DTU Wind Energy
Department of Wind Energy

Inflow estimation using LiDAR measurements and physics-informed machine learning

Master Thesis - Project Plan

Contents

1 Learning Objectives

The learning objectives for this MSc project are compliant with the expected learning outcomes shared by DTU Wind Energy department. The learning objectives are the following:

General objectives

- Communicate the challenges faced during the development of the project and explain the possible solutions considered
- Collaborate with the group members involved in the CONTINUE project to carry out the investigation.

Related to LIDAR measurements

- Understand how LIDARs work and the governing differences between the Continuous Wave (CW) technology and pulsed technology.
- Identify and discuss the limitations of LIDAR technology, and how it impacts the measurements
- Determine the most relevant LIDAR configuration to support the simulations that will be used for the inflow estimation

Related to the establishment of the data-set

- Understand the requirements to produce a FAIR database (Findable, Accessible, Interoperable and Reusable) [1]
- Understand the assumptions behind the Mann model, and the parameters needed to create turbulence boxes.
- Collaborate to create a FAIR database containing Mann boxes and simulated LIDAR measurements from the chosen configurations.

Related to physics-informed machine learning

- Understand how the structure of physics-informed neural networks is relevant for real-time inflow estimation
- Apply the current methods employed in PIML to develop a neural network relevant for inflow estimation based on LIDAR measurements
- Validate the estimation method by simulating loads on the NREL 5MW turbine and comparing the reconstructed turbulence box to the original.

2 Summary of the project

The control of wind turbines is currently mostly relying on feedback methods, where the controller reacts to the state of the turbine and acts upon it. It implies that the turbine needs to be out of equilibrium for the controller to change. That is why feed-forward techniques are being explored, where the controller would now react to information of the flow happening ahead of the turbine, to anticipate the changes. This is expected to reduce loads and lifetime fatigue.

In this regard, LiDAR measurements could prove to be useful, as this measurement technique is able to scan the space ahead of the turbine and provide wind measurements in real time of the incoming flow. The possibility of mounting such devices on the nacelle or the hub of a turbine could prove to be beneficial. However, these measurements remain sparse and do not depict the full rotor plane.

This project thus aims at reconstructing the full incoming inflow, using data measurements from a hub-mounted LiDAR, in movement with the rotor. The measurements will originate from simulated Mann turbulence boxes, where different configurations of the LiDAR will be tested.

The estimation will be achieved through the use of physics-informed machine learning, that will rely on both the simulated measurements and the Navier-Stokes equations to output the velocities. The resulting inflow will be then compared to the initial Mann box to evaluate the performance of the network. Furthermore, simulations of the response of the NREL 5MW [2] turbine will be performed in the aeroelastic tool HAWC2 [3], to evaluate the effect of the errors in estimation.

3 State of the Art

3.1 Introduction

As wind turbines grow bigger in power and size, it has become increasingly important to tackle the resulting loads on the structure, to either achieve a longer lifetime, or optimize the structure to cut cost. The Light Detection and Ranging (LiDAR) technology has been emerging recently as a method to improve the control of wind turbines. The goal is to drift away from the current feedback methods where the turbine is only reacting to a disturbance, and implement feed-forward strategies where the controller would anticipate the changes in wind. Moreover, it is becoming increasingly relevant as the turbines grow bigger and a singular measurement from the hub is no longer a reliable information for the entirety of the rotor plane area. Meteorological masts are also increasingly expensive as they grow taller, especially in offshore conditions [4]; thus LiDARs could become an attractive replacement.

3.2 Wind measurement - LiDAR technology

3.2.1 LiDAR measurements

LiDAR systems are employed to measure the wind speed remotely, in the direction of its beams. The measurement relies on the signal received from back scattered light: the fraction of the beam reflected from the moving particles in the air will have a change in frequency from the Doppler effect. By measuring this difference in frequency in the resulting signal, the speed of the particle can be derived. By assuming the speed of the particle matches the wind, the wind speed in the line of sight is thus acquired [5]. Two main technologies exist to acquire the speed at a specific point in space: Continuous Wave (CW) systems and Pulsed LiDAR. CW LiDARs will focus the beam through a lens to acquire data at a certain distance, while pulsed devices emit laser pulses and rely on time of flight to separate the different distances of measurements. In this project, a pulsed LiDAR is considered, a technology that has the advantage of measuring several distances in a quasi-instantaneous way, but consequently with a slower sampling frequency from one beam to the next[6].

The devices can be ground based, as seen in [7] where the measurements were used to reconstruct the wind field. However, for the purpose of determining indicators at hub height, where the longitudinal component is the most critical, it is more practical to mount the LiDAR on the turbine itself. Two positions can be looked at: on the nacelle behind the blades, pointing forward ('nacelle-mounted'), or rotating with the hub, in front of the blades ('hub mounted'). The main difference between both will be the potential blockage with the blades passing through the line of sight.

The paper [8] reconstructs a wind field by using a simulated LiDAR. The authors model the blockage of the blades by filtering out some data. They estimate that the availability would drop as low as 60% for the lower beams. The article [9] presents an algorithm to improve data reliability, and among other improvements detect when the LiDAR hits the blades, to exclude these points. With this method, it is claimed that for the V52 Vestas turbine, where a Spinner LiDAR was mounted on, data availability from field measurements ranged from 57% to 89%, depending on the position of the beam. It was found that the lower part of the pattern, as expected with increased blade thickness, experienced on average lower availability. Mounting the LiDAR on the hub is expected to raise the availability, as one

clear and constant source of data falsification will be removed. However, it will be harder to define regularly spaced patterns, as the beams for the project are expected to be fixed at a given angle to the hub, and spinning at this fixed angle.

3.2.2 Challenges and shortcomings

Though LiDARs offer clear new use cases by their advantages compared to a cup anemometer, some shortcomings, mainly coming from the spatial averaging and line of sight, are important to underline, to best understand the challenges to overcome when using LiDAR data.

A LiDAR will only measure the line of sight wind velocity, meaning the velocity perceived along the beam direction. This is often referred in the literature as the 'Cyclops Dilemma', that can be circumvented using several assumptions [10]:

- the wind is purely one directional, and this direction is known: the wind can be calculated based on the line of sight wind and the beam angle.
- the wind has no vertical component and the flow is considered homogeneous at a given height. the wind can then be deduced from two separate measurements of the same height.

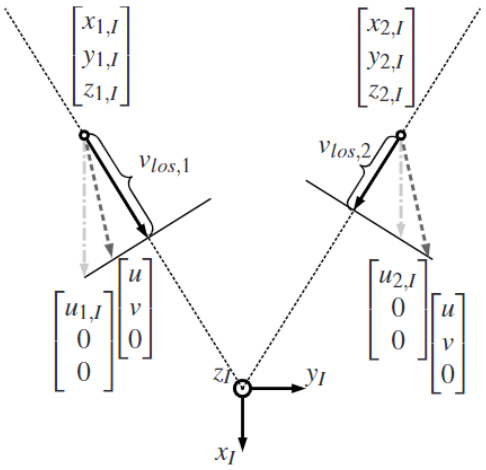


Figure 3.1: The cyclops dilemma: how to derive the actual wind speed using only line-of-sight wind speeds, from [10]

In Figure 3.1, the dash-dotted lines represent the values of the wind derived from the assumption that $v = w = 0$, the vertical and cross-horizontal components of the wind. The dashed lines represent the wind value derived from the assumption of a homogeneous flow and no vertical component, which is thus obtained by using both LOS wind speeds to derive u and v. It can be assumed that adding a third beam at a different angle could allow to remove the assumption $w = 0$, if the homogeneous simplification is kept.

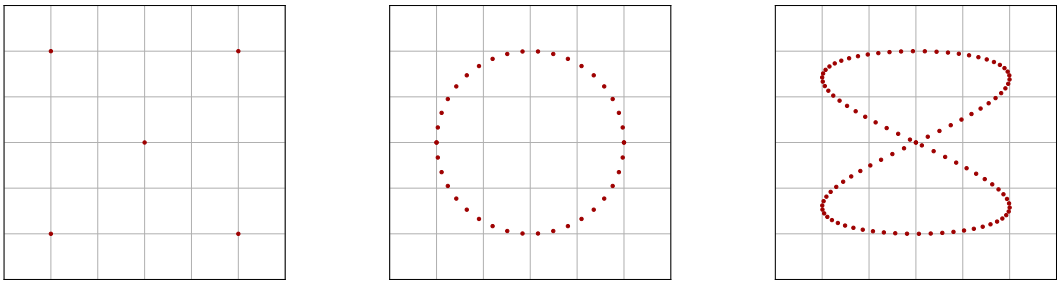
Furthermore, as the pulse used to measure the wind velocity cannot be infinitely small, the resulting wind speed will be an average of the line of sight velocity within a cylinder the size of the beam. This will result in the loss of higher frequency content, due to the spatial averaging.

However, as can be seen in Figure 3.1, the configuration of the beams has a major role to play in improving the reliability of the results, as a singular beam with a fixed angle would likely result in partial information on the wind only.

3.2.3 LiDAR configuration

The article [11] simulates several LiDAR configurations, where the turbulence field is reconstructed for load validation using constrained Gaussian turbulence field. The constraints are imposed using simulated LiDAR data from 6 different patterns,the performance of each is evaluated based on the variance in loads observed after simulating cases with and without constraints on the turbulence boxes. The simulations are based on Mann turbulence boxes and the simulations are run in HAWC2. The main categories of LiDAR configurations studied are:

- point-based measurements: singular central one, but also four points placed as a square with a central point as seen in Figure 3.2a
- moving patterns: from a simple circular pattern with no central beam as seen in Figure 3.2b, to a more intricate Lissajous curve plotted in Figure 3.2c. The paper explores as well the Spinner LiDAR configuration presented in detail in [12].



(a) Square design (b) Circle design (c) Lissajous curve

Figure 3.2: Three types of LiDAR configuration, derived from [11]

It is shown that for the purpose of load validation, the smallest uncertainty was reached for the moving beam patterns, with the exception of the simple circular pattern. However, the authors underline how their specific use case (load validation) could influence the choice of pattern: an example they cite found the circular pattern to be ideal for yaw realignment. Thus no optimal pattern can be chosen for inflow estimation without sensitivity analyses.

3.3 LiDAR-based inflow estimation methods

A substantial effort is currently put forward to use LiDAR measurements from in the wind energy sector, with goals varying from extracting wind characteristics to recreate the entire inflow.

3.3.1 Partial inflow reconstruction

A first approach to the use of sparse LiDAR measurements lies in determining spatially averaged values, or characteristics from the flow that used to be estimated on a singular measurement from a cup anemometer, such as the free stream velocity, or the Rotor-Averaged Wind Speed (RAWS).

An example of this approach can be found in [8], where the RAWS and the Rotor-Averaged Wind Direction (RAWD) are determined from simulated LiDAR data. The fictitious wind field is generated based on the Kaimal spectrum with added vertical sheer following a power law. To simulate the actual measurements, the authors added Gaussian white noise and a Gaussian weighting function was used to account for the spatial averaging

of the device. The reconstruction of the wind flow is treated as an optimization problem, where a discrete-time Kalman filter is implemented´and several simplification assumptions are made on the 3D flow, among which the Taylor frozen turbulence hypothesis. The assumptions lead to having full knowledge of the global correlation matrix of the wind components at specific locations. This matrix is then used in the Kalman filter to estimate the flow. However, only averaged quantities are displayed in the article, so it is not known if such a method would be applicable for the full 3D reconstruction. The performance of the method was evaluated by calculating the Root Mean Squared Error (RMSE) of the selected indicators, which varied from $0.405\,\mathrm{m\,s^{-1}}$ to $0.818\,\mathrm{m\,s^{-1}}$ for the RAWS.

In [13], the free-stream velocity is the target of the estimation, by using experimental LiDAR measurements, despite the measurements being taken in the induction zone. The author use the Levenberg-Macquardt algorithm to minimize the error between the measurement and the wind model from which they deduce the free stream velocity. This wind model assumes horizontal homogeneity, a known vertical shear profile and homogeneous wind direction. Furthermore, a one-dimensional induction model is added to account for the effect of the rotor onto the stream. With this method, was found that the modeled wind speed was withing 1-1.5% of the mast-measured results.

Alternatively, full reconstruction of the wind field without focusing on perfect replication has been explored in [11]. The emphasis is here not put on simulating a flow field very accurately, but rather getting the subsequent loads close, in order to give a better prediction of energy prediction and loads than the current methods prescribed the IEC standard[14]. The method used to produce the turbulent field relies on constrained Gaussian turbulence, where only the longitudinal component of the wind u is constrained, for simplicity. Mann turbulence boxes are created, and their constrained versions are derived from it by selecting what would correspond to LiDAR measurements, and modifying the cross correlation matrices used in the turbulence model accordingly. A graphical representation of the constrained Gaussian method is displayed in Figure 3.3

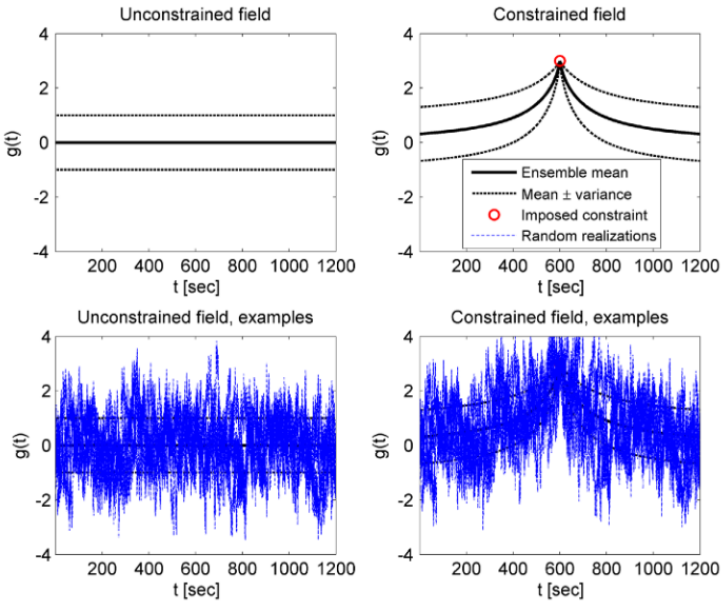


Figure 3.3: One-dimensional, zero-mean and unit-variance Gaussian field on which a single constraint is imposed, from [11]

3.4 Physics informed machine learning

As the previous section showed, achieving full flow reconstruction based only the sparse data LiDAR measurements provide is a complex task, that result in heavy assumptions (flow homogeneity for example), or settle to compute aggregated variables. Machine Learning-based method could potentially bridge the gap, as they are known to be more flexible with the use of noisy data, with a smaller computational cost in some instances [15]. However, such methods generally require a lot of data for the training of the neural network before being able to output results using the trained network. Such datasets can be prohibitively expensive to get to apply this method. However, Physics-informed Machine Learning (PIML) methods offer the alternative to both rely on data and known equations that define the state to estimate. Thus less data would be needed, as the solutions are already more constrained using Partial Differential Equations (PDEs).

3.4.1 Method and motivation

A generic neural network can be defined as L layers, containing each $n_{l\in[1,L]}$ neurons. For a fully connected network, each neuron of a given layer is connected to each neuron of the next layer; the connection is materialized by a weight $w^l_{o,p}$ where $l \in [1, L]$ materializes the l-th layer, $o \in [1, n_l]$ the o-th neuron of this layer, connected to the p-th neuron of the next layer: $p \in [1, n_{l+1}]$. Each neuron has additionally a biais attached to itself, b^l_o, that acts as an activation threshold. To train the network means to adjust each weight and biais for every data fed to the network during the learning phase, so that the input matches the desired output, by minimizing a defined loss function [16].

In the case of Physics-informed Neural Network (PINN), the loss function is not just relying on data, but on a defined set of (partial) differential equations. By taking the general form given in [17], let us define the problem:

$$u_t + \mathcal{N}[u(t,x)] = 0, x \in \Omega, t \in [0,T] \tag{3.1}$$

Where $u(t,x)$ is the solution to estimate, and $\mathcal{N}[\cdot]$ is a nonlinear differential operator, Ω is a subset of $\mathbb{R}^D$.

The addition of the differential operator, compared to a conventional neural network, result in a change in the lean squared error loss that will now be:

$$MSE = MSE_u + MSE_f \tag{3.2}$$

Where $f = u_t + \mathcal{N}[u(t,x)]$, and:

$$MSE_u = \frac{1}{N_u}\sum_{i=1}^{N_u} \left|u\left(t_u^i, x_u^i\right) - u^i\right|^2$$

is the root mean squared error to minimize to get close to the data points;

$$MSE_f = \frac{1}{N_f}\sum_{i=1}^{N_f} \left|f\left(t_f^i, x_f^i\right)\right|^2$$

is the root mean squared error to minimize to comply with the partial differential equations.

Such a method can provide solutions to otherwise complex or analytically unsolvable equations, such as the coupled Stokes-Darcy equation in [18]. A PINN was used to successfully model a complex interface problem with heterogeneous conditions.

As a flow field can be described by the Navier-Stokes equations, PINN can be a method to explore in the case of flow reconstruction.

3.4.2 Application to inflow estimation

The estimation of an incoming flow has been carried out using PINN, in two consecutive papers by the same main author. In [19], the method is first tested using only two-dimensional Navier Stokes equations, and simulated LiDAR measurements. The research scope is then extended in [20], where the estimated flow is now three-dimensional, and the effective viscosity from the equations is no longer assumed but becomes an estimated parameter. If the results are promising, with the RMSE for all three wind components being consistently below 13% of the mean wind speed, only a single LiDAR configuration is explored. Furthermore, no explanation or hypothesis is given for the increasing error away from the rotor center, which could become challenging for bigger scopes: their rotor diameter was only of about 60m, while the NREL 5MW as a diameter of 126m.

4 Research Questions

RQ1: What LiDAR configurations can be achieved using a rotating hub-mounted device?

The configurations should be compliant with the LiDAR that will be used in the CONTINUE project. The specifications can be found in [21]. This device is a pulsed LiDAR with up to 6 beams, and a sampling frequency of 1Hz. The measurements cannot be taken from several beams simultaneously. However, it is assumed that all distances from a given beam are measured at the same time.

RQ2: Can the inflow be reconstructed for a Mann box large enough to fit the NREL 5MW turbine, using PIML?

The papers [19] and the continued work in [20] used a method relying on PINN to reconstruct a turbulent flow. If the results looked promising, the wind turbine considered was however small (with a diameter of 60 meters), and the error in estimation was progressively worse the further away from the hub center. It will thus be crucial in this work to adapt the method used to improve this effect.

RQ3: Is there a LiDAR configuration that minimizes the error in the estimation of the flow?

Several LiDAR configurations will be explored in this work, and an assessment of the optimal setup will be carried out.

RQ4: Is the error in estimation acceptable with regards to the resulting difference in loads observed in simulation?

Once the estimations have been carried out, load simulations should be performed in HAWC2 on the NREL 5MW, to compare the response of the turbine between the original Mann box and its estimated counterpart. This analysis would provide additional performance indicator, mainly on the influence of errors on the structure.

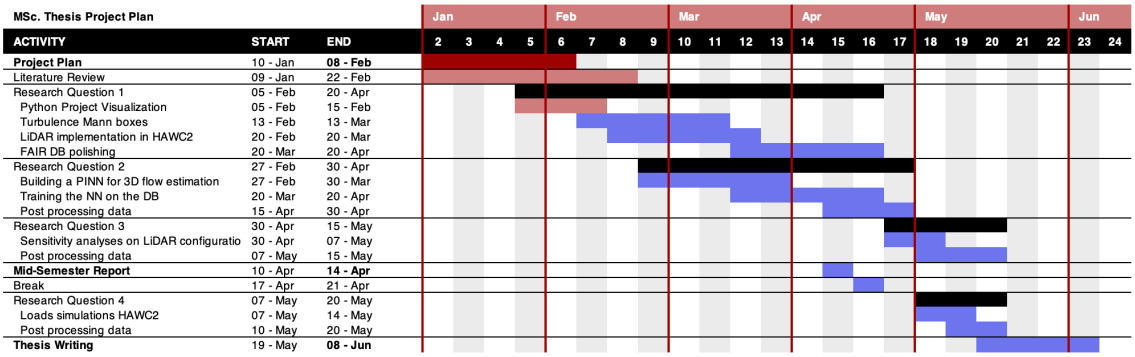
If time allows for it, the following research question could also be explored:

RQ5: Can the method be improved by changing the parameters of the neural network?

A neural network can be defined by its number of layers and the number of nodes per layer (the 'neurons'), which will result in the degrees of freedom of the training variables at hand. The network used in [19] uses 12 hidden layers of 128 neurons each, resulting in 149378 possible weights and biases to change during the training phase. These parameters are chosen during the elaboration of the structure, and a sensitivity analysis could be carried out to estimate a combination that would improve the training time and the estimation results.

5 Project Management

The provisional calendar for the thesis tasks is displayed in Figure 5.1



MSc. Thesis Project Plan		
ACTIVITY	**START**	**END**
Project Plan	10 - Jan	**08 - Feb**
Literature Review	09 - Jan	22 - Feb
Research Question 1	05 - Feb	20 - Apr
Python Project Visualization	05 - Feb	15 - Feb
Turbulence Mann boxes	13 - Feb	13 - Mar
LiDAR implementation in HAWC2	20 - Feb	20 - Mar
FAIR DB polishing	20 - Mar	20 - Apr
Research Question 2	27 - Feb	30 - Apr
Building a PINN for 3D flow estimation	27 - Feb	30 - Mar
Training the NN on the DB	20 - Mar	20 - Apr
Post processing data	15 - Apr	30 - Apr
Research Question 3	30 - Apr	15 - May
Sensitivity analyses on LiDAR configuratio	30 - Apr	07 - May
Post processing data	07 - May	15 - May
Mid-Semester Report	10 - Apr	**14 - Apr**
Break	17 - Apr	21 - Apr
Research Question 4	07 - May	20 - May
Loads simulations HAWC2	07 - May	14 - May
Post processing data	10 - May	20 - May
Thesis Writing	19 - May	**08 - Jun**

Figure 5.1: Thesis plan. In dark red, are the completed tasks; light red the tasks in progress and in blue the penging tasks.

Nomenclature

CW Continuous Wave.

LiDAR Light Detection and Ranging.

PDEs Partial Differential Equations.

PIML Physics-informed Machine Learning.

PINN Physics-informed Neural Network.

RAWD Rotor-Averaged Wind Direction.

RAWS Rotor-Averaged Wind Speed.

RMSE Root Mean Squared Error.

Bibliography

[1] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, *et al.*, "Comment: The FAIR Guiding Principles for scientific data management and stewardship," *Scientific Data*, vol. 3, Mar. 2016, ISSN: 20524463. DOI: 10.1038/sdata.2016.18.

[2] J. Jonkman, S. Butterfield, W. Musial, and G. Scott, "Definition of a 5-MW Reference Wind Turbine for Offshore System Development," Tech. Rep., 2009. [Online]. Available: http://www.osti.gov/bridge.

[3] A. Melchior Hansen, T. Juul Larsen, and Danmarks Tekniske Universitet. Risø DTU, *How 2 HAWC2, the user's manual*. Risø DTU - National Laboratory for Sustainable Energy, 2007, ISBN: 9788755035836.

[4] W. Fu, A. Sebastiani, A. Peña, and J. Mann, "Dependence of turbulence estimations on nacelle-lidar scanning strategies," DOI: 10.5194/wes-2022-85. [Online]. Available: https://doi.org/10.5194/wes-2022-85.

[5] C. Slinger and M. Harris, "Introduction to continuous-wave Doppler lidar," Tech. Rep., 2012.

[6] F. Dunne, E. Simley, and L. Y. Pao, "LIDAR Wind Speed Measurement Analysis and Feed-Forward Blade Pitch Control for Load Mitigation in Wind Turbines," Tech. Rep., 2010. [Online]. Available: http://www.osti.gov/bridge.

[7] C. Stock-Williams, P. Mazoyer, and S. Combrexelle, "Wind field reconstruction from lidar measurements at high-frequency using machine learning," in *Journal of Physics: Conference Series*, vol. 1102, Institute of Physics Publishing, Oct. 2018. DOI: 10.1088/1742-6596/1102/1/012003.

[8] F. Guillemin, H. N. Nguyen, G. Sabiron, D. Di Domenico, and M. Boquet, "Real-time three dimensional wind field reconstruction from nacelle LiDAR measurements," in *Journal of Physics: Conference Series*, vol. 1037, Institute of Physics Publishing, Jun. 2018. DOI: 10.1088/1742-6596/1037/3/032037.

[9] N. Angelou and M. Sjöholm, "Data Reliability Enhancement for Wind-Turbine-Mounted Lidars," *Remote Sensing*, vol. 14, no. 13, Jul. 2022, ISSN: 20724292. DOI: 10.3390/rs14133225.

[10] S. Raach, D. Schlipf, F. Haizmann, and P. W. Cheng, "Three dimensional dynamic model based wind field reconstruction from lidar data," in *Journal of Physics: Conference Series*, vol. 524, Institute of Physics Publishing, 2014. DOI: 10.1088/1742-6596/524/1/012005.

[11] N. Dimitrov and A. Natarajan, "Application of simulated lidar scanning patterns to constrained Gaussian turbulence fields for load validation," *Wind Energy*, vol. 20, no. 1, pp. 79–95, Jan. 2017, ISSN: 10991824. DOI: 10.1002/we.1992.

[12] A. ; Peña, J. ; Mann, and R. Thorsen, "General rights SpinnerLidar measurements for the CCAV52," *DTU Wind Energy*, vol. 0177, 2019.

[13] A. Borraccino, D. Schlipf, F. Haizmann, and R. Wagner, "Wind field reconstruction from nacelle-mounted lidar short-range measurements," *Wind Energy Science*, vol. 2, no. 1, pp. 269–283, 2017, ISSN: 23667451. DOI: 10.5194/wes-2-269-2017.

[14] International Electrotechnical Commission., *IEC 61400-1 Part 1 : design requirements*, p. 168, ISBN: 9782832262535.

[15] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, *Physics-informed machine learning*, Jun. 2021. DOI: 10.1038/s42254-021-00314-5.

[16] Michael A. Nielsen, *Neural Networks and Deep Learning*. Determination Press, 2015.

[17] M. Raissi, P. Perdikaris, and G. E. Karniadakis, “Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations,” *Journal of Computational Physics*, vol. 378, pp. 686–707, Feb. 2019, ISSN: 10902716. DOI: 10.1016/j.jcp.2018.10.045.

[18] R. Pu and X. Feng, “Physics-Informed Neural Networks for Solving Coupled Stokes–Darcy Equation,” *Entropy*, vol. 24, no. 8, Aug. 2022, ISSN: 10994300. DOI: 10.3390/e24081106.

[19] J. Zhang and X. Zhao, “Spatiotemporal wind field prediction based on physics-informed deep learning and LIDAR measurements,” *Applied Energy*, vol. 288, Apr. 2021, ISSN: 03062619. DOI: 10.1016/j.apenergy.2021.116641.

[20] J. Zhang and X. Zhao, “Three-dimensional spatiotemporal wind field reconstruction based on physics-informed deep learning,” *Applied Energy*, vol. 300, Oct. 2021, ISSN: 03062619. DOI: 10.1016/j.apenergy.2021.117390.

[21] Sensup by Lumibird, *The Windfield: For wind measurement*, 2023.