100}*

(d) WRF-PALM a_{25} *(e) WRF-PALM a_{50}* *(f) WRF-PALM a_{100}*

Figure 4.9: Gaussian distribution (red curve) fitted to a subset of the time-dependent coefficient a_i at the i^{th} mode, of the retrieved data ((a, b, and c) and WRF-PALM data (d, e, and f). Figures of mode of (a, d) i = 25, (b, e) i =50, (c, f) i = 100.

Model Selection. In order to get an impression of suitable hyper-parameters and kernels (explained in section 3.3.2) for the temporal weighting coefficients, obtained from the standard POD of the retrieved and WRF-PALM data, we conduct a brief sensitivity study.

Due to convenience and clarity in the plots, a limited time interval of 200 snapshots is chosen to investigate closed for both data sets. This corresponds to the period ∼00:16 to ∼10:15 at 01.10.2016 for the retrieved data and ∼00:00 to ∼00:07 at 22.11.2015 for the WRF-PALM data. In section 4.2.1, we found that ∼67% and ∼93% of the total energy is captured by including 7 and 5 modes for the retrieved and WRF-PALM data, respectively. The following sensitivity study utilizes the temporal weighting coefficients at these modes for the results.

The following figures, 4.10, 4.11,4.12, 4.13, and 4.14, present a selection of covariance matrix with one associated random predicted function for each kernel presented in section 3.3.2; the Squared Exponential (SE), exponential, γ-exponential, Rational Quadratic (RQ), and random periodic kernel, respectively. The figures are all structured pairwise, with the covariance matrix in the upper images and the corresponding predicted function in the lower images. The covariance matrix image visualize the covariance, $k_i(t,t)$ at mode i, between the input values at known time incidents, t. Green dots in the associated predicted functions indicates already observed values at time, t. The red dots denote test data, which are observed values that have not been used as input in the model, indicating the performance of the random predicted function for forecasts. The grey areas of 1, 2, and 3 standard deviations follow the predicted mean curve, and all predicted functions are present within this grey area.

The noise and variance are kept constant for all kernels, for both the retrieved and WRF-PALM data at 0.2 and 1, respectively. A "base case" with $l = 1$ been defined for all kernels (shown in figure a in all subfigures), and moreover $\alpha = 1$ and $\gamma = 1$ for the RQ and γ-exponential kernel, respectively.

In figure 4.10, we see that increasing the hyper-parameter l increases the width of covariance (figures 4.10 a, b, c, g, h, and i), which results in a smoother predicted function (figures 4.10 d, e, f, j, k, and l), and visa-versa. Increasing fluctuations result in a decreased standard deviation range, restricting the predicted functions more to the training data. The SE kernel appears to be able to construct both relatively smooth and fluctuating functions. Despite that the observed WRF-PALM data points (figure 4.10 j, k, and l) display a much smoother curve than the retrieved data (figure 4.10 d, e, and f) does, the SE kernel appears to be able to fit both curves. The predicted area, indicated by the red dots, appears to be more or less independent of the kernel.

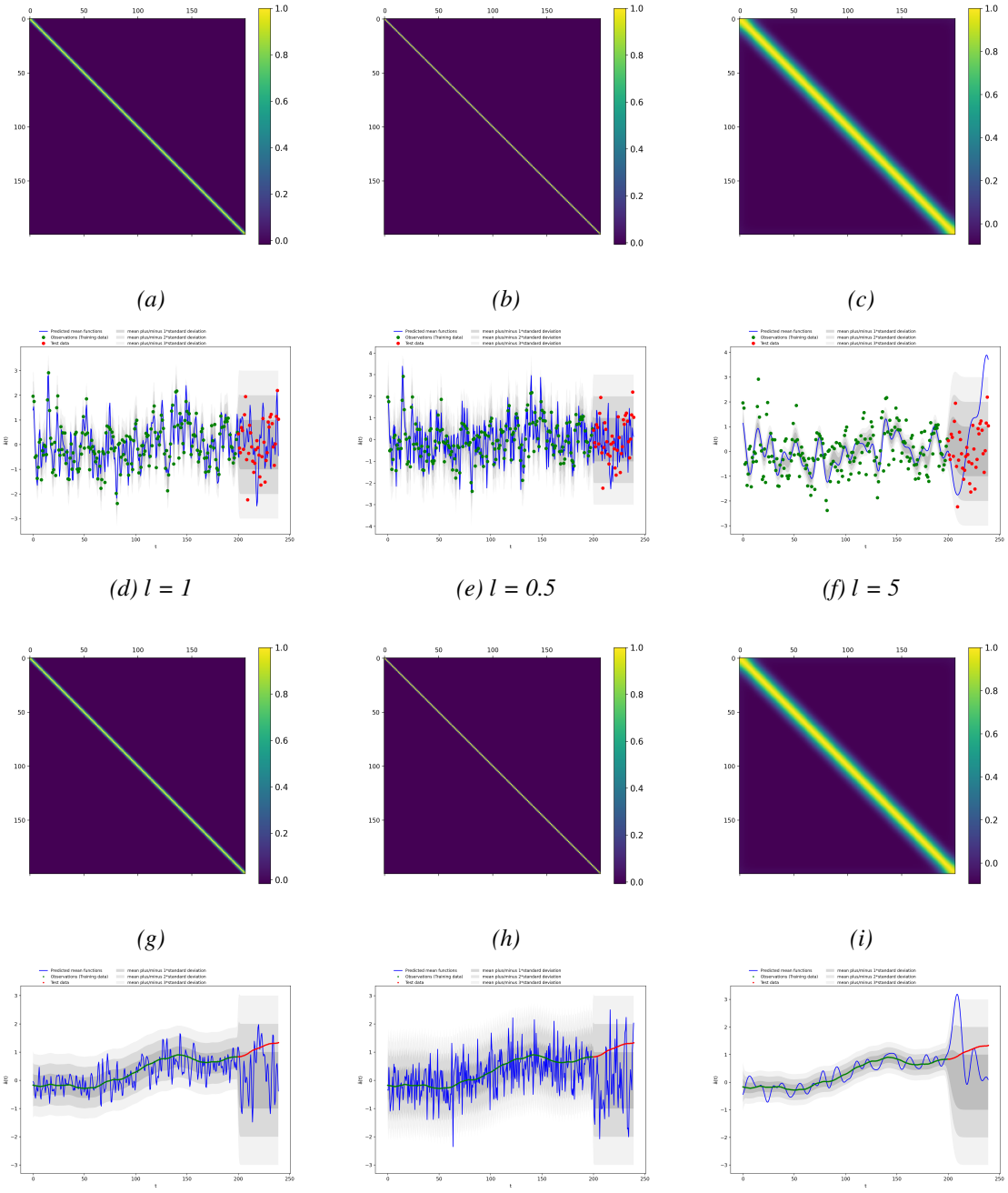
Figure 4.11 indicated that the exponential kernel generally constructs a highly fluctuating predicted function for the retrieved and WRF-PALM data. By increasing the length scale, l, we see the smoothening effect on the predicted functions. However, the effect is not nearly the same as observed for the SE kernel. Moreover, the flattening effect is more prominent for the WRF-PALM data presented in figure 4.11 l, than for the retrieved data in figure 4.11 f. It also appears as, for the WRF-PALM data, that the test data area in the predicted functions are somewhat more affected by the kernel hyper-parameters, using the the exponential kernel. From figure 4.11 l, the predictions at unknown time incidents, t^* are noticed to be much broader than what it is for the fluctuating function presented in figure 4.11 j. For the retrieved data, it appears to be very much the same as seen for the SE kernel.

The results utilizing the γ-exponential kernel, depicted in figure 4.12, is reminiscent of the exponential kernel. For the γ-exponential kernel is the length scale, l, kept constant at 1. The predicted functions are highly fluctuating for both data sets. Moreover, is the same pattern seen for the predicted functions at unknown time incidents, t^*, for the SE kernel, also found for this kernel. For the WRF-PALM data an $\gamma = 0.5$ (figure 4.12 k), the predicted function

is somewhat more narrow in the test data area, compared with the other hyper-parameters. The γ-exponential kernel is more adaptable then the exponential kernel, as it can vary both the hyper-parameters, l and γ. Therefore, one can specify the fluctuations on a small and large-scale more precisely.

The RQ kernel is displayed in figure 4.13. It is somewhat similar to the results presented for the SE kernel. For this kernel, only α is changed. The effect of changing α is less visible than what was by changing γ for the γ-exponential function. The predicted functions for the retrieved data (figure 4.13 d, e, and f) and the WRF-PALM data (figure 4.13 j, k, and l) appear to be very much alike. Extremities of the hyper-parameter was tested, and it was challenging to identify any difference in the predicted functions because the functions are random. A slight difference was detected by investigating the covariance matrices, which is visible from the inspection of the covariance matrices. We found that the covariance matrix becomes somewhat more narrow with increasing α. Additionally, we found that it performed similar to the SE when altering the length scale, l.

The random, periodic kernel, illustrated in figure 4.14, behaves quite differently from the other kernels. In general, it appears to be somewhat more sensitive to changes. The pattern of the temporal weighting coefficients at mode 7 for the retrieved data appears to have a random behavior. Hence, a smooth periodic function was challenging to fit the data, as no apparent symmetry was found. By visually inspecting the predicted functions, we find that a small length scale best fits the retrieved data, as seen in figure 4.14 b and e. The covariance function depicted in figure 4.14 c also performs well. In this case, however, only one period is used, as seen in figure 4.14 f. For the WRF-PALM data, a fit for the period of 200 snapshots investigated was not found (the results can be seen in section 6). However, by increasing the period to 600 snapshots ($\sim$20 minutes), the periodic kernel performed relatively well. As for the retrieved data, a small length scale value was found to perform the best.



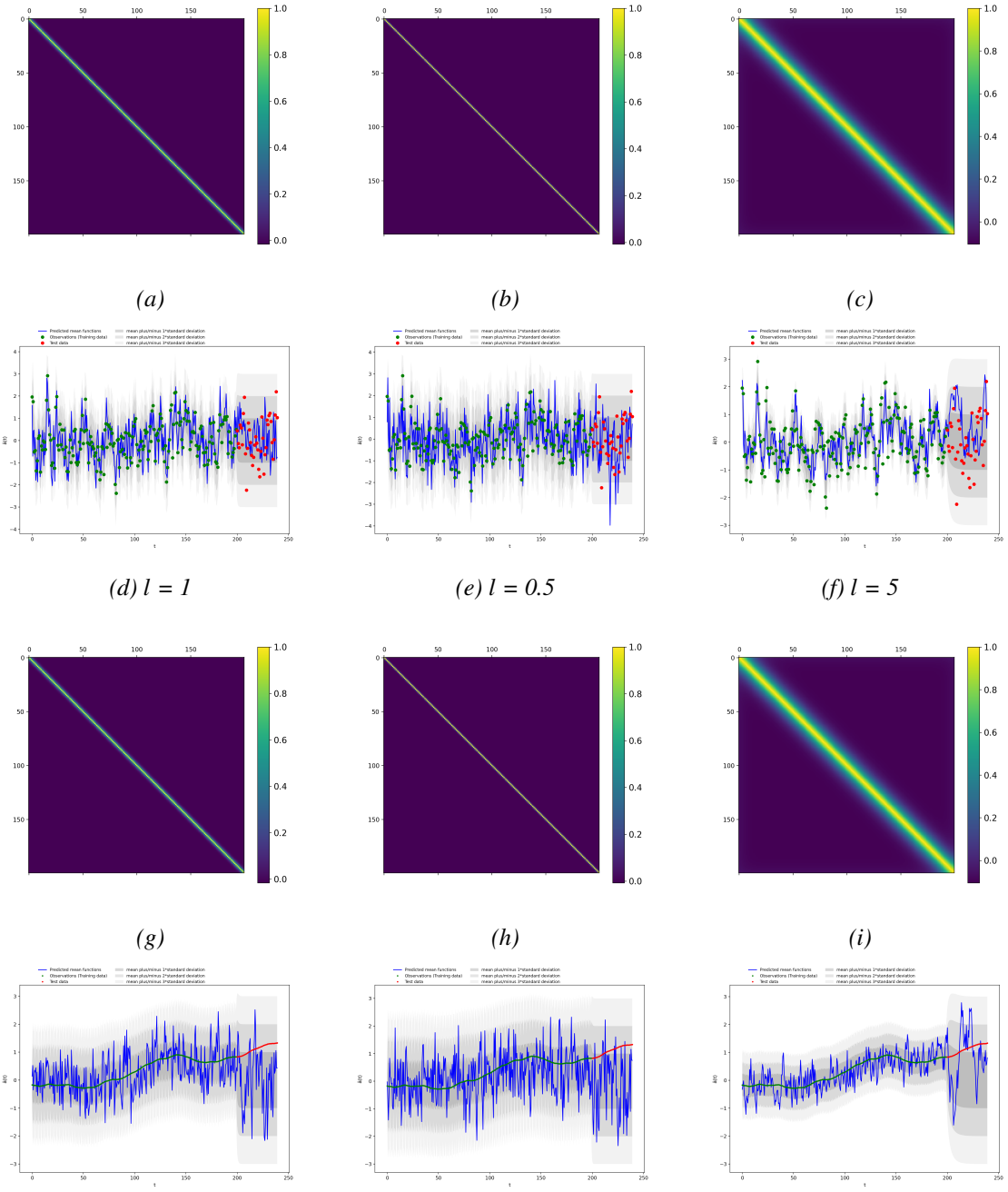
(a) *(b)* *(c)*

(d) $l = 1$ *(e) $l = 0.5$* *(f) $l = 5$*

(g) *(h)* *(i)*

(j) $l = 1$ *(k) $l = 0.5$* *(l) $l = 5$*

Figure 4.10: Covariance matrices (upper) and corresponding predicted functions (lower) using the Squared Exponential (SE) covariance function for constant noise = 0.2 and $\sigma^2 = 1$. The six upper paired figures ((a, d), (b, e), and (c, f)) are obtained from the retrieved data at a_7, whereas the six lower figures ((g, j), (h, k), and (i, l)) are from the WRF-PALM data at a_5. (a, d) and (g, j) represents the "base case", with $l = 1$. (b, e) and (h, k) are with $l = 0.5$, and finally (c, f) and (i, l) has $l = 5$.



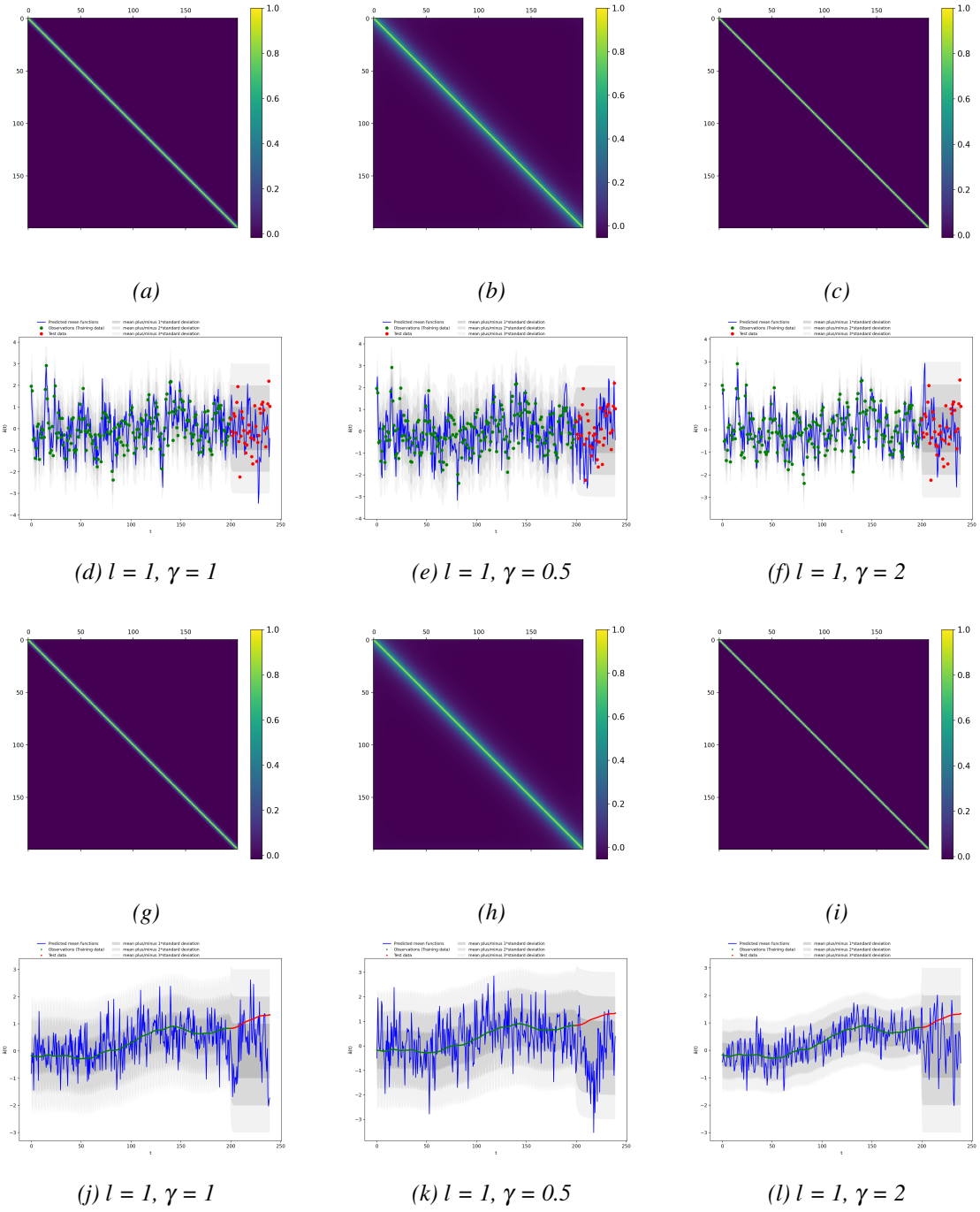
(a) *(b)* *(c)*

(d) l = 1 *(e) l = 0.5* *(f) l = 5*

(g) *(h)* *(i)*

(j) l = 1 *(k) l = 0.5* *(l) l = 5*

Figure 4.11: Covariance matrices and corresponding predicted functions using the exponential covariance function for constant noise = 0.2 and $\sigma^2 = 1$. The six upper paired figures ((a, d), (b, e), and (c, f)) are using the retrieved data at a_7, whereas the six lower figures ((g, j), (h, k), and (i, l)) are from the WRF-PALM data at a_5. (a, d) and (g, j) represents the "base case", with $l = 1$. (b, e) and (h, k) are with $l = 0.5$, and finally (c, f) and (i, l) has $l = 5$.



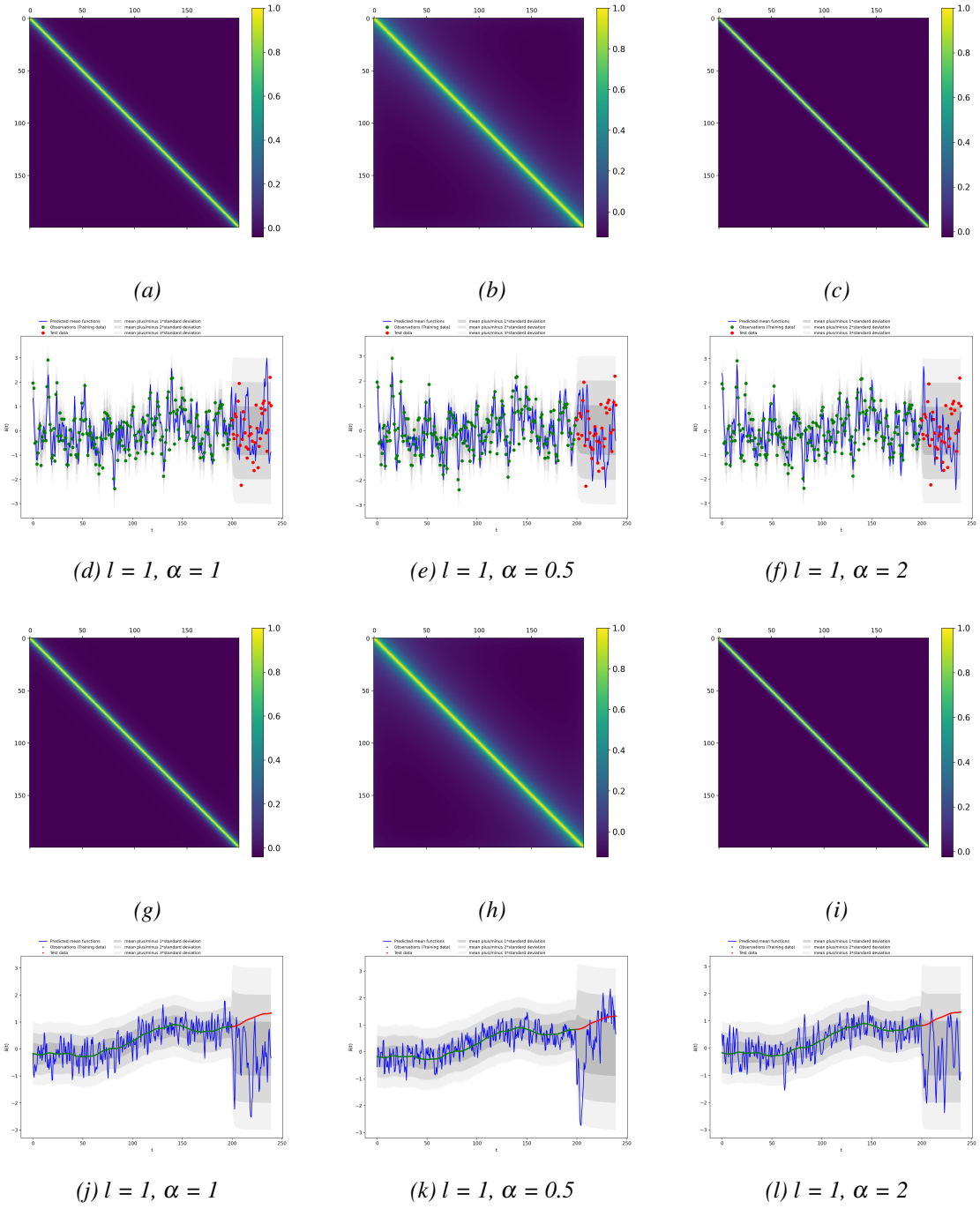
(a) *(b)* *(c)*

(d) $l = 1$, $\gamma = 1$ *(e) $l = 1$, $\gamma = 0.5$* *(f) $l = 1$, $\gamma = 2$*

(g) *(h)* *(i)*

(j) $l = 1$, $\gamma = 1$ *(k) $l = 1$, $\gamma = 0.5$* *(l) $l = 1$, $\gamma = 2$*

Figure 4.12: Covariance matrices and corresponding predicted functions using γ-exponential covariance function for constant noise = 0.2, $\sigma^2 = 1$, and $l = 1$. The six upper paired figures ((a, d), (b, e), and (c, f)) are using the retrieved data at a_7, whereas the six lower figures ((g, j), (h, k), and (i, l)) are from the WRF-PALM data at a_5. (a, d) and (g, j) represents the "base case", with $\gamma = 1$. Next, (b, e) and (h, k) have $\gamma = 0.5$, and finally (c, f) and (i, l) have $\gamma = 2$.



(a) *(b)* *(c)*

(d) $l = 1$, $\alpha = 1$ *(e) $l = 1$, $\alpha = 0.5$* *(f) $l = 1$, $\alpha = 2$*

(g) *(h)* *(i)*

(j) $l = 1$, $\alpha = 1$ *(k) $l = 1$, $\alpha = 0.5$* *(l) $l = 1$, $\alpha = 2$*

Figure 4.13: Covariance matrices and corresponding predicted functions using Rational Quadratic (RQ) covariance function for constant noise = 0.2, $\sigma^2 = 1$, and $l = 1$. The six upper paired figures ((a, d), (b, e), and (c, f)) are using the retrieved data at a_7, whereas the six lower figures ((g, j), (h, k), and (i, l)) are from the WRF-PALM data at a_5. (a, d) and (g, j) represents the "base case", with $\alpha = 1$. (b, e) and (h, k) have $\alpha = 0.5$, and finally (c, f) and (i, l) have $\alpha = 2$.

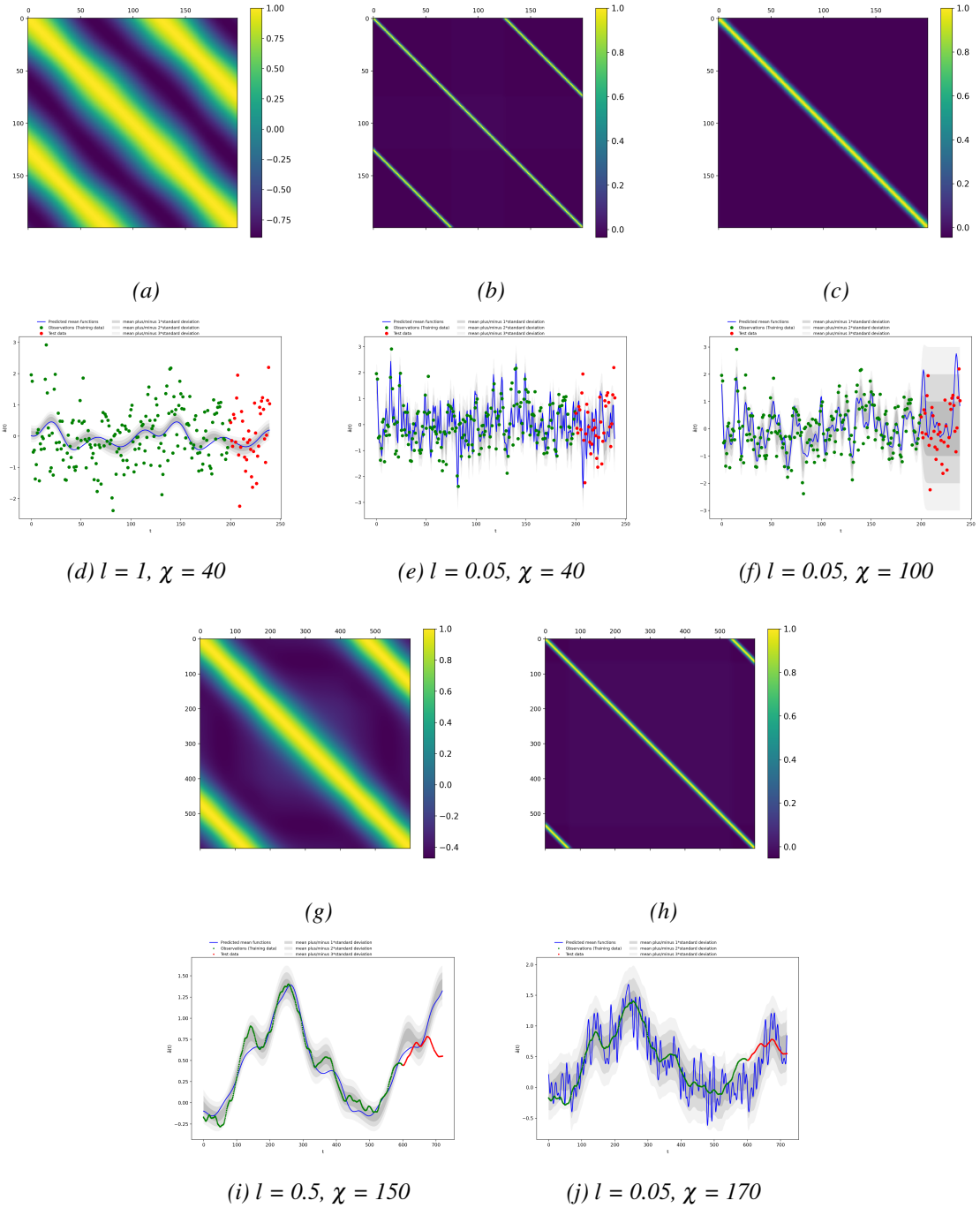


Figure 4.14: Covariance matrices and corresponding predicted functions, using the random periodic covariance function with constant noise = 0.2 and $\sigma^2 = 1$. The six upper paired figures ((a, d), (b, e), and (c, f)) are using the retrieved data at a_7, whereas the six lower figures ((g, j), (h, k), and (i, l)) are from the WRF-PALM data at a_5. (a, d) have $l = 1$, $\chi = 40$, (b, e) are with $l = 0.05$, $\chi = 40$, and finally (c, f) have $l = 0.05$, $\chi = 100$. (g, i) have $l = 0.2$ and $\chi = 250$, and (h, j) is with $l = 0.05$, $\chi = 600$.

Kernel Performance Summary

Findings from the sensitivity study in section 4.2.2 indicates that l greatly dominates the shape of the predicted function for all kernels. Increasing the length scale results in decreased fluctuations. On this basis, the length scale appears to be the dominant hyper-parameter in all kernels.

Decreasing γ for the γ-exponential kernel increases the fluctuations. The hyper-parameter α, for the RQ kernel, impacts the predicted functions minimally. We also learned that the value of χ seems to be more data-dependent than the other hyper-parameters, and greatly affects the pattern of the predicted curve. Moreover, the RQ and SE kernel appear to behave similarly, and the same goes for the γ-exponential and the exponential kernel.

In common for the kernels, we found that increasing fluctuations generally decrease the standard deviation range. Hence, generating 10 random predicted functions of the fluctuating kind would probably be more alike, than what 10 random smooth functions would have been. Moreover, the standard deviation range in the area of test data (red points) was observed to be wider than the area for training data.

All kernels appear to perform somewhat similarly on both the fluctuating retrieved data at mode 7 and the relatively smooth WRF-PALM data at mode 5. It is recognized that the observed retrieved data is outside the standard deviation ranges for the smooth curves (e.g. figure 4.10 c and 4.14 a), whereas the fluctuating curves seem to fit better. As a consequence, the predicted functions are more restricted and not that free to vary in this case. The observed WRF-PALM data is, on the other hand, always observed to be within 3 standard deviations of the predicted curve.

On a scale from 1-5, is the expected performance from visual interpretation of the various kernels for the stochastic POD presented in Table 4.4, with $k = k_i(t,t')$ at the respective mode i.

Table 4.4: Evaluation of the performance of various kernels based on visual interpretation using both retrieved and WRF-PALM data, accompanied with the respective kernel formula, where $k = k(t,t')$.

Kernel performance			
		Performance of dataset	
Kernel	**Formula**	Retrieved	WRF-PALM
SE	$k = \sigma^2 \exp\left(-\frac{(t-t')^2}{2l^2}\right)$	4	5
Exponential	$k = \sigma^2 \exp\left(-\frac{t-t'}{l}\right)$	4	2
γ-exponential	$k = \sigma^2 \exp\left(-\left(\frac{t-t'}{l}\right)^{\gamma}\right)$	4	3
RQ	$k = \sigma^2 \left(1 + \frac{(t-t')^2}{2\alpha l^2}\right)^{-\alpha}$	5	4
Periodic	$k = \sigma^2 \exp\left(-\frac{2\sin^2(\frac{t-t'}{\chi})}{l^2}\right)$	1	2

Comments to Kernel Performance. The performance of each kernel determined based on visual interpretation of the figures provided in section 4.2.2 is based on a rather subjective evaluation, and can easily be interpreted in other ways. It is therefore conceivable that the interpretation is misread. Moreover, as the grade of optimal fluctuation was unknown at the time the sensitivity study was conducted, the kernels were evaluated mainly based on flexibility.

Reconstruction using Stochastic POD

The reconstructed fields using the temporal weighting coefficients obtained from GPR are presented with the corresponding reconstructed fields using standard POD and the original snapshot. For the stochastic POD, the noise is kept at 0.2 and σ^2 at 1. The entire snapshot matrix is utilized for both data sets, at mode 5 to enable comparison. Hence, the following section assumes that the effect of the hyper-parameters identified in section 4.2.2 for a limited period, holds for the entire period time at 5 modes, for both data sets.

Since the stochastic temporal weighting coefficients rely on stochasticity, a stationary solution calculated from one single predicted curve may become unstable. In order to address this, an ensemble approach is utilized. By the ensemble approach, we mean that an average of over five reconstructed velocity fields using the temporal weighting coefficients obtained from the GPR and the truncated POD modes are used to find a stable solution [151].

Table 4.2.2 indicates that the SE and RQ kernel is expected to provide the greatest range of reconstructed fields, hence possibly fitting both data sets. Therefore, the reconstruction of the stochastic temporal weighting coefficients is calculated using the SE kernel, with the length scale $l = 0.5$.

The results are presented in figure 4.15 and 4.16 for the retrieved and WRF-PALM data, respectively. Both figures are arranged with the u and v fields on the left and right, respectively. Each subplot is separated into three snapshots, where the leftmost snapshot is the reconstruction from the standard POD, the middle from the stochastic POD, and the rightmost for the initial snapshot.

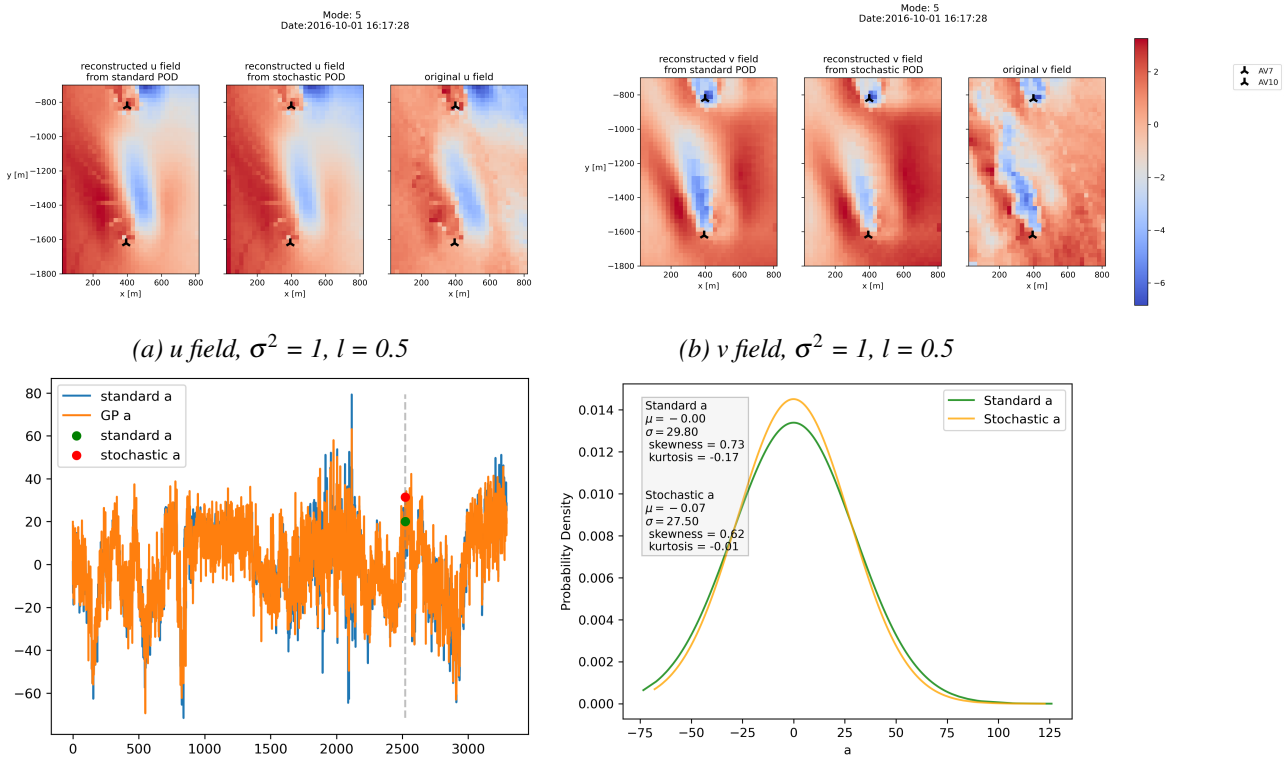
Comparing the three snapshots for the retrieved data in figure 4.15 a, we find some clear deviations in the shape of the areas marked in blue. The two reconstructions (left and middle) appear to be of a different shape than what is found for the original snapshot (right). This is perhaps due to the low number of modes used in the reconstruction. Firstly, comparing the reconstruction utilizing GP and the original snapshot, we find that the stochastic POD appears to reconstruct the blue area in the middle in a rightward skewed orientation. The blue area in the upper right corner seems to be overestimated, as for the free-flow (red). We find tendencies of the blue area to the right for turbine 10 in the original snapshot present in the reconstruction as well. By investigating the reconstructions from the standard and stochastic POD, we find that the shape of the wake is comparable. It also appears as if the features from the standard POD are reinforced. E.g. the lines in the free-flow area for the standard POD are more striking in the stochastic POD. The same goes for the wake in the upper right corner.

In figure 4.15 b, we find that the values for the stochastic POD are more extreme than what is found in the original snapshot. Hence, the blue area indicates the wake is smaller at both turbine locations, and the free-flow wind speed appears to be increased. Investigating the standard POD, we clearly see that features of the standard POD are enhanced in the stochastic POD. Figure 4.15 c indicates that the standard and stochastic temporal weighting coefficients, at the various points in time (snapshots), follow one another rather well. From the histograms in figure 4.15 d, with associated statistics, we find that the functions are very much alike.

The results for the WRF-PALM data are presented in figure 4.16. From the first glance at figure 4.16 a and b, the stochastic GP appears to be less prominent as the retrieved data. However, especially for the v field, it seems that the wake is directed towards the right for the three

uppermost turbines.

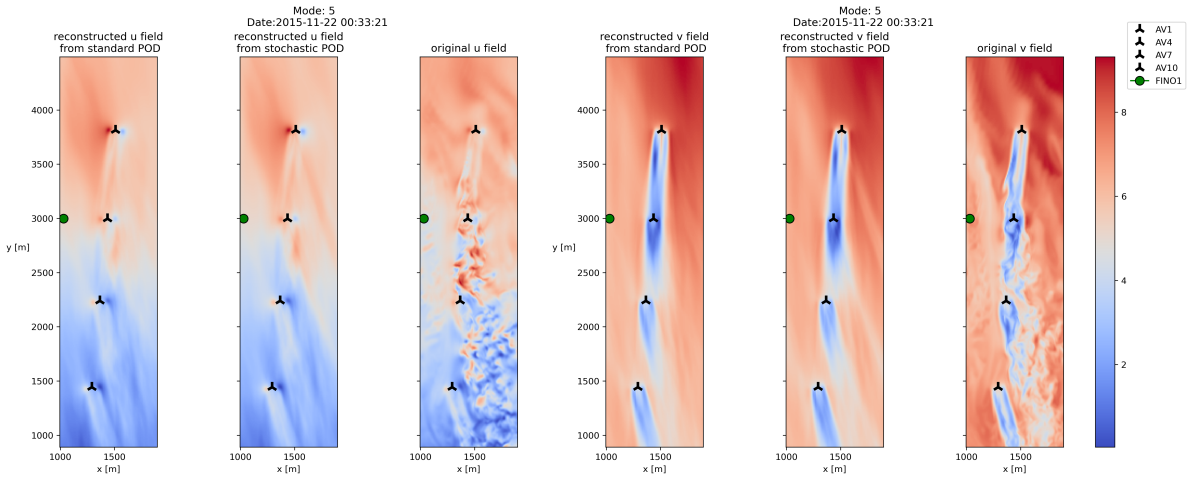
Figure 4.16 c visualizes that the standard temporal weighting coefficients are totally covered by the stochastic coefficients, indicating that fluctuations are added using the SE kernel with the parameters specified previously. Despite that, the histograms in 4.16 d indicates that the statistical distributions of the stochastic coefficients, $\tilde{a}_5$ and deterministic coefficients a_5 are comparable.



(a) u field, $\sigma^2 = 1$, $l = 0.5$ *(b) v field, $\sigma^2 = 1$, $l = 0.5$*

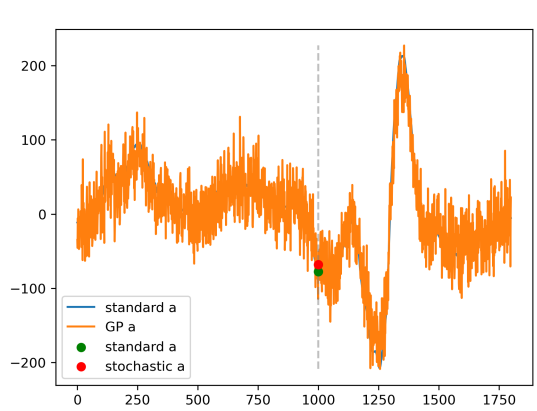
(c) a_5 and predicted GP function $\tilde{a}_5$ *(d) Histogram of a_3 and $\tilde{a}_3$*

Figure 4.15: Reconstruction of the (a) u and (b) v fields for the retrieved data using stochastic POD. The associated figures (c) and (d) show statistics of the temporal weighting coefficients, a_5, in relation to the stochastic temporal weighting coefficients, $\tilde{a}_3$. (c) the predicted GP function for $\tilde{a}_3$ and (d) the histograms.

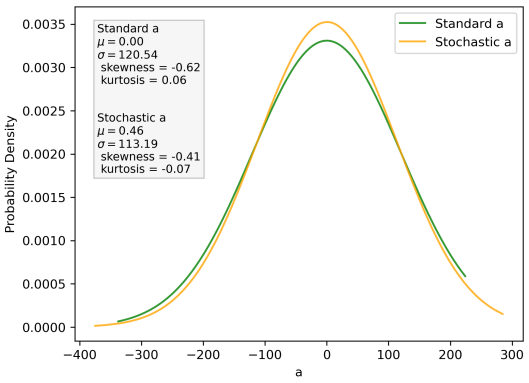


(a) u field, $\sigma^2 = 1$, $l = 0.5$

(b) v field, $\sigma^2 = 1$, $l = 0.5$



(c) Temporal weighting coefficients a_5 and stochastic temporal weighting coefficients $\tilde{a}_5$



(d) Histogram of temporal weighting coefficients a_3 and stochastic temporal weighting coefficients $\tilde{a}_3$

Figure 4.16: Reconstruction of the (a) u and (b) v fields for the WRF-PALM data using stochastic POD. The associated (c) predicted curve for the temporal weighting coefficients at mode 5, a_5 and (d) histogram of the temporal weighting coefficients, a_3 and the stochastic temporal weighting coefficients, $\tilde{a}_3$.

The respective RMSE errors are presented in Table 4.5. By comparing the numbers with the standard POD, we find clear increased relative error for the WRF-PALM data. Somewhat unexpected, the relative errors for the retrieved data are significantly lower.

Table 4.5: RMSE and cumulative eigenvalues for five POD modes, for retrieved and WRF-PALM data. RMSE is calculated for the entire snapshot, area 1, 2, and 3. For further information about the areas, see figure 4.7.

RMSE for stochastic									
		RMSE							
		u field				*v* field			
Data set	**N modes**	**Total**	**Area 1**	**Area 2**	**Area 3**	**Total**	**Area 1**	**Area 2**	**Area 3**
Retrieved	5	0.585	0.3	0.311	0.529	1.007	0.679	1.81	1.067
WRF-PALM	5	1.128	1.033	1.055	1.415	1.665	2.341	1.603	2.448

Chapter 5

Comparison and Discussion

Within the scope of this study, the results are discussed and evaluated in the context of previous work. In the following chapter, data used for validation of the retrieved data is discussed firstly. Thereafter, the processing of the retrieved data is concerned. The retrieved and WRF-PALM data are addressed next, with the focus on the differences separating them. Two types of reconstructed fields using POD have been presented in the result section; standard POD and stochastic POD. The procedures were conducted from both retrieved and WRF-PALM data, and the number of modes considered as required for an efficient reconstruction is discussed. The combination of POD and LiDAR data contributes to a more comprehensive understanding of the POD technique in general, and moreover provides new insight into POD applied to more realistic data sets such as LiDAR data, especially of higher value to the industry.

5.1 Data

Data for Validation

By separating the data presented in section 3.2 into data for POD and validation, the retrieved and WRF-PALM data are used for POD, whereas the remaining cup anemometer, wind vane, and SCADA data are used for validation. Common for the latter is that they are modestly processed, where values detected outside a fixed threshold value labeled as outliers were removed. As mentioned in the respective sections 3.2.2 and 3.2.3, various factors may impact the data sets and reduce the data's validation ability.

Based on the lack of information and inspection of the data, an assumption that the cup anemometer and wind vane are not distorted was made (as no peaks were detected), which is most likely an idealized assumption. Since this data is used for the retrieval and for data processing of the retrieved data, this could impact the results presented. Moreover, the reference wind direction is used to determine cross-wind directions, which appears to be the main contributor to the bad-quality data. Therefore, corrupted reference wind direction measurements may remove high-quality data while leaving bad-quality data.

For the SCADA data, some outliers were detected and removed, as shown in the Table 3.4. As explained in section 3.2.3, based on the sensor's location on top of the turbine, it must be assumed that the data is distorted to some extent from the rotor and/or neighboring wakes (see

figure 3.4). Therefore, the SCADA data could potentially be highly corrupted. By investigating the (reference) wind speed distribution, with the knowledge that southwest is the dominant wind direction in this period (see section 3.2.2), the effects from neighboring wakes on the SCADA data are probably moderate. The probability of disturbances caused by the rotor can, on the other hand, not be eliminated.

Another aspect to consider is the vertical and horizontal displacement of the various data sets, enabling local deviations to appear without necessarily being erroneous. The horizontal distance between the SCADA and LiDAR data (and reference data) is nearly one kilometer (919m), offering the possibility for local differences [142]. Keeping in mind that the LiDAR ranges from 23.5m to 290m .a.s.l. (see Table 3.4) and the area of interest (see figure 3.6), we can roughly estimate the LiDAR scanning range to be from 90m to 190m. Hence, the vertical difference between the LiDAR and the reference wind speed and direction, and SCADA data are in a range from 0m to 100m, roughly speaking. We know, from section 2.3.3, that the velocity of the wind varies with height. Therefore, it could be reasonable to think that height differences like these could cause deviations in the various data sets. However, as found in section 2.3, the profile is generally steep offshore, with a relatively homogeneously distributed wind velocity with height. Moreover, the differences are greatest near the surface due to surface friction. On this basis, and considering the figure 4.1, it is reasonable to assume that the differences caused by vertical displacement in the SCADA and LiDAR data are minimal.

Overall, both the wind wane, cup anemometer, and SCADA data appear to coincide relatively nicely, as visualized in figure 4.1. Thus, they are assumed to be applicable for providing a relatively correct validation of the retrieved data. However, given the potential for distortion (especially considering the SDADA data), it is somewhat unclear how the data sets behave this similar.

LiDAR Processing

In section 4.1, the processing of the LiDAR data was elaborated on. After a subdomain was chosen, and scans of data gaps were removed, the retrieval was enabled. From the analysis of the retrieved data, we find that it is within reasonable values according to the validation data. Moreover, the gaps in the retrieved data were found to be caused mainly by the cross-wind effect. It's debatable whether all of the scans removed in this range were of bad-quality. However, as no algorithm was developed to identify low and high-quality snapshots, this was determined to be a necessity to obtain the best quality data possible, as the number of low-quality snapshots detected in the region was high. By developing an algorithm for such, one could possibly reduce the number of snapshots removed to some degree.

One can speculate on whether a simple retrieval algorithm such as the one utilized in the work of Krutova M. et al. [142], would have been sufficient in this work. This could make the entire process even less complex, thus, more applicable to the industry. Moreover, the simple retrieval algorithm would not remove every 25^{th} scan.

Finally, a local peak removal and interpolation were conducted. As the local peak removal algorithm was developed based on visual inspection of the time series together with the snapshot, on a randomly selected sample of the data, the possibility that it is not applicable for other data sets, operates faulty, and/or decreases the quality in cases that were not inspected, is possible. The interpolation method used in this work was simple linear interpolation. Hence, the discussion of whether a more sophisticated peak removal and interpolation method would

have increased the quality of the retrieved data is highly relevant.

Comparison of Retrieved and WRF-PALM Data used for POD

Before investigating the results provided by the POD, it is important to bear in mind the difference in the types of data sets that are being compared in this work. First of all, from the previous section, we know that the retrieved data has undergone data processing, whereas the WRF-PALM data was provided in high-quality. Therefore, there is a probability that the processing of the LiDAR data was conducted fault, insufficiently, or excessively.

From Table 3.4, some of the major characteristics of the retrieved ad WRF-PALM data are presented. The first clear difference emerging from the table is the time period, where the retrieved data is collected for 10 days and one hour is collected for the WRF-PALM data. In addition, produces the retrieved data one snapshot every 50^{th} second, whereas one snapshot for the WRF-PALM data is produced every 2^{nd} second. Considering the periods investigated, the result is a total number of 6000 and 1800 snapshots, respectively. On a smaller scale, one hour of LiDAR data is described by initially 25 snapshots, whereas one hour of the WRF-PALM data is described by 1800 snapshots. Additionally, since the LiDAR data only measures the first 20 minutes in the hour, and not the remaining 40 minutes, the LiDAR data is measured in an irregular temporal manner, whereas the WRF-PALM data is simulated on a regular temporal resolution. On this basis, the retrieved data is expected to be relatively much more fluctuating compared with the WRF-PALM data. Moreover, and enhancing the fluctuation pattern from the retrieved data, is the fact that 45% of the retrieved data had to be removed due to bad-quality. Thus, the final number of snapshots being compared, as explained in section 3.2.2 and 3.2.1, respectively, are 3293 (initially 6000) and 1800 number of snapshots.

As a result of the high percentage of data removal for the retrieved data, scans where the turbines are turned off have not been removed (wind speed below 4m/s). Figure 4.2 reflects the data within these constraints. As the lowest mean wind speed value for the WRF-PALM data was detected as 6.16m/s, these kinds of snapshots are assumed not to be present in the WRF-PALM reconstruction. By not removing these scans from the retrieved data, we expect them to affect the energy in the POD modes somewhat, but to what extent remains unknown.

The next observation is the spatial resolution, which reports a grid resolution of 25m for the retrieved data, and 10m for the WRF-PALM data. Based on this information, we expect finer results including more small-scale wake dynamics from the WRF-PALM data. By reading further in Table 3.4, we find that the subdomain investigated for the retrieved data is very much like the WRF-PALM data in the x-direction, whereas the WRF-PALM data covers over three times the length of the retrieved data in the y-direction. Considering the x and y ranges in conjunction with the resolution, we find that the WRF-PALM area is described by 23 times more grid points than what the retrieved data is. The retrieved and WRF-PALM subdomains can be seen in figure 3.6 and 3.2, respectively. Therefore, we can conclude that this will affect the overall visual quality of the snapshots, and the number of small-scale dynamics captured. Moreover, the relatively small subdomain chosen to investigate for the retrieved data makes the visual interpretation somewhat complicated. In periods when the wake in the retrieved data is not aligned with the y-direction, the wake is not covered by the subdomain. Hence, the dynamics of the meandering are somewhat disrupted for these scans, in particular, complicating the interpretation somewhat. Since the wake effect is a key interest in this work, the subdomain is somewhat poorly chosen. On the other hand, is the performance of the POD

algorithm unaffected by the area investigated, but rather the quality of the data.

One can in general say that for the purpose of exploring wake, the WRF-PALM data is definitely most suitable for this work. However, with the aim of the work in mind, the data quality of the LiDAR is experienced and considered as of high-quality, enabling investigation of the possibility of creating an applicable method for the industry, outscoring the currently analytical models. Considering the fact that one could never obtain the same temporal and spatial resolution for real-life LiDAR measurements as what can be done for simulations, it appears as if the snapshots moved forward with, are representative.

5.2 Proper Orthogonal Decomposition

5.2.1 Standard POD

As already elaborated on, is POD a widely used technique for reducing the order of a high-dimensional system [10, 12, 27–35]. The aim of this work is to obtain the lowest rank possible, while creating a useful low order description, in order to drive down the computational costs as much as possible. Thus, a simple visual inspection is inadequate to determine the number required. Naturally, as done in [12, 29, 31, 32, 35, 36, 43, 156, 159], the eigenvalues are inspected together with the visual inspection. Moreover, as seen in [31, 33, 35, 151, 156, 157, 159], we have utilized a statistical tool, RMSE in this work, to gain a deeper understanding of the data. The results of POD applied to the retrieved and WRF-PALM data in its entirety clearly mirror the differences discussed in section 5.1.

An immediate visual inspection of figure 4.5 and 4.6 indicate that increasing the number of modes of the reconstructions enhances the accuracy of the description of the wake and wind flow in the original snapshots for the retrieved data and WRF-PALM data, respectively, as seen in e.g. [31, 33, 34, 159] as well. Thus, for a precise description of the original snapshot, a substantial amount of modes is required in the reconstruction.

Energetic Contribution (Eigenvalues) Firstly, the energy captured with the respective POD mode is investigated in section 4.2.1. We find that the energy of the POD modes for the WRF-PALM data (figure 4.4 c) decreases more rapidly than for the retrieved data (figure 4.4 a). As stated in the works [10, 28], the rate at which the energetic contribution decreases with the increasing mode is heavily dependent on the width of the range the energy in the flow is distributed over (see figure 2.2). Hence, if the energy is distributed over many modes, as seen in the retrieved data, it may indicate a wide range of turbulent structures present in the snapshots. Moreover, as found in the work of Bastine, D. et al. [31], the slow decaying variance with modes indicate that more POD modes are required to cover a large part of the total energy. On the contrary, a relatively narrow distribution of the energy, as we see for the WRF-PALM data, can be interpreted as if there are not as many turbulent structures in the data. However, from a brief look at figure 4.5, presenting a range of reconstructed wind fields of the retrieved and WRF-PALM data, we get the impression that the matter is the opposite, and that the WRF-PALM data appears to capture more scales of turbulence than what the retrieved data does. A reasonable explanation for the wide distribution of energetic contribution for the retrieved data can be due to the wide range of various snapshots (large period of time), measured during varying atmospheric stability regimes. Hence, one can assume a great variation in snapshots captured by the LiDAR. On the other hand, the snapshots obtained from the WRF-PALM

simulation for one hour, dominated by unstable conditions, may be somewhat more similar in dynamics. Hence, the nature of real-life measurements versus simulations, as discussed in the previous section, is perhaps reflected in the distribution. A final aspect to consider is that the grid points used to produce the fields are 1408 and 32400 for the retrieved and WRF-PALM data, respectively. Intuitively, one finds more dynamics for the WRF-PALM data. Therefore, one could speculate in whether the trend of the eigenvalues would have been more similar for the two data sets if the wake dynamics covered in the snapshots from the retrieved data were more similar.

The values of the eigenvalues reported in figure 4.4 are considered to be in line with the findings of [35], where the energy captured was found to be approximately 21%, 72%, and 82% for the first, 15^{th}, and 30^{th} mode respective. In comparison with the retrieved data used in this work, the values were found to be approximately 28%, 63%, and 77% for the first, seventh and 21^{st} mode, respectively.

The eigenvalues for the WRF-PALM data are in a different range, where the first mode accounts for as much as 83% of the total energetic contribution in the covariance matrix, displayed in equation (3.17). These kinds of values are found in work of S. S. Ravindran [29]. The energetic contribution was found to be approximately 97% for the first three modes and 99.69% for the first sixth modes. In this work, one and seven modes were found to account for approximately 83.3% and 93.9%.

RMSE As already mentioned, numerous works have incorporated the study of relative error for validation of POD performance. Two areas within the wake region were investigated in [33] for one specified mode using histograms and corresponding statistics. However, for the majority of the error inspections regarding POD performance, various types of errors for the entire snapshots are being compared [35, 36, 151, 156, 157, 159]. In the work of G. Dimitriu et al., [156], RMSE was used for validation. The RMSE values were somewhat larger in this study. By including 4 modes with respective cumulative energy of 91.74%, they found RMSE for regular POD to be 0.1193. In contrast, we found 89% energy coverage for 101 modes and 93% energy coverage for 5 modes, for retrieved and WRF-PALM data, respectively. The former had total RMSE calculations for the u and v fields of 0.26 and 0.48, whereas the latter had 0.53 and 0.88, respectively. However, this could be justified that both the retrieved are real-life measurements, and the WRF-PALM data is generally less idealized compared with other numerical techniques [156] as those used in the presented work.

As emerges from the studies is that the overall (total) error decreases with increasing POD number, as is shown to be the case for this study as well, apart from a relatively small increase for the retrieved data. Similar behavior as the one unexpected increase in total RMSE for the retrieved u field at mode 101, displayed in Table 4.2, has been found in the previous study of A. G. Buchan et al., [159]. As for that work, and considering the relatively large increase in energy added, the reason is unknown. However, as the visual inspection indicates that the reconstruction is improved, and the error seems to be relatively small, the error is perceived as not interfering with the overall results. Moreover, the general trend for the areas (areas 1, 2, and 3) investigated is decreasing as well. However, a larger number of alternating behavior was found here, especially for the WRF-PALM data. As no previous work investigating relative error of smaller areas is found to compare with, and the total error generally follows the patterns from previous work, it is regarded as acceptable in this work. However, due to this uncertainty, the RMSE is investigated together with the total RMSE and visual inspection in order to de-

termine the data's credibility. Some possible explanations are small temporal displacements in the dynamic features of the wake. The increased observations for the WRF-PALM data are most likely due to narrow energy distribution over POD modes.

Hence, from inspection of previous work [10, 29, 31, 33, 35, 45], it appears as the eigenvalues obtained from POD of both the retrieved and WRF-PALM data are in a realistic range, and reflect the variety of data included in the data sets. As emerges from the results provided by previous studies, it appears to be highly data-dependent.

5.2.2 Stochastic POD

Comparison. From the results obtained by utilizing the stochastic POD, we find one case for the u and v field for both data sets investigated, accompanied by one time series and one histogram at 5 and 3 modes, respectively. Comparing the histograms, they both offer information that the statistics models generally preserve the statistical properties of the standard POD. The largest deviations are found for the WRF-PALM data, and only kurtosis deviates more for the retrieved data. From inspection of the predicted functions (figure 4.15 c and 4.16 c), we find that there is included significantly more fluctuations into the stochastic temporal weighting coefficients for the WRF-PALM data, than what is the case for the retrieved data. This is because the same kernel and parameters are used for both data sets. From the sensitivity study conducted in section 4.2.2, we found that the retrieved data fluctuates greatly, whereas the WRF-PALM data has smooth characteristics. Hence, the predicted functions created for the SE kernel are not providing any surprising aspects. It is, therefore, somewhat suspect that changing the standard with the stochastic temporal weighting coefficients for the WRF-PALM data appears to have a smaller effect than for the retrieved data. This may be, again, due to the different spatial resolutions.

As the results from the stochastic POD are not striking as improving the results from the standard POD greatly, one could debate whether this is due to the fact that the number of modes suitable for a reconstruction were found to be in the range from 50 to 101 modes for both data sets. However, by comparing the RMSE calculations for the stochastic POD found in Table 4.5 and the standard pod (from Table 4.2 and 4.3), we find that the total errors have been increased for both data sets. For the retrieved data, the total RMSE is moderate for both the u and v fields. We notice, from inspection of the u field, that the error for area 1 and 3 have increased, whereas for area 2 it has decreased. However, no clear improvement emerges from a visual inspection of area 2. Hence, a suggested explanation is that the error decrease is mainly caused by the decreased values for the blue area detected for the stochastic POD. Hence, as a result from that the blue area is somewhat skewed for the stochastic POD, the error in this area become smaller as well. The same assumptions for the v field.

For the WRF-PALM data, on the other hand, the errors are greatly increased for both wind fields in all areas. As the stochastic fields were not visually interpreted as improved, this was expected. However, an interesting observation is that the RMSE deviations from the standard and stochastic WRF-PALM reconstructions are found to be relatively much greater than what was the case for the retrieved data. Considering that the visual changes seemed to be larger for the retrieved data. On the other hand, from the statistics provided by the histogram and predicted functions, the difference between the standards and stochastic temporal weighting coefficients was clearly found to be the largest for the WRF-PALM data.

Gaussian Process (GP). In this work, GPR is found to decrease in performance of the standard POD, as the deterministic weights were changed with the stochastic weights, both statistically and visually. The results are not in line with previous studies such as [33]. However, keeping in mind the large range of possible kernels indicated by the sensitivity study, this is possibly explained by the that only one kernel was used in this work.

Despite that a wide range of kernels can be created, it can be debated whether other methods perhaps are more appropriate. Three methods were investigated in [10, 33], and two, including GP, are found in [33]. Moreover, more sophisticated methods have been investigated, such as those proposed by [160, 161]. These methods account for non-linear moments of the temporal weighting coefficients. In common for these methods is that they require relatively much data, which in turn is not applicable for efficient reconstructions. In addition to these methods, is the Galerkin approach well-documented approach to including non-linearities into the system, e.g. in [10, 26, 29, 35, 152].

Moreover, GPR has been used extensively in previous studies [27, 30] in combination with the data-driven ROM. One debate addressed in [39], is that the prior function is supposed to be determined by priory beliefs of the function. It is however determined most frequently based on mathematical convenience.

The GP is, as seen throughout this study, very adaptable. There is both a positive and negative aspect to this. It is beneficial due to its ability to costume the kernels for the specific system investigated. However, as seen in this work, finding the optimal kernel is challenging and time-consuming. The GP is not sparse and uses the entire time series given to it. In this work, as the processes must be calculated several times for the number of ensembles used, the computational time becomes unpractical and disables the efficient estimations applicable to the industry.

A final aspect to keep in mind for the stochastic reconstruction is that the decomposition of the wind field consists of two components. In this work, we only address the temporal component of the total system. Therefore, the spatial components remain untreated.

Analytical models

From a brief comparison between the analytical models seen in section 2.4.1 and the reconstructed wind fields visualized in figure 4.5, we get the impression that the POD performs better when it comes to e.g. wake meandering, velocity recovery distance, and wake interactions with shear and turbulence in the atmospheric boundary layer are undoubtedly more present. By a brief inspection of the centerlines, we find that the analytical models are completely symmetrical. The reconstructed fields, on the other hand, we find several details added in the wake dynamics, even at a few modes. However, as no validation study has been conducted, the improvements are still unknown. Moreover, despite that the POD is a relatively simple procedure for conducting wake estimations, it is not inevitable that the data input for the analytical models is far easier to get a hold of. Another aspect to consider is that LiDAR scans are limited to the range of the LiDAR. Thus, an entire wind farm cannot be investigated with this procedure, whereas this is possible with the analytical models.

Chapter 6

Conclusions and Future Work

The capabilities of POD have been investigated by using the technique on both LiDAR measurements obtained from FINO1 during varying atmospheric conditions, and WRF-PALM data simulated for one hour during unstable conditions. By utilizing standard POD, we have decomposed the wind fields into temporal weighting coefficients and spatial POD modes. The novelty of this work shows that POD is able to reconstruct wind flows captured by a LiDAR in a reduced manner while preserving the large-scale structures of the wake flow. Additionally, is the method shown to be efficient for the standard POD.

In order to incorporate stochasticity to the reconstructed wind fields, the deterministic temporal components were replaced by predicted coefficients using the GPR. A sensitivity study further substantiates the method's dependence on kernel selection, with associated hyper-parameters, for POD application. Moreover, the reconstruction of stochastic coefficients was showed to be highly dependent on the original POD, as it appears that the GPR emphasizes patterns found in the reconstruction of the standard POD. It emerges from the study that stochastic coefficients may include dynamics that are not present using the deterministic temporal weights. In this work, one single kernel was investigated, which decreased the visual performance and increased the relative error.

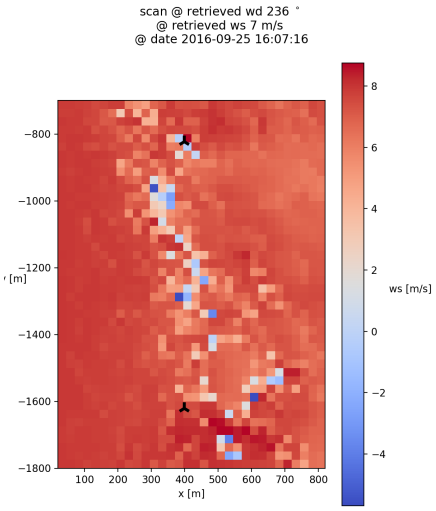
Hence, a more extensive investigation must be done in order to determine the ability of the stochastic POD for real data such as LiDAR measurements. Despite the different data sets investigated in the work, the results for both standard and GPR POD on LiDAR are considered to be in line with the results obtained for the WRF-PALM data. Furthermore, it appears as the stochastic POD potentially is applicable for LiDAR data.

Future work includes further investigation of stochastic models for LiDAR data and extending to explore the spatial stochastic modes. A suggestion is to separate the LiDAR snapshots based on stability, which was an initial thought for this study. Moreover, great insight is gained by utilizing the reconstruction fields in power and load calculations.

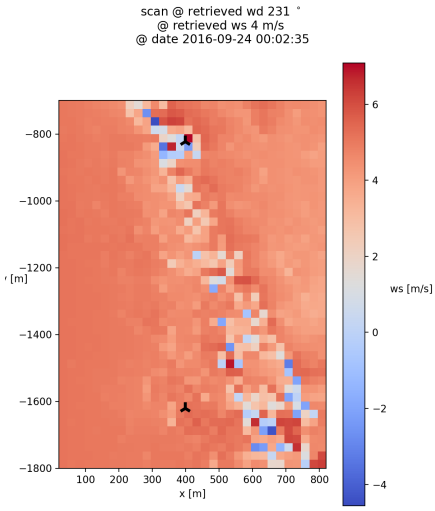
List of Figures

List of Tables

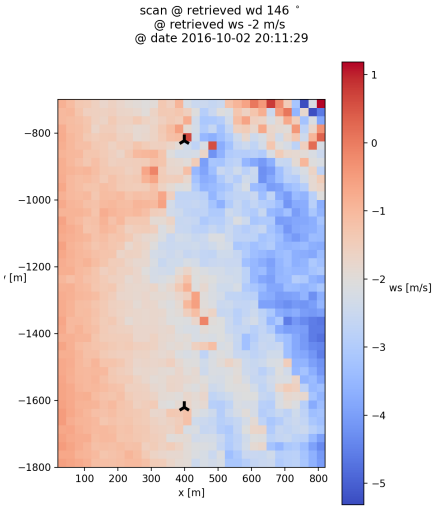
Additional Figures



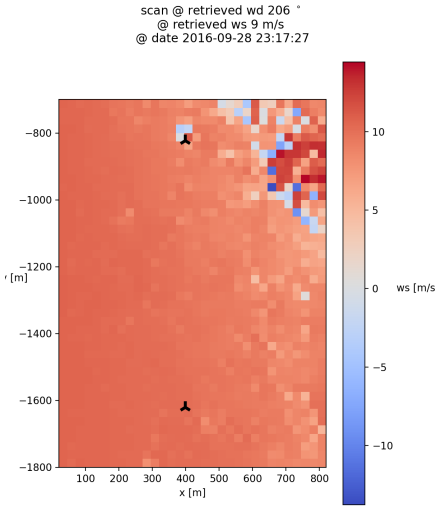
(a) u field



(b) v field



(c) u field



(d) v field

Figure 1: LiDAR snapshots in the cross wind range.

LIST OF TABLES

Bibliography

[1] U. N. C. Change, "Paris Agreement," Tech. Rep., 2015. [Online]. Available: https://unfccc.int/sites/default/files/english_paris_agreement.pdf

[2] I. E. Agency, "Net Zero by 2050 - A Roadmap for the Global Energy Sector," Tech. Rep. 4, oct 2021. [Online]. Available: https://www.iea.org/reports/net-zero-by-2050

[3] United Nations, "For a livable climate: Net-zero commitments must be backed by credible action." [Online]. Available: https://www.un.org/en/climatechange/net-zero-coalition

[4] M. Jaganmohan, "Global installed wind energy capacity from 2001 to 2020," apr 2021. [Online]. Available: https://www.statista.com/statistics/268363/installed-wind-power-capacity-worldwide/

[5] IEA, "Wind Power," 2021. [Online]. Available: https://www.iea.org/reports/wind-power

[6] Z. Ti, X. W. Deng, and H. Yang, "Wake modeling of wind turbines using machine learning," *Applied Energy*, vol. 257, no. July 2019, p. 114025, 2020. [Online]. Available: https://doi.org/10.1016/j.apenergy.2019.114025

[7] P. McKay, R. Carriveau, D. S.-K. Ting, and T. Newson, *Turbine Wake Dynamics*, R. Carriveau, Ed. London, United Kingdom: IntechOpen, 2012 [Online], nov 2012. [Online]. Available: http://dx.doi.org/10.5772/53968

[8] T. Dasari, Y. Wu, Y. Liu, and J. Hong, "Near-wake behaviour of a utility-scale wind turbine," *Journal of Fluid Mechanics*, vol. 859, pp. 204–246, jan 2019. [Online]. Available: https://doi.org/10.1017/jfm.2018.779

[9] T. Göçmen, P. V. D. Laan, P. E. Réthoré, A. P. Diaz, G. C. Larsen, and S. Ott, "Wind turbine wake models developed at the technical university of Denmark: A review," *Renewable and Sustainable Energy Reviews*, vol. 60, pp. 752–769, jul 2016.

[10] D. Bastine, L. Vollmer, M. Wächter, and J. Peinke, "Stochastic wake modelling based on POD analysis," *Energies*, vol. 11, no. 3, pp. 1–29, feb 2018. [Online]. Available: https://doi.org/10.3390/en11030612

[11] M. Krutova, M. B. Paskyabi, F. G. Nielsen, and J. Reuder, "Evaluation of Gaussian wake models under different atmospheric stability conditions: Comparison with large

eddy simulation results," *Journal of Physics: Conference Series*, vol. 1669, no. 1, pp. 0–15, 2020.

[12] M. Zhang and R. J. Stevens, "Characterizing the Coherent Structures Within and Above Large Wind Farms," *Boundary-Layer Meteorology*, vol. 174, no. 1, pp. 61–80, 2020.

[13] M. Yeghikian, A. Ahmadi, R. Dashti, F. Esmaeilion, A. Mahmoudan, S. Hoseinzadeh, and D. A. Garcia, "Wind Farm Layout Optimization with Different Hub Heights in Manjil Wind Farm Using Particle Swarm Optimization," *Applied Sciences 2021*, vol. 11, no. 20, pp. 1–19, oct 2021. [Online]. Available: https://doi.org/10.3390/app11209746

[14] R. J. Barthelmie, S. T. Frandsen, O. Rathmann, K. Hansen, E. S. Politis, D. Cabezon, K. Rados, S. P. van der Pijl, J. M. Prospathopoulos, J. G. Schepers, W. Schlez, J. Phillips, and A. Neubert, "Flow and wakes in large wind farms in complex terrain and offshore," Tech. Rep., mar 2014. [Online]. Available: https://www.researchgate.net/publication/253382422_Flow_and_wakes_in_large_wind_farms_in_complex_terrain_and_offshore

[15] R. J. Barthelmie, K. Hansen, S. T. Frandsen, O. Rathmann, J. G. Schepers, W. Schlez, J. Phillips, K. Rados, A. Zervos, E. S. Politis, and P. K. Chaviaropoulos, "Modelling and measuring flow and wind turbine wakes in large wind farms offshore," *Wind Energy*, vol. 12, no. 5, pp. 431–444, jul 2009. [Online]. Available: https://onlinelibrary.wiley.com/doi/epdf/10.1002/we.348

[16] J. Schmidt and B. Stoevesandt, "Wind Farm Layout Optimization with Wakes from Fluid Dynamics Simulations," *EWEA 2014*, pp. 1–11, mar 2014. [Online]. Available: https://doi.org/10.13140/2.1.2544.3847

[17] M. M. Pedersen and G. C. Larsen, "Integrated wind farm layout and control optimization," *Wind Energy Science*, vol. 5, no. 4, pp. 1551–1566, nov 2020.

[18] R. Krishnamurthy, "Wind Farm Characterization and Control using Coherent Doppler Lidar," ARIZONA STATE UNIVERSITY, Tech. Rep., may 2013. [Online]. Available: https://keep.lib.asu.edu/items/151874

[19] K. E. Johnson, L. Y. Pao, M. J. Balas, and J. F. Lee, "Control of Variable-Speed Wind Turbines: Standard and Adaptive Techniques for Maximizing Energy Capture," *IEEE Control Systems*, vol. 26, no. 3, pp. 70–81, jun 2006. [Online]. Available: https://doi.org/10.1109/MCS.2006.1636311

[20] F. Porté-Agel, M. Bastankhah, and S. Shamsoddin, "Wind-Turbine and Wind-Farm Flows: A Review," *Boundary-Layer Meteorology*, vol. 174, pp. 1–59, sep 2019. [Online]. Available: https://doi.org/10.1007/s10546-019-00473-0

[21] A. Sebastiani, F. Castellani, G. Crasto, and A. Segalini, "Data analysis and simulation of the Lillgrund wind farm," Institute of Technology, Stockholm, Sweden, Tech. Rep. 6, jun 2021. [Online]. Available: https://onlinelibrary.wiley.com/doi/epdf/10.1002/we.2594

[22] N. O. Jensen, "A note on wind generator interaction," Risø National Laboratory, DK-4000 Roskilde, Denmark, Tech. Rep., sep 1983. [Online]. Available: https://backend.orbit.dtu.dk/ws/portalfiles/portal/55857682/ris{_}m{_}2411.pdf

[23] M. Bastankhah and F. Porté-Agel, "A new analytical model for wind-turbine wakes," *Renewable Energy*, vol. 70, pp. 116–123, oct 2014.

[24] P. Kundu, I. Cohen, and D. Dowling, "Computational Fluid Dynamics," in *Fluid Mechanics*, 5th ed., P. Kundu, I. Cohen, and D. Dowling, Eds. Elsevier Inc., jul 2012, ch. 10, pp. 421–472. [Online]. Available: https://www.sciencedirect.com/book/9780123821003/fluid-mechanics{#}book-info

[25] K. G. Rados, J. M. Prospathopoulos, E. S. Politis, P. Chaviaropoulos, and A. Zervos, "CFD modeling issues of wind turbine wake under stable atmospheric conditions," in *European Wind Energy Conference and Exhibition*, p. 9.

[26] D. Chapelle, A. Gariah, and J. Sainte-Marie, "Galerkin approximation with proper orthogonal decomposition : new error estimates and illustrative examples," *ESAIM: Mathematical Modelling and Numerical Analysis*, vol. 46, no. 4, pp. 731–757, jul 2012. [Online]. Available: https://www.esaim-m2an.org/articles/m2an/abs/2012/04/m2an110053/m2an110053.html

[27] M. Guo and J. S. Hesthaven, "Reduced order modeling for nonlinear structural analysis using Gaussian process regression," *Computer Methods in Applied Mechanics and Engineering*, vol. 341, pp. 807–826, nov 2018. [Online]. Available: https://doi.org/10.1016/j.cma.2018.07.017

[28] S. J. Andersen, J. N. Sørensen, and R. Mikkelsen, "Reduced order model of the inherent turbulence of wind turbine wakes inside an infinitely long row of turbines," *Journal of Physics: Conference Series*, vol. 555, no. 1, pp. 1–11, dec 2014. [Online]. Available: https://doi.org/10.1088/1742-6596/555/1/012005

[29] S. S. Ravindran, "A reducedorder approach for optimal control of fluids using proper orthogonal decomposition," *International Journal for Numerical Methods in Fluids*, pp. 425–448, nov 2000. [Online]. Available: https://doi.org/10.1002/1097-0363(20001115)34:5<425::AID-FLD67>3.0.CO;2-W

[30] M. Guo and J. S. Hesthaven, "Data-driven reduced order modeling for time-dependent problems," *Computer Methods in Applied Mechanics and Engineering*, vol. 345, pp. 75–99, mar 2019.

[31] D. Bastine, B. Witha, M. Wächter, and J. Peinke, "POD Analysis of a Wind Turbine Wake in a Turbulent Atmospheric Boundary Layer," *Journal of Physics: Conference Series*, vol. 524, pp. 1–11, 2014. [Online]. Available: https://iopscience.iop.org/article/10.1088/1742-6596/524/1/012153

[32] J. Weiss, "A Tutorial on the Proper Orthogonal Decomposition." American Institute of Aeronautics and Astronautics, 2019, pp. 17–21. [Online]. Available: https://depositonce.tu-berlin.de/handle/11303/9456

[33] M. B. Paskyabi, M. Krutova, F. G. Nielsen, J. Reuder, and O. E. Guernaoui, "On Stochastic Reduced-Order and LES-based Models of Offshore Wind Turbine Wakes," *Journal of Physics: Conference Series*, vol. 1669, no. 1, 2020.

[34] Y. D. Lang, A. Malacina, L. T. Biegler, S. Munteanu, J. I. Madsen, and S. E. Zitney,

"Reduced order model based on principal component analysis for process simulation and optimization," *Energy and Fuels*, vol. 23, no. 3, pp. 1695–1706, mar 2009.

[35] M. S. Siddiqui, S. T. M. Latif, M. Saeed, M. Rahman, A. W. Badar, and S. M. Hasan, "Reduced order model of offshore wind turbine wake by proper orthogonal decomposition," *International Journal of Heat and Fluid Flow*, vol. 82, no. January, p. 108554, 2020. [Online]. Available: https://doi.org/10.1016/j.ijheatfluidflow.2020.108554

[36] D. Bastine, B. Witha, M. Wächter, and J. Peinke, "Towards a simplified dynamic wake model using POD analysis," *Energies*, vol. 8, no. 2, pp. 895–920, jan 2015.

[37] C. E. Rasmussen and C. K. I. Williams, "Gaussian Processes for Machine Learning," Tech. Rep. [Online]. Available: www.GaussianProcess.org/gpml

[38] G. Valerio Iungo, R. Maulik, S. Ashwin Renganathan, and S. Letizia, "Machine-learning identification of the variability of mean velocity and turbulence intensity for wakes generated by onshore wind turbines," *Journal Of Fluid Mechanics*, nov 2021. [Online]. Available: https://arxiv.org/pdf/2109.01646.pdf

[39] C. Bishop, *Pattern Recognition and Machine Learning*. Springer, January 2006. [Online]. Available: https://www.microsoft.com/en-us/research/publication/pattern-recognition-machine-learning/

[40] G. Ortali, N. Demo, G. Rozza, G. Ortali, N. Demo, and G. Rozza, "A Gaussian Process Regression approach within a data-driven POD framework for engineering problems in fluid dynamics," *Mathematics in Engineering 2022 3:1*, vol. 4, no. 3, pp. 1–16, 2022. [Online]. Available: http://www.aimspress.com/article/doi/10.3934/mine.2022021

[41] G. Berkooz, P. Holmes, and J. L. Lumley, "The proper orthogonal decomposition in the analysis of turbulent flows," *Annual Review of Fluid Mechanics*, vol. 25, no. 1, pp. 539–575, 1993. [Online]. Available: https://doi.org/10.1146/annurev.fl.25.010193.002543

[42] S. J. Andersen, J. N. Sørensen, and R. Mikkelsen, "Simulation of the inherent turbulence and wake interaction inside an infinitely long row of wind turbines," *Journal of Turbulence*, vol. 14, no. 4, pp. 1–24, jun 2013. [Online]. Available: http://dx.doi.org/10.1080/14685248.2013.796085

[43] M. Debnath, C. Santoni, S. Leonardi, and G. V. Iungo, "Towards reduced order modelling for predicting the dynamics of coherent vorticity structures within wind turbine wakes," *Phil. Trans. R. Soc. A*, vol. 375, no. 2091, pp. 1–14, mar 2017. [Online]. Available: https://doi.org/10.1098/rsta.2016.0108

[44] G. V. Iungo, C. Santoni-Ortiz, M. Abkar, F. Porté-Agel, M. A. Rotea, and S. Leonardi, "Data-driven Reduced Order Model for prediction of wind turbine wakes," *Journal of Physics: Conference Series*, vol. 625, no. 1, 2015.

[45] A. Chatterjee, "An introduction to the proper orthogonal decomposition," vol. 78, no. 7, 2000.

[46] N. W. Cherukuru, R. Calhoun, R. Krishnamurthy, S. Benny, J. Reuder, and M. Flügge,

"2D VAR single Doppler lidar vector retrieval and its application in offshore wind energy," *Energy Procedia*, vol. 137, pp. 497–504, oct 2017.

[47] P. H. Jensen, T. Chaviaropolos, and A. Natarajan, "LCOE reduction for the next generation offshore wind turbines," *INNWIND.EU project*, pp. 1–104, oct 2017.

[48] M. Shields, P. Beiter, J. Nunemaker, A. Cooperman, and P. Duffy, "Impacts of turbine and plant upsizing on the levelized cost of energy for offshore wind," *Applied Energy*, vol. 298, pp. 1–13, june 2021. [Online]. Available: https://doi.org/10.1016/j.apenergy.2021.117189

[49] G. Sieros, P. Chaviaropoulos, J. D. Sørensen, B. H. Bulder, and P. Jamieson, "Upscaling wind turbines: theoretical and practical aspects and their impact on the cost of energy," *Wind Energy*, pp. 3–17, jan 2012. [Online]. Available: https://www.researchgate.net/publication/258724726_Upscaling_wind_turbines_Theoretical_and_practical_aspects_and_their_impact_on_the_cost_of_energy

[50] I. p. o. C. Change(ipcc), "Climate Change 2022 - Impacts, Adaptation and Vulnerability," ipcc, Tech. Rep., 2022.

[51] I. International Renewable Energy Agency, "Future of wind: Deployment, investment, technology, grid integration and socio-economic aspects (A Global Energy Transformation paper),," Tech. Rep., oct 2019. [Online]. Available: www.irena.org/publications.

[52] V. Smil, *Power Density : A Key to Understanding Energy Sources and Uses*, 1st ed., V. Smil, Ed. MIT Press, 2015. [Online]. Available: https://doi.org/10.7551/mitpress/10046.001.0001

[53] B. E. Layton, "A COMPARISON OF ENERGY DENSITIES OF PREVALENT ENERGY SOURCES IN UNITS OF JOULES PER CUBIC METER," *International Journal of Green Energy*, vol. 5, pp. 438–455, 2008.

[54] H. Ritchie and M. Roser, "Energy," 2020, https://ourworldindata.org/energy.

[55] T. M. T. M. Letcher, *Wind energy engineering : a handbook for onshore and offshore wind turbines*, 6th ed., Elsevier Science & Technology, Ed. London, England: Academic Press, may 2017. [Online]. Available: https://ebookcentral.proquest.com/lib/bergen-ebooks/detail.action?docID=4858115.

[56] D. A. Spera, *Wind Turbine Technology: Fundamental Concepts in Wind Turbine Engineering, Second Edition*, 2nd ed., D. A. Spera, Ed. Three Park Avenue, New York, NY 10016, USA: ASME Press, jun 2009. [Online]. Available: https://doi.org/10.1115/1.802601

[57] T. Burton, D. Sharpe, N. Jenkins, and E. Bossany, *Wind Energy Handbook*. Baffins Lane, Chichester West Sussex, PO19 1UD, England: JOHN WILEY & SONS, LTD, 2001.

[58] J. Annoni and P. Seiler, "A low-order model for wind farm control," *Proceedings of the American Control Conference*, vol. 2015-July, no. 1, pp. 1721–1727, 2015.

[59] M. Dörenkämper, J. Tambke, G. Steinfeld, D. Heinemann, and M. Kühn, "Atmospheric Impacts on Power Curves of Multi-Megawatt Offshore Wind Turbines," *Journal of Physics: Conference Series*, vol. 555, no. 1, pp. 1–11, dec 2014. [Online]. Available: https://iopscience.iop.org/article/10.1088/1742-6596/555/1/012029https://iopscience.iop.org/article/10.1088/1742-6596/555/1/012029/meta

[60] R. Barthelmie, M. Churchfield, P. Moriarty, J. Lundquist, G. Oxley, S. Hahn, and S. Pryor, "The role of atmospheric stability/turbulence on wakes at the Egmond aan Zee offshore wind farm," *Journal of Physics: Conference Series*, vol. 625, pp. 1–11, jun 2015. [Online]. Available: https://iopscience.iop.org/article/10.1088/1742-6596/625/1/012002

[61] J. Blom and L. Wartena, "The Influence of Changes in Surface Roughness on the Development of the Turbulent Boundary Layer in the Lower Layers of the Atmosphere in: Journal of the Atmospheric Sciences Volume 26 Issue 2 (1969)," *Journal of the Atmospheric Sciences*, vol. 26, no. 2, pp. 255–265, mar 1969. [Online]. Available: https://doi.org/10.1175/1520-0469(1969)026{%}3C0255:TIOCIS{%}3E2.0.CO;2

[62] P. Davidson, *Turbulence: An Introduction for Scientists and Engineers*, 2nd ed., P. Davidson, Ed. Oxford University Press, aug 2015. [Online]. Available: https://oxford.universitypressscholarship.com/view/10.1093/acprof:oso/9780198722588.001.0001/acprof-9780198722588

[63] S. Lee, M. Churchfield, P. Moriarty, J. Jonkman, and J. Michalakes, "Atmospheric and Wake Turbulence Impacts on Wind Turbine Fatigue Loadings," National Renewable Energy Lab. (NREL), Golden, CO (United States), United States, Tech. Rep., jan 2012. [Online]. Available: https://doi.org/10.2514/6.2012-540

[64] J. C. Y. Lee and M. J. Fields, "An overview of wind-energy-production prediction bias, losses, and uncertainties," *Wind Energy Science*, vol. 6, no. 2, pp. 311–365, mar 2021. [Online]. Available: https://doi.org/10.5194/wes-6-311-2021

[65] P. Veers, K. Dykes, E. Lantz, S. Barth, C. L. Bottasso, O. Carlson, A. Clifton, J. Green, P. Green, H. Holttinen, D. Laird, V. Lehtomäki, J. K. Lundquist, J. Manwell, M. Marquis, C. Meneveau, P. Moriarty, X. Munduate, M. Muskulus, J. Naughton, L. Pao, J. Paquette, J. Peinke, A. Robertson, J. S. Rodrigo, A. M. Sempreviva, J. C. Smith, A. Tuohy, and R. Wiser, "Grand challenges in the science of wind energy," *Science*, vol. 366, no. 6464, 2019.

[66] A. Crespo and J. Hernández, "Turbulence characteristics in wind-turbine wakes," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 61, no. 1, pp. 71–85, jun 1996.

[67] D.C. Quarton and J.F. Ainslie, "Turbulence in Wind Turbine Wakes," *European Wind Energy Conferance EWEC 89*, vol. Vol.14, no. No.1, pp. 15–23, jul 1989. [Online]. Available: https://www.jstor.org/stable/43749409?seq=1{#}metadata{_}info{_}tab{_}contents

[68] S. Emeis, *Wind energy meteorology :Atmospheric physics for wind power generation*, 1st ed. Springer Science & Business Media,, jul 2013. [Online]. Available: https://doi.org/10.1007/978-3-642-30523-8

[69] L. J. Vermeer, J. N. Sørensen, and A. Crespo, "Wind turbine wake aerodynamics," *Progress in Aerospace Sciences*, vol. 39, no. 6-7, pp. 467–510, aug 2003. [Online]. Available: https://doi.org/10.1016/S0376-0421(03)00078-2

[70] B. Sanderse, "Aerodynamics of wind turbine wakes Literature review," ECN Wind Energy, Tech. Rep., apr 2009. [Online]. Available: https://www.researchgate.net/publication/24272650

[71] D. Elliott, "Status of wake and array loss research," Pacific NorthwestLabor_tory, Richland, Washington 99352, Tech. Rep., oct 1991. [Online]. Available: https://www.osti.gov/servlets/purl/6211976

[72] A. Crespo and J. Hernaâ Ndez, "Survey of Modelling Methods for Wind Turbine Wakes and Wind Farms," *Wind Energy*, pp. 1–24, aug 1999. [Online]. Available: https://doi.org/10.1002/(SICI)1099-1824(199901/03)2:1{%}3C1::AID-WE16{%}3E3.0.CO;2-7

[73] J. Wallace and P. Hobbs, *Atmospheric Science: An Introductory Survey*, ser. International geophysics series. Elsevier Academic Press, 2006. [Online]. Available: https://books.google.no/books?id=k4shngEACAAJ

[74] J. F. Ainslie, "Calculating the flowfield in the wake of wind turbines," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 27, no. 1-3, pp. 213–224, jan 1988.

[75] P. Frohboese and G. Hassan, "Thrust coefficients used for estimation of wake effects for fatigue load calculation," in *European Wind Energy Conference 2010, Warsaw, Poland*, P. Frohboese, G. Hassan, and C. Schmuck, Eds. Warsaw, Poland: WINDTEST Kaiser-Wilhelm-Koog GmbH, Brooktorkai 18, 20457 Hamburg, Germany, 2010, pp. 1–10. [Online]. Available: https://www.researchgate.net/publication/267564457_Thrust_coefficients_used_for_estimation_of_wake_effects_for_fatigue_load_calculation

[76] F. G. Schmitt, "Turbulence from 1870 to 1920: The birth of a noun and of a concept," *Comptes Rendus Mécanique*, vol. 345, no. 9, pp. 620–626, sep 2017. [Online]. Available: https://doi.org/10.1016/j.crme.2017.06.003

[77] P. Bradshaw, *Turbulence*, 1st ed., ser. Topics in Applied Physics, P. Bradshaw, Ed. Berlin, Heidelberg: Springer, 1976, vol. 12. [Online]. Available: http://link.springer.com/10.1007/978-3-662-22568-4

[78] R. B. Stull, *An introduction to boundary layer meteorology*, 1st ed., R. B. Stull, Ed. VancouverCanada: Atmospheric Sciences Library, 1988, vol. 13. [Online]. Available: https://doi.org/10.1007/978-94-009-3027-8

[79] J. Å. Dahlberg, M. Poppen, and S. E. Thor, "Load/fatigue effects on a wind turbine generator in a wind farm," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 39, no. 1-3, pp. 199–209, jan 1992. [Online]. Available: https://www.sciencedirect.com/science/article/pii/016761059290546M?via{%}3Dihub

[80] L. Zhan, S. Letizia, and G. Valerio Iungo, "LiDAR measurements for an onshore wind farm: Wake variability for different incoming wind speeds and atmospheric stability regimes," *Wind Energy*, vol. 23, no. 3, pp. 501–527, mar 2020.

[81] Y. Jooss, L. Li, T. Bracchi, and R. J. Hearst, "Spatial develop-

ment of a turbulent boundary layer subjected to freestream turbulence," *Journal of Fluid Mechanics*, vol. 911, 2021. [Online]. Available: https://www.cambridge.org/core/journals/journal-of-fluid-mechanics/article/spatial-development-of-a-turbulent-boundary-layer-subjected-to-freestream-turbulence/ABD1B65685A0779E8407FD3B09607054

[82] E. Cheynet, M. Flügge, J. Reuder, J. B. Jakobsen, Y. Heggelund, B. Svardal, P. Saavedra Garfias, C. Obhrai, N. Daniotti, J. Berge, C. Duscha, N. Wildmann, I. H. Onarheim, and M. Godvik, "The COTUR project: Remote sensing of offshore turbulence for wind energy application," *Atmospheric Measurement Techniques*, vol. 14, no. 9, pp. 6137–6157, sep 2021.

[83] M. Kozul, R. J. Hearst, J. P. Monty, B. Ganapathisubramani, and D. Chung, "Response of the temporal turbulent boundary layer to decaying free-stream turbulence," *Journal of Fluid Mechanics*, vol. 896, p. 11, 2020. [Online]. Available: https://www.cambridge.org/core/journals/journal-of-fluid-mechanics/article/response-of-the-temporal-turbulent-boundary-layer-to-decaying-freestream-turbulence/28A048E2E5A0B4259B7A18F563749834

[84] P. Davidson, *Turbulence: An Introduction for Scientists and Engineers*, 2nd ed., P. Davidson, Ed. Oxford University Press, aug 2015. [Online]. Available: https://oxford.universitypressscholarship.com/view/10.1093/acprof:oso/9780198722588.001.0001/acprof-9780198722588

[85] J. Wallace and P. Hobbs, *Atmospheric Science: An Introductory Survey*, ser. International geophysics series. Elsevier Academic Press, 2006. [Online]. Available: https://books.google.no/books?id=k4shngEACAAJ

[86] A. N. Kolmogorov, "The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers," *Mathematical and Physical Sciences*, vol. 434, no. 1890, pp. 9–13, jul 1991. [Online]. Available: https://doi.org/10.1098/rspa.1991.0075

[87] A. Yoshizawa, "Nonequilibrium effect of the turbulent-energy-production process on the inertial-range energy spectrum," Institute ofIndustrial Science, Uniuersity of Tokyo,, Roppongi, Minato ku, -Tokyo I06, Japan, Tech. Rep. 5, nov 1994. [Online]. Available: https://journals.aps.org/pre/pdf/10.1103/PhysRevE.49.4065

[88] D. G. Ortiz-Suslow and Q. Wang, "An Evaluation of Kolmogorov's 5/3 Power Law Observed Within the Turbulent Airflow Above the Ocean," in *Geophysical Research Letters*, 24th ed. Monterey, CA, USA: Blackwell Publishing Ltd, dec 2019, vol. 46, no. 24, pp. 14 901–14 911. [Online]. Available: https://doi.org/10.1029/2019GL085083

[89] A. Niayifar and F. Porté-Agel, "Analytical modeling of wind farms: A new approach for power prediction," *Energies*, vol. 9, pp. 1–13, oct 2016. [Online]. Available: https://doi.org/10.3390/en9090741

[90] S. Emeis, "Current issues in wind energy meteorology," *Meteorological Applications*, vol. 21, no. 4, pp. 803–819, oct 2014.

[91] A. Pena Diaz, "Sensing the wind profile," National Laboratory for Sustainable Energy, Roskilde, Denmark, Tech. Rep., mar.

[92] M. Mohan and T. A. Siddiqui, "Analysis of various schemes for the estimation of atmospheric stability classification," *Atmospheric Environment*, vol. 32, no. 21, pp. 3775–3781, nov 1998.

[93] D. Wagner, G. Steinfeld, B. Witha, H. Wurps, and J. Reuder, "Low level jets over the southern North Sea," *Meteorologische Zeitschrift*, vol. 28, no. 5, pp. 389–415, jul 2019. [Online]. Available: https://www.researchgate.net/publication/334542824{_}Low{_}Level{_}Jets{_}over{_}the{_}Southern{_}North{_}Sea

[94] M. Imberger, ·. Xiaoli, G. Larsén, and ·. N. Davis, "Investigation of Spatial and Temporal Wind-Speed Variability During Open Cellular Convection with the Model for Prediction Across Scales in Comparison with Measurements," *Boundary-Layer Meteorology*, vol. 179, pp. 291–312, 2021. [Online]. Available: https://doi.org/10.1007/s10546-020-00591-0

[95] E. M. Agee, "Mesoscale cellular convection over the oceans," *Dynamics of Atmospheres and Oceans*, vol. 10, no. 4, pp. 317–341, jan 1987. [Online]. Available: https://doi.org/10.1016/0377-0265(87)90023-6

[96] C. Kamath, "Associating weather conditions with ramp events in wind power generation," *2011 IEEE/PES Power Systems Conference and Exposition, PSCE 2011*, 2011.

[97] M. B. Paskyabi, M. Krutova, H. Bui, and X. Ning, "Multiscale Simulation of Offshore Wind Variability During Frontal Passage: Brief Implication on Turbines Wakes and Load." IOP Publishing, 2022, pp. 1–13.

[98] E. M. Agee, T.S. Chen, and K.E. Dowell, *A review of mesoscale cellular convection*, 10th ed., Bulletin of the American Meteorological Society, Ed. West Lafayette, Ind. 47907: American Meteorological Society, oct 1973, vol. 54. [Online]. Available: http://www.jstor.org/stable/26255239

[99] H. M. Helfand and E. Kalnay, "A Model to Determine Open or Closed Cellular Convection," *Journal of the Atmospheric Sciences*, vol. 40, no. 3, pp. 631–650, mar 1983. [Online]. Available: https://doi.org/10.1175/1520-0469(1983)040{%}3C0631:AMTDOO{%}3E2.0.CO;2

[100] B. Lange, S. Larsen, J. Højstrup, and R. Barthelmie, "Importance of thermal effects and sea surface roughness for offshore wind resource assessment," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 92, no. 11, pp. 959–988, sep 2004.

[101] P. H. Alfredsson, J.-A. Dahlberg, and P. E.J. Vermeulen, "A Comparison Between Predicted and Measured Data from Wind Turbine Wakes," *Wind Engineering*, vol. 6, no. 3, pp. 149–155, 1982. [Online]. Available: https://www.jstor.org/stable/43749217

[102] F. C. Fuertes, C. D. Markfort, and F. Porté-Agel, "Wind Turbine Wake Characterization with Nacelle-Mounted Wind Lidars for Analytical Wake Model Validation," *Remote Sensing 2018, Vol. 10, Page 668*, vol. 10, no. 5, p. 668, apr 2018. [Online]. Available: https://doi.org/10.3390/rs10050668

[103] S. Frandsen, R. Barthelmie, S. Pryor, O. Rathmann, S. Larsen, J. Højstrup, and M. Thøgersen, "Analytical modelling of wind speed deficit in large offshore

wind farms," *Wind Energy*, vol. 9, pp. 39–53, jan 2006. [Online]. Available: https://doi.org/10.1002/we.189

[104] J. Cleijne, "Results of Sexbierum Wind Farm ," Instituut voor Milieu- en Energietechnologie TNO, 7300 AH Apeldoorn, The Netherlands, Tech. Rep., mar 1993. [Online]. Available: http://resolver.tudelft.nl/uuid:7bb5570d-e223-4c4b-9448-bec95453f061

[105] R. G. Flay, "Model tests of wind turbines in wind tunnels," *Czasopismo Techniczne*, vol. 2015, no. Budownictwo Zeszyt 2-B (12) 2015, pp. 63–81, dec 2015. [Online]. Available: https://www.ejournals.eu/Czasopismo-Techniczne/2015/Budownictwo-Zeszyt-2-B-(12)-2015/art/6077/

[106] S. Rodrigues, P. Bauer, and P. A. Bosman, "Multi-objective optimization of wind farm layouts Complexity, constraint handling and scalability," *Renewable and Sustainable Energy Reviews*, vol. 65, no. 1, pp. 587–609, nov 2016.

[107] P. Davidson, *Turbulence: An Introduction for Scientists and Engineers*, 2nd ed., P. Davidson, Ed. Oxford University Press, aug 2015. [Online]. Available: https://oxford.universitypressscholarship.com/view/10.1093/acprof:oso/9780198722588.001.0001/acprof-9780198722588

[108] J. N. Sørensen and W. Z. Shen, "Numerical Modeling of Wind Turbine Wakes," *Journal of Fluids Engineering*, vol. 124, no. 2, pp. 393–399, jun 2002. [Online]. Available: https://asmedigitalcollection.asme.org/fluidsengineering/article/124/2/393/444521/Numerical-Modeling-of-Wind-Turbine-Wakes

[109] S. Raasch and M. Schröter, "PALM - A large-eddy simulation model performing on massively parallel computers," *Meteorologische Zeitschrift*, vol. 10, no. 5, pp. 363–372, oct 2001. [Online]. Available: https://www.schweizerbart.de/papers/metz/detail/10/49174/PALM{_}A{_}large{_}eddy{_}simulation{_}model{_}performing{_}on{_}m?af=crossref

[110] A. Hellsten, K. Ketelsen, M. Sühring, M. Auvinen, B. Maronga, C. Knigge, F. Barmpas, G. Tsegas, N. Moussiopoulos, and S. Raasch, "A nested multi-scale system implemented in the large-eddy simulation model PALM model system 6.0," *Geoscientific Model Development*, vol. 14, no. 6, pp. 3185–3214, jun 2021.

[111] G. C. Larsen, "A simple stationary semi-analytical wake model," Risø National Laboratory for Sustainable Energy, Technical University of Denmark, DK-4000 Roskilde, Denmark., Tech. Rep., mar 2009. [Online]. Available: https://orbit.dtu.dk/en/publications/a-simple-stationary-semi-analytical-wake-model

[112] T. Duc, O. Coupiac, N. Girard, G. Giebel, and T. Göçmen, "Local turbulence parameterization improves the jensen wake model and its implementation for power optimization of an operating wind farm," *Wind Energy Science*, vol. 4, no. 2, pp. 287–302, may 2019.

[113] A. Penã, P. E. Réthoré, and M. P. Van Der Laan, "On the application of the Jensen wake model using a turbulence-dependent wake decay coefficient: The Sexbierum case," *Wind Energy*, vol. 19, no. 4, pp. 763–776, apr 2016.

[114] N. Mittelmeier, J. Allin, T. Blodau, D. Trabucchi, G. Steinfeld, A. Rott, and M. Kühn,

"An analysis of offshore wind farm SCADA measurements to identify key parameters influencing the magnitude of wake effects," *Wind Energy Science*, vol. 2, no. 2, pp. 477–490, jul 2017.

[115] L. Zhan, S. Letizia, and G. V. Iungo, "Optimal tuning of engineering wake models through lidar measurements," *Wind Energy Science*, vol. 5, no. 4, pp. 1601–1622, nov 2020.

[116] M. M. Pedersen, P. van der Laan, M. Friis-Møller, J. Rinker, and P.-E. Réthoré, "Dtuwindenergy/pywake: Pywake," Feb 2019.

[117] B. Gagakuma, A. P. Stanley, and A. Ning, "Reducing wind farm power variance from wind direction using wind farm layout optimization:," *gnireenignE dniW*, vol. 45, no. 6, pp. 1517–1530, jan 2021. [Online]. Available: https://doi.org/10.1177/0309524X20988288

[118] A. Klein and D. Haugland, "Optimization of reliable cyclic cable layouts in offshore wind farms," *Engineering Optimization*, vol. 53, no. 2, pp. 258–276, 2020. [Online]. Available: https://doi.org/10.1080/0305215x.2020.1717482

[119] S. R. Reddy, "Wind Farm Layout Optimization (WindFLO) : An advanced framework for fast wind farm analysis and optimization," *Applied Energy*, vol. 269, p. 115090, jul 2020. [Online]. Available: https://doi.org/10.1016/j.apenergy.2020.115090

[120] J. Fehr and B. Haasdonk, *IUTAM Symposium on Model Order Reduction of Coupled Systems*, 1st ed., ser. IUTAM Bookseries, J. Fehr and B. Haasdonk, Eds. Stuttgart, Germany: Springer International Publishing, may 2018, vol. 36. [Online]. Available: https://doi.org/10.1007/978-3-030-21013-7

[121] D. Hdidouan and I. Staffell, "The impact of climate change on the levelised cost of wind energy," *Renewable Energy*, vol. 101, pp. 575–592, feb 2017. [Online]. Available: https://doi.org/10.1016/j.renene.2016.09.003

[122] J. F. Herbert-Acero, O. Probst, P. E. Réthoré, G. C. Larsen, and K. K. Castillo-Villar, "A Review of Methodological Approaches for the Design and Optimization of Wind Farms," *Energies 2014, Vol. 7, Pages 6930-7016*, vol. 7, no. 11, pp. 6930–7016, oct 2014. [Online]. Available: https://doi.org/10.3390/en7116930

[123] B. A. Sengers, G. Steinfeld, D. Heinemann, and M. Kühn, "A new method to characterize the curled wake shape under yaw misalignment," *Journal of Physics: Conference Series*, vol. 1618, no. 6, pp. 1–11, sep 2020. [Online]. Available: https://iopscience.iop.org/article/10.1088/1742-6596/1618/6/062050

[124] X. xing Yin, Y. gang Lin, W. Li, Y. jing Gu, X. jun Wang, and P. fei Lei, "Design, modeling and implementation of a novel pitch angle control system for wind turbine," *Renewable Energy*, vol. 81, pp. 599–608, sep 2015. [Online]. Available: https://doi.org/10.1016/j.renene.2015.03.042

[125] O. Apata and D. T. Oyedokun, "An overview of control techniques for wind turbine systems," *Scientific African*, vol. 10, pp. 1–13, nov 2020. [Online]. Available: https://doi.org/10.1016/j.sciaf.2020.e00566

[126] FINO1, "FINO1 - research platform in the North Sea and the Baltic No. 1." [Online]. Available: https://www.fino1.de/en/

[127] M. Krutova, M. Bakhoday-Paskyabi, J. Reuder, F. Gunnar Nielsen, and M. Krutova mariakrutova, "Development of an image processing method for wake meandering studies and its application on data sets from scanning wind lidar and large-eddy simulation," *Wind Energy Science*, pp. 1–25, 2022. [Online]. Available: https://doi.org/10.5194/wes-2021-90

[128] RAVE, "Data - RAVE: Research at alpha ventus." [Online]. Available: https://rave-offshore.de/en/data.html

[129] Alpha Ventus, "Alpha ventus." [Online]. Available: https://www.alpha-ventus.de/english

[130] EDR, "Construction of the New alpha ventus Offshore Wind Park to the North of the Island of Borkum." [Online]. Available: https://www.edr.de/en/projects/project/construction-of-the-new-alpha-ventus-offshore-wi/?cHash=8f2bdafb47abf22cfb91d67e38969604&tx_edrprojects_projects{%}5Baction{%}5D=show&tx_edrprojects_projects{%}5BcategoriesSelection{%}5D{%}5B0{%}5D=5&tx_edrprojects_projects{%}5Bcontroller{%}5D=Project

[131] L. Bauer and S. Matysik, "Adwen AD 5-116." [Online]. Available: https://en.wind-turbine-models.com/turbines/1808-adwen-ad-5-116

[132] ——, "REpower 5M Onshore." [Online]. Available: https://en.wind-turbine-models.com/turbines/264-repower-5m-onshore

[133] NORCOWE, "Concluding the OBLEX-F1 campaign," pp. 1–2. [Online]. Available: http://www.norcowe.no/index.cfm?id=433048

[134] European Centre for Medium-Range Weather Forecasts, "ERA5 | ECMWF." [Online]. Available: https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5

[135] Nybø A, Nielsen F G, and Reuder J, "Processing of sonic anemometer measurements for offshore wind turbine applications," *16th Deep Sea Offshore Wind R&D conference*, no. Series 1356 (2019) 012006, 2006. [Online]. Available: https://iopscience.iop.org/article/10.1088/1742-6596/1356/1/012006/pdf

[136] E. Cheynet, J. B. Jakobsen, and C. Obhrai, "Spectral characteristics of surface-layer turbulence in the North Sea," *Energy Procedia*, vol. 137, pp. 414–427, oct 2017.

[137] L. Kristensen, "The Cup Anemometer and Other Exciting Instruments," pp. 1–85, 03 1993.

[138] R. J. Barthelmie, H. Wang, P. Doubrawa, and S. Pryor, "BEST PRACTICE FOR MEASURING WIND SPEEDS AND TURBULENCE OFFSHORE THROUGH IN-SITU AND REMOTE SENSING TECHNOLOGIES," Department of Earth and Atmospheric Sciences, Upson Hall, Ithaca NY 14853, Tech. Rep., jul 2016. [Online]. Available: https://doi.org/10.7298/X4QV3JGF

[139] B. Hennemuth and A. Lammert, "DETERMINATION OF THE ATMOSPHERIC

BOUNDARY LAYER HEIGHT FROM RADIOSONDE AND LIDAR BACKSCATTER," *Boundary-Layer Meteorology*, vol. 120, pp. 181–200, apr 2006.

[140] M. Harris, M. Hand, and A. Wright, "Lidar for Turbine Control," National Renewable Energy Laboratory, 1617 Cole Boulevard, Golden, Colorado 80401-3393, Tech. Rep., jan 2006. [Online]. Available: https://www.researchgate.net/publication/255665204_Lidar_for_Turbine_Control

[141] A. Development Bank, "GUIDELINES FOR WIND RESOURCE ASSESSMENT: BEST PRACTICES FOR COUNTRIES INITIATING WIND DEVELOPMENT ," Tech. Rep., 2014. [Online]. Available: https://www.adb.org/sites/default/files/publication/42032/guidelines-wind-resource-assessment.pdf

[142] M. Krutova, M. Bakhoday-Paskyabi, J. Reuder, F. Gunnar Nielsen, and M. Krutova mariakrutova, "Development of an automatic thresholding method for wake meandering studies and its application to the data set from scanning wind lidar," *Wind Energy Science*, p. 859, 2022. [Online]. Available: https://doi.org/10.5194/wes-2021-90

[143] S. Kongara, R. Calhoun, A. Choukulkar, and M. O. Boldi, "Velocity retrieval for coherent Doppler lidar," *http://dx.doi.org/10.1080/01431161.2011.631948*, vol. 33, no. 11, pp. 3596–3613, 2012. [Online]. Available: https://www.tandfonline.com/doi/abs/10.1080/01431161.2011.631948

[144] M. Debnath, P. Doubrawa, T. Herges, L.A. Martt'nez-Tossas, D.C. Maniaci, and P. Moriarty, "Evaluation of Wind Speed Retrieval from Continuous-Wave Lidar Measurements of a Wind Turbine Wake Using Virtual Lidar Techniques," *Journal of Physics: Conference Series*, 2019. [Online]. Available: https://iopscience.iop.org/article/10.1088/1742-6596/1256/1/012008

[145] M. Krutova, M. Bakhoday-Paskyabi, and J. Reuder, "LiDAR based retrieval algorithm and its verification using SCADA wind data."

[146] D. Van Dinther, C. R. Wood, O. K. Hartogensis, A. Nordbo, and E. J. O'connor, "Observing crosswind over urban terrain using scintillometer and Doppler lidar," *Atmos. Meas. Tech*, vol. 8, pp. 1901–1911, 2015. [Online]. Available: www.atmos-meas-tech.net/8/1901/2015/

[147] P. Benner, S. Grivet-Talocia, A. Quarteroni, G. Rozza, W. Schilders, and L. e. a. Miguel Silveira, *Model Order Reduction*, P. Benner, S. Grivet-Talocia, A. Quarteroni, G. Rozza, W. Schilders, and L. Miguel Silveira, Eds. Berlin, Boston: De Gruyter, dec 2020, vol. 3. [Online]. Available: https://doi.org/10.1515/9783110499001

[148] P. A. LeGresley and J. J. Alonso, "Airfoil design optimization using reduced order models based on proper orthogonal decomposition," *Fluids 2000 Conference and Exhibit*, jun 2000. [Online]. Available: https://www.semanticscholar.org/paper/Airfoil-design-optimization-using-reduced-order-on-LeGresley-Alonso/2abe64371577d794bea2ec0e90de6d7dd950b421

[149] Z. Wang, B. McBee, and T. Iliescu, "Approximate partitioned method of snapshots for POD," *Journal of Computational and Applied Mathematics*, vol. 307, pp. 374–384, dec 2016. [Online]. Available: http://dx.doi.org/10.1016/j.cam.2015.11.023

[150] K. Taira, S. L. Brunton, S. T. M. Dawson, C. W. Rowley, T. Colonius, B. J. Mckeon, O. T. Schmidt, S. Gordeyev, V. Theofilis, and L. S. Ukeiley, "Modal Analysis of Fluid Flows: An Overview," vol. 55, no. 12, pp. 4013–4041, dec 2017. [Online]. Available: https://arc.aiaa.org/doi/10.2514/1.J056060

[151] M. Gunzburger, N. Jiang, and M. Schneier, "An Ensemble-Proper Orthogonal Decomposition Method for the Nonstationary Navier–Stokes Equations," *SIAM Journal on Numerical Analysis*, vol. 55, no. 1, pp. 286–304, feb 2017. [Online]. Available: http://dx.doi.org/10.1137/16M1056444

[152] S. Ullmann, M. Rotkvic, and J. Lang, "POD-Galerkin reduced-order modeling with adaptive finite element snapshots," *Journal of Computational Physics*, vol. 325, pp. 244–258, nov 2016.

[153] J. Quiñonero, Q. Quiñonero-Candela, C. E. Rasmussen, and C. M. De, "A Unifying View of Sparse Approximate Gaussian Process Regression," *Journal of Machine Learning Research*, vol. 6, pp. 1939–1959, 2005.

[154] V. Nasteski, "An overview of the supervised machine learning methods," Faculty of Information and Communication Technologies, Bitola, Macedonia, Tech. Rep., dec 2017.

[155] D. Kristjanson Duvenaud, "Automatic Model Construction with Gaussian Processes," Ph.D. dissertation, University of Cambridge, Pembroke College, jun 2014.

[156] G. Dimitriu, N. Apreutesei, and R. tefnescu, "Numerical simulations with data assimilation using an adaptive POD procedure," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 5910 LNCS, pp. 165–172, 2010.

[157] N. Hamilton, M. Tutkun, and R. B. Cal, "Anisotropic character of low-order turbulent flow descriptions through the proper orthogonal decomposition," *PHYSICAL REVIEW FLUIDS*, vol. 2, p. 14601, 2017.

[158] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)?-Arguments against avoiding RMSE in the literature," *Geosci. Model Dev*, vol. 7, pp. 1247–1250, 2014. [Online]. Available: www.geosci-model-dev.net/7/1247/2014/

[159] A. G. Buchan, C. C. Pain, F. Fang, and I. M. Navon, "A POD reduced-order model for eigenvalue problems with application to reactor physics," *International Journal for Numerical Methods in Engineering*, vol. 95, no. 12, pp. 1011–1032, sep 2013.

[160] R. Friedrich, J. Peinke, M. Sahimi, and M. Reza Rahimi Tabar, "Approaching complexity by stochastic methods: From biological systems to turbulence," *Physics Reports*, vol. 506, no. 5, pp. 87–162, sep 2011.

[161] H. Kantz and T. Schreiber, *Nonlinear time series analysis*. Cambridge, UK: Cambridge University Press, 2004, vol. 2.