

TNO PUBLIC

TNO report

TNO 2021 R11759

**Application of Gaussian Processes to
Dual-Doppler LiDAR scanning measurements
for high frequency wind field reconstruction**

Westerduinweg 3
1755 LE Petten
P.O. Box 15
1755 ZG Petten
The Netherlands

www.tno.nl

T +31 88 866 50 65

Date	11 November 2021
Author(s)	Nassir Cassamo, Marco Turrini, Dennis Wouters, Arno van der Werff, Wim Castricum, Jan Willem Wagenaar
Copy no	
No. of copies	
Number of pages	49 (incl. appendices)
Number of appendices	0
Sponsor	TNO
Project name	GLOBE
Project number	060.46522

All rights reserved.

No part of this publication may be reproduced and/or published by print, photoprint, microfilm or any other means without the previous written consent of TNO.

In case this report was drafted on instructions, the rights and obligations of contracting parties are subject to either the General Terms and Conditions for commissions to TNO, or the relevant agreement concluded between the contracting parties. Submitting the report for inspection to parties who have a direct interest is permitted.

© 2021 TNO

TNO PUBLIC

Revision history

Rev.	Date	Description
1	07-10-2021	First draft version of the report
2	11-11-2021	Report reviewed by Fraunhofer IWES
3	12-11-2021	Final version of report
4	12-11-2021	Report released

Summary

The work here presented aims to demonstrate the potential of applying Gaussian Processes regression, a machine learning technique, for the purposes of extracting more information from scanning LiDARs measurements.

With such goal in mind, two major tasks constitute the bulk of the work here presented: (1) application of Gaussian Processes to one single scanning LiDAR for data interpolation and (2) application of Gaussian Processes to two scanning LiDARs operating in dual-doppler mode for wind field reconstruction.

For such tasks, data measurements from a test campaign at Bremerhaven test field from Fraunhofer IWES are used. These were collected from the 14th of March until the 9th April. The data measurements collected and used for this work include data from 2 scanning LiDARs operating in dual-doppler mode, a cup anemometer positioned at 55 meters of height and a wind vane positioned at 110 meters of height, the last two installed in a IEC-compliant met mast.

Application of Gaussian Processes to one single scanning LiDAR for data interpolation.

The Gaussian Processes are first fitted onto one single scanning LiDAR measurements and the GP is then used to predict the radial wind speed at the location of the cup anemometer. The two, predicted value and measured value are then compared.

- **It is shown that the GP can provide accurate radial wind speed predictions.** A linear regression between the predicted and actual measurements for the campaign duration shows an R^2 value of 0.934 and linear fit of $y = 1.01x + 0.08$. For 10 minute statistics the regression shows an R^2 value of 0.979 and linear fit of $y = 1.01x + 0.05$.
- **It is shown that proper filtering of the radial wind speeds measured by the LiDAR is necessary to ensure accurate predictions.** Despite using simple filters based on Carrier to Noise Ratio (CNR) thresholds and quality flags of the LiDAR radial wind speed measurements, the data may still contain erroneous measurements¹. The origin of these is sometimes unknown. Filtering such values leads to having more precise GP models, which in turn are able to provide more accurate predictions. Although the improvements are not major (erroneous measurements represent a small part of the overall measurements), if the user wishes to have the most accurate predictions yielded by the GP, this step may prove useful.

Application of Gaussian Processes to two scanning LiDARs operating in dual-doppler mode for wind field reconstruction

A method based on applying two independent GPs to two scanning LiDARs operating in dual-doppler mode is presented. The 2 independent GPs are able to predict the value of radial wind speed for each one of the LiDARs. The predictions are done over the same cartesian grid, which means that for a certain time instant, all points in space will contain information of the radial wind speed as seen by the 2 LiDARs. The

¹the erroneous measurements were used for the full comparison as they are only present in minor time instances. However the impact of filtering them out was studied in separate.

problem of calculating the u and v wind field components boils down to a problem of calculating two unknowns - the u and v - based on two known values - the predicted radial wind speeds yielded by each GP - by using simple trigonometry.

- **It is shown that the developed method is able to accurately reconstruct wind field measurements.** For the analysed time window of 4 days, the comparison of the measured (at the met-mast) and numerically calculated u value yielded an R^2 value of 0.937 and a linear fit of $y = 1.02x - 0.19$. The comparison of the measured and numerically calculated v value yielded an R^2 value of 0.892 and a linear fit of $y = 0.96x - 0.45$. After filtering the data on which the GP was fitted, these values changed. More specifically, the comparison of the measured and numerically calculated u value increased to an R^2 value of 0.941 and a linear fit of $y = 1.02x - 0.20$ and the comparison of the measured and numerically calculated v value increased to an R^2 value of 0.904 and a linear fit of $y = 0.95x - 0.42$.

The results presented in this report are further evidence that the Gaussian Processes are a suitable machine learning technique that can be applied to scanning LiDARs to accurately interpolate radial wind speeds at points in space and time where direct measurements from the LiDAR are not available. In this way, the results of this work provide added credibility to the application of the GPs in the context of radial wind speed interpolation. The results and applicability of the GPs may vary depending on the type of scanning pattern chosen.

Furthermore, the developed method that applies the GPs to dual doppler scanning measurements to perform high frequency wind field reconstruction can be of immense use to further gather insights into the wind field at a certain locations. This method will be applied in the future to an offshore campaign with the final aim to gather insight into global blockage effects upstream and downstream of a wind farm.

Although the development of the GPs (i.e., the complete algorithmic implementation developed within TNO) has reached an advanced stage, further points of improvement have been noticed. The first would aim at increasing the robustness of the GPs to issues in the data (gaps in the data complicate the usage of the GPs in an online fashion). A second point of improvement would be to standardize the application of filtering techniques to radial wind speed measurements. Another relevant point is to further corroborate the accuracy of the wind field reconstructions, that could be achieved in simulations or in an experimental setup.

Acknowledgements: The AFFABLE project is co-financed by TKI-Energy from the 'Toeslag voor Topconsortia voor Kennis en Innovatie (TKI's)' of the Ministry of Economic Affairs and Climate. Fraunhofer IWES supported some of this work by providing access to data and infrastructure within the "Testfeld BHV" project funded by the German Federal Ministry for Economic Affairs and Energy on the basis of a decision by the German Bundestag (FKZ 0324148). Results and conclusions are responsibility of TNO.

Contents

	Summary	3
1	Introduction	6
1.1	Project context	6
1.2	Background on Gaussian Processes Regression	6
1.3	Application of Gaussian Processes to LiDAR	7
1.4	Methodology	9
2	Onshore campaign	10
2.1	Setup and instrumentation	10
2.2	Data acquisition and visualisation	14
2.3	Baseline result and scanning pattern impact	15
3	Application of Gaussian Processes	18
3.1	Goal and methodology	18
3.2	Results	18
3.3	Conclusions and discussion	32
4	Combination of LiDARS	33
4.1	Goal and methodology	33
4.2	Application of the Gaussian based Multi Lidar method	34
4.3	Conclusions and discussions	38
5	Uncertainty evaluation	39
5.1	Goal and methodology	39
5.2	Results	39
5.3	Conclusions and discussions	44
6	Conclusions and future work	46
6.1	Conclusions	46
6.2	Future work	48
7	Bibliography	49

1 Introduction

The purpose of the AFFABLE project is twofold. Firstly, to test and validate the application of Gaussian Process regression to one scanning LiDAR. Secondly to develop a new methodology that makes use of the GPs and measurements from two scanning LiDARs operating in dual Doppler mode to reconstruct the wind field.

These two goals are of relevance to further gain confidence in using the GPs to interpolate radial wind speed measurements from scanning LiDARs and to be able to do two dimensional wind field reconstructions (i.e., planar reconstructions without the vertical wind component) to gather insight into the wind behavior. The latter is applicable within the scope of the GLOBE project, where blockage effects may be visible from planar reconstructions upstream of the wind farm.

To achieve the above-mentioned goals, the GP radial wind speed prediction and the wind field reconstruction are compared to measurements from a cup anemometer and wind vane. This data was gathered from an onshore campaign with an approximate duration of 1 month.

This introductory chapter (1) contextualizes this report in terms of the projects it is related to (2) provides background information on the Gaussian Processes regression (GPs), the Machine Learning technique used as the backbone of the presented work (3) describes the main steps in which it is validated and applied (4) inform the reader on the methodology and structure of this report.

1.1 Project context

The OWA GLoBE project aims to assess Global Blockage Effects (GBE) in a large offshore measurement campaign, where the role of TNO is to support the campaign design. The AFFABLE project is a project between RWE and TNO, where RWE represents the OWA GLoBE partner.

The scope of the AFFABLE project includes the following:

1. Show how the use of GPs in combination with scanning LiDAR data can increase the accuracy from the Dual-Doppler baseline technique.
2. Execute an onshore test to demonstrate the capability of GPs in estimating a wind velocity field and its associated uncertainty from sparse measurements.
3. Validate wind farm modeling towards GBE as a blueprint for the Independent Technology Review Group (ITRG), a group of experts in the field of wind farm modeling.

The contents of this report are within the context of the second point above, covering the application, analysis, and investigation of the added value of the GPs to data from the onshore campaign.

1.2 Background on Gaussian Processes Regression

Scanning LiDARs measure at different points in space by scanning through a certain pre-determined area. Compared to other types of LiDARs, there is only one Line Of Sight (LOS) that scans through the area. This means that information of radial wind

speed will be missing at certain locations in some time instances. The gap in information and the unique patterns in which data is gathered set the perfect conditions to perform interpolation and prediction of measurements at times and space locations where information is unavailable.

Model-based interpolation, i.e, the classical approach that makes use of the underlying physical equations to derive a model, becomes cumbersome due to the complexity of the underlying equations governing the wind behavior. Data-driven models, where the underlying equations are reverse-engineered from data, offer a workaround to this. Gaussian Processes are one of the possibilities to perform the desired interpolation

Gaussian Processes (GPs) regression is an interpolation method based on Gaussian processes. The goal is to predict the value of a function at a given point by computing a weighted average of the known values of the function in the neighborhood of the point. The distribution obtained at this point is defined by a mean function $\mu(x)$ and a positive definite covariance function $k(x, x')$. The covariance function k captures the relationship between the variables involved in the process. The selection and preparation of this covariance function define the prior knowledge of the Gaussian Process model. The method has underlying hyperparameters that are optimised through a fitting process. This locates Gaussian Processes regression within the field of Machine Learning

As opposed to the standard regression methods - where a certain function (or model) describing the behavior of the underlying system is assumed (for example, a linear or quadratic one) and the known parameters are calculated from data (what is often referred to as training data set in the context of machine learning) in order to minimize a cost function which penalizes the distance between the data and predictions given by the model - the GP does not assume an underlying function [1].

For a more thorough technical description of the GPs the reader is referred to [2]. For a hands-on introduction to GPs, the reader is referred to [3].

1.3 Application of Gaussian Processes to LiDAR

The process of reconstructing wind velocity from LiDAR measurements using GPs interpolation can be summarised in two main steps:

1. **Prediction of radial wind speed (RWS)**, i.e. the creation of a *virtual lidar*. This terminology refers to a GP model that can be used to estimate² the RWS in, theoretically, any point in space and time.
2. **Wind velocity reconstruction**, i.e. the inference of the real wind field from radial wind speed virtual measurements at key locations and time instants. For this step, several assumptions can be made and methodologies used, and it is out of the scope of this project to investigate this. A list of possible methodologies taken in the literature can be consulted here [2].

To accomplish the first step, i.e., to create the virtual LiDAR and obtain radial wind speed predictions, five steps must be taken [2, 1].

1. **Data cleaning:** A pre-processing step of data cleaning by performing filtering should be done, with no in-built rules for filtering within the GPs application. The

²The terms estimation/interpolation/prediction are hereafter used as synonyms.

type of filtering technique to use is to be decided by the end-user depending on the application.

2. **Estimation of bulk wind velocity:** Infer a mean wind velocity during each scanning pattern repeat period which separates the slow-varying bulk velocity flow through time from the local turbulence.
3. **Fitting of GP1 hyperparameters:** A GP is fitted to the radial speed measurements (turbulent component only) over a short period of time, thus ensuring that the GP learns only the local and systematic turbulent fluctuations in the flow field. The GP fitted into this short time scale is referred to as a GP1. [1].
4. **Fitting of GP2 hyperparameters:** A second GP is fitted to the hyperparameters of the local GP1s and the local bulk wind velocities estimates. This separate GP, referred to as a GP2, is fitted onto a longer time window.
5. **Radial wind speed prediction:** By making use of the fitted GPs and the original data, radial wind speeds at other points than the ones measured, for pre-specified time instants, can then be inferred.

These steps are further illustrated in Figure 1. Using the typical terminology in the context of Machine Learning, the training of the model is seen in the color grey and the prediction in the color blue.

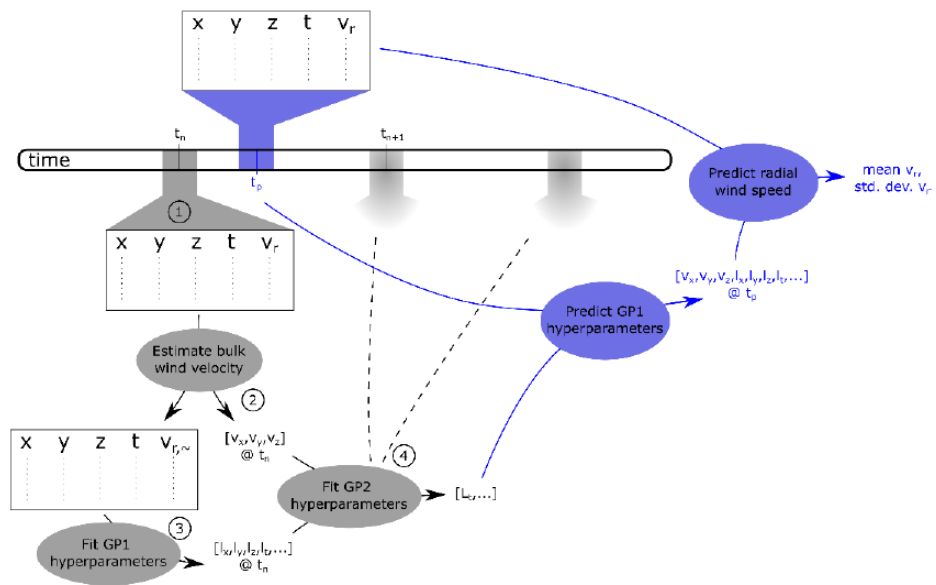


Figure 1: Schematic representation of the steps involved in the creation of a virtual LiDAR for the purposes of predicting radial wind speed. 1: data from a certain short time window is taken. 2: the turbulent fluctuations in the same short time window are identified by removing the bulk wind velocity. 3: A GP1 is fitted to this data and its hyperparameters are optimized. 4: A GP2 is fitted onto the hyperparameters of different GP1s and the blue wind velocity for the same time windows. This fitting is done over a longer time window. Schematics taken from [2].

Steps (2) and (3) are the core of the method. The tool is a python package named GPyDAR developed by TNO. The methodology mentioned has been tested and validated in previous works by Stock-Williams [2, 4, 5].

In the work in [4], where the GPs are applied to data from TNO's test site EWTW, four scanning patterns were tested, altering the speed and azimuthal scanning range. The results found indicated that the scanning LiDAR measurements - which were

filtered only to remove pulse reflections from the met mast - processed through the GP reconstruction method, could be used to reproduce reliable 1Hz measurements. Different scanning patterns were here tested (fast-narrow, fast-wide, slow-narrow, slow-wide) for different amounts of time. The application of the GPs to the slow-wide data set (data set with a length of 25 days) showed a bias of -0.05 m/s and a scatter of 0.91.

1.4 Methodology

The final result of this work is to make use of the Gaussian Processes and dual-Doppler LiDAR measurements to perform wind field reconstructions. For that end result, several steps of increased complexity are taken. Starting from a pure comparison between the LiDAR and cup anemometer measurements without making use of the GPs to comparison of the GPs predictions to the cup anemometer and wind vane measurements and only then the wind field reconstruction with an associated uncertainty of the prediction yielded by the GP.

The report follows these steps and is structured in four sections:

1. **On-shore campaign:** the onshore campaign is described. The setting is specified along with available instrumentations and the acquired data is visualised. A baseline result that compares the cup anemometer wind speed with the LiDAR without the usage of Gaussian Processes is also shown.
2. **Application of Gaussian Processes:** the Gaussian Processes are applied to a scanning LiDAR and the accuracy of the predictions is evaluated by comparing the latter with data from a cup anemometer and wind vane.
3. **Combination of LiDARs:** the information contained in the 2 scanning LiDARs available at the site is combined to go from radial wind speed predictions to instantaneous flow field reconstruction. The high-frequency reconstruction of wind speed at several spatial points simultaneously is only possible by leveraging the GPs. This methodology has been dubbed GPbML (Gaussian Processes based Multi LiDAR).
4. **Uncertainty evaluation:** the uncertainty yielded by the GP is evaluated for both point and grid predictions. The uncertainty is propagated from the radial wind speed to the wind field reconstruction, allowing to compute high-frequency wind field reconstructions onto a large spatial grid (approximate size of 40 thousand square meters) and the corresponding confidence interval of that same prediction.

2 Onshore campaign

The onshore campaign is performed as a test setup for the offshore campaign that will take place within the GLOBE project with the final aim of assessing Global Blockage Effects. In the offshore campaign, measurements of dual Doppler LiDARs will be made available. The rationale for the onshore campaign is then to test the application of the Gaussian Processes similarly as in the offshore campaign.

2.1 Setup and instrumentation

The onshore campaign took place at the test field of Bremerhaven, in a collaborative effort with Fraunhofer Institute for Wind Energy Systems (IWES).

The campaign started on 12 March 2021 and ended on 9 April 2021.

Available at the site where three WindCube 200s LiDARs (two scanning LiDARs functioning in Dual-Doppler mode and a Line Of Sight (LOS) reference) and a 115 m IEC-compliant meteorological mast equipped with cup anemometers, sonic anemometers, and wind vanes at different height levels.

The LiDAR setup concerning the mast can be visualised in Figures 2 and 3. The two WindCubes operating in dual-doppler mode are placed in the northern and western positions relative to the mast. The azimuthal range of the scanning patterns of each one is qualitatively represented by blue lines.



Figure 2: LiDARs at the Bremerhaven test field site. Image copyrights Fraunhofer. On the left, the western scanning and staring LiDARs. On the right, the northern scanning LiDAR.

The northern LiDAR from Fraunhofer was placed approximately 388 m from the met mast and the western LiDAR is placed approximately 280 m from the met mast in the horizontal direction.

Both LiDARs' position was measured manually. The location of the northern LiDAR is approximately 5928611 meters north, 472084 meters east, and 8 meters above sea level. The location of the western LiDAR is approximately 5928200 meters north, 471731 meters east, and 5 meters above sea level (WGS84 UTM coordinate system).



Figure 3: LiDARs setup with respect to the met mast. Scanning patterns are qualitatively represented in blue lines.

Scan		A_TNO	B_TNO	C_TNO
scan type		PPI	PPI	PPI
ToF (Time of Flight)		direct	indirect	direct
repetitions	(#)	1	1	1
azimuth	start(deg)	81,67	111,67	81,67
	stop(deg)	111,67	81,67	111,67
	speed(deg/s)	2	2	2
elevation	start(deg)	9,0	11,0	13,0
sample time	(ms)	1000	1000	1000
distance	start(m)	100	100	100
	number(#)	10	10	10
	step size(m)	50	50	50
range gate	(m)	50	50	50

Table 1: West positioned LiDAR scanning patterns specification. Three scanning patterns are presented.

The third, staring³ LiDAR is positioned next to the western scanning LiDAR. The initial goal was to compare the measurements of both mentioned LiDARs so that the impact of scanning when measuring radial wind speed could be assessed. The measurements recovered from the staring LiDAR proved not useful for the analysis, since a post-verification hard target mapping evidenced a high uncertainty in the elevation angle ($>0.15\text{deg}$).

For this work, measurements from the cup anemometer at 55 meters and wind vane at 110 meters are used. The relative positions of the three with relation to the western LiDAR scanning pattern can be observed in Figure 4, where yellow lines indicate the scanning pattern around the 55-meter boom.

The boom lengths at 55m and the sensor specifications are detailed in Figure 5. On the left, the cup anemometer, a Thies First Class 4.3351.00.000, and on the right is the sonic anemometer, a Gill WindMaster Pro 1590-PK-020. The cup is positioned

³a LiDAR operating in a staring mode will measure the radial wind speed at only a specific point in space.

Scan		A_IWES	B_IWES	C_IWES
scan type		PPI	PPI	PPI
ToF (Time of Flight)		direct	indirect	direct
repetitions	(#)	1	1	1
azimuth	start(deg)	194,66	164,66	194,66
	stop(deg)	164,66	194,66	164,66
	speed(deg/s)	2	2	2
elevation	start(deg)	5,9	7,9	9,9
sample time	(ms)	1000	1000	1000
distance	start(m)	100	100	100
	number(#)	10	10	10
	step size(m)	50	50	50
range gate	(m)	50	50	50

Table 2: North positioned LiDAR scanning patterns specification. Three scanning patterns are presented.

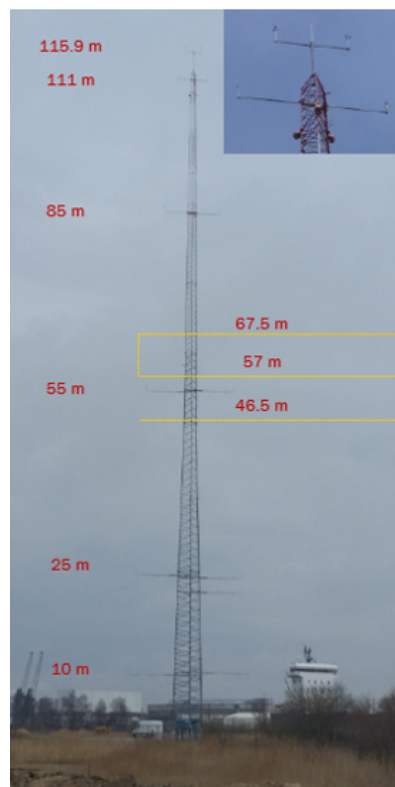


Figure 4: 115m IEC conform met mast and western LiDAR scanning pattern with relation to the existing instrumentation (yellow line). Image copyright Fraunhofer IWES.

at the north-west end, with an orientation of $301,95^\circ$ and at an exact height of 55,50 meters. The sonic is positioned at the southeast end, with an orientation of $121,84^\circ$ and at an exact height of 55,07 meters.

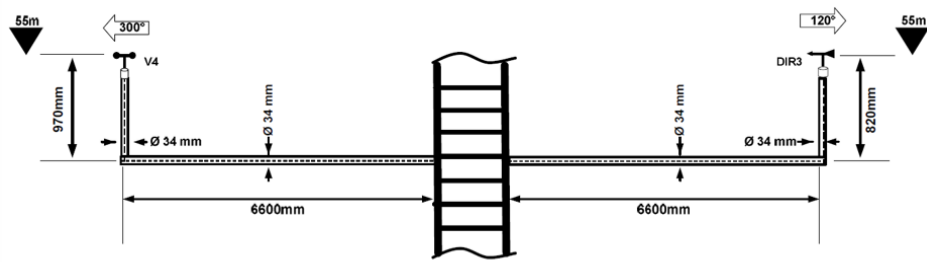


Figure 5: Detailed metmast boom at 55 meters and sensors specification. Image copyright Fraunhofer IWES.

The position of the sonic anemometer was mapped onto the west LiDAR by using the hard target detection method^a. This can be seen in Figure 6. The location of the sonic anemometer, with relation to the western LiDAR, using the LiDAR's spherical coordinate system is the following:

- **Azimuth, ϕ_{sonic} :** 87.67 degrees (with relation to the north).
- **Elevation, θ_{sonic} :** 10.6 degrees (with relation to the ground).
- **Range, R_{sonic} :** 300 meters.

^aThe hard target detection method maps the location of the anemometer by identifying the reflection of the LiDAR's pulse from the solid structure/surface through the backscatter signals and associated CNR value.

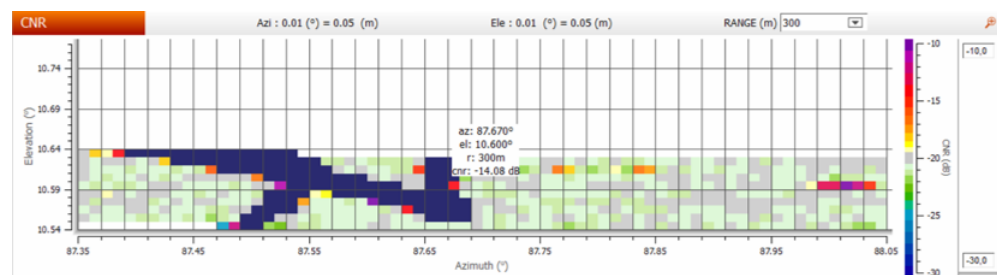


Figure 6: Sonic anemometer hard targeting using the western LiDAR's software.

The difference in height between the sonic anemometer and western LiDAR was calculated to be 52.60 meters and the horizontal distance between the western LiDAR and the sonic anemometer to be 292.23 meters.

The cup anemometer was not subject to hard targeting, but its location is expected to be the following:

- **Azimuth, ϕ_{sonic} :** 86.44 degrees (with relation to the north).
- **Elevation, θ_{sonic} :** 10.69 degrees (with relation to the ground).

The difference in height between the cup anemometer and western LiDAR was calculated to be 52.88 meters and the horizontal distance between the western LiDAR and cup anemometer to be 280.23 meters.

2.2 Data acquisition and visualisation

The scanning patterns described in section 2.1 are plotted in Figures 8 and 7. On the lower-left plot, the points where the radial wind speed is registered have been scattered. The middle line, i.e., the points that were taken at the middle elevation angle, appear to be slightly sifted, when compared to the upper and lower elevation angles. This is explained by the averaging of the radial wind speeds throughout a 2-degree sector, and such averaging is then matched to the last point of the same 2-degree sector.

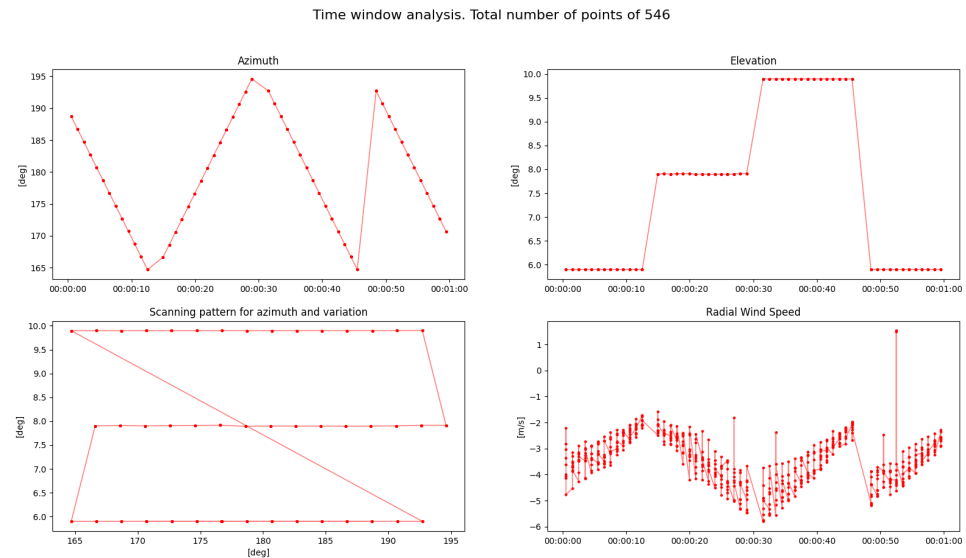


Figure 7: Western Windcube 200s scanning LiDAR pattern. On the two upper plots, the azimuth and elevation patterns are shown, following the PPI patterns specified in Table 2.

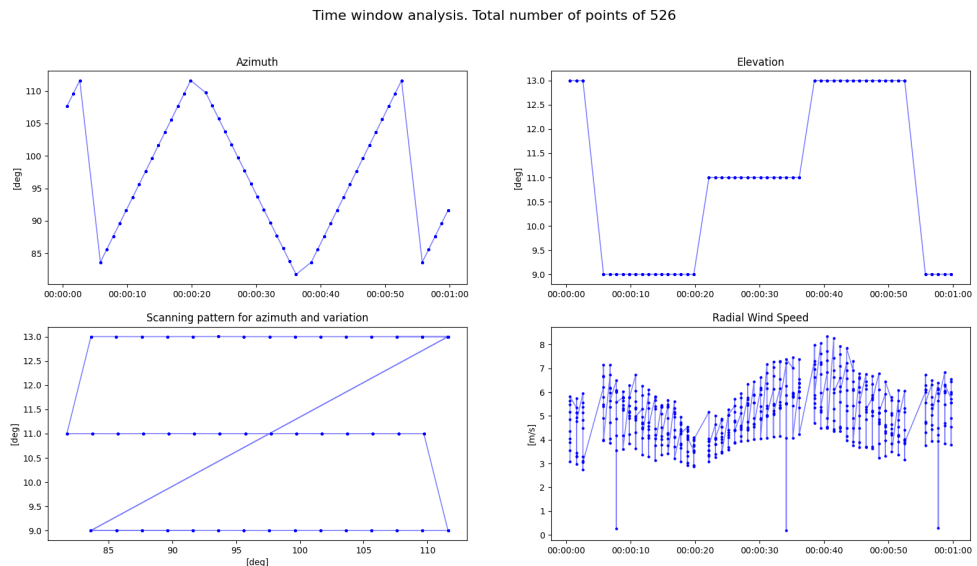


Figure 8: Western Windcube 200s scanning LiDAR pattern. On the two upper plots, the azimuth and elevation patterns are shown, following the PPI patterns specified in Table 1.

The points taken by both LiDARs are converted to a Cartesian coordinate system by making use of a simple coordinate transformation. The coordinates of the points taken by the north positioned LiDAR are written with relation to the western LiDAR.

Figure 9 puts in evidence the overlapping are covered by both LiDARs⁴.

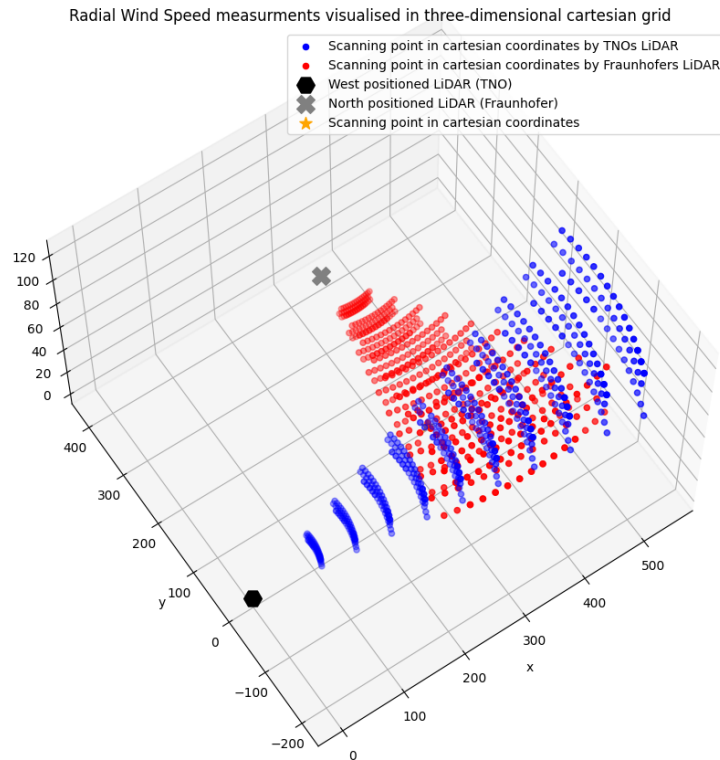


Figure 9: Radial wind speed measurement locations by both north and western LiDARs when written in the same cartesian reference system. The overlap of both measurements can be clearly seen.

The data measured by the cup anemometer and wind vane have been combined into the wind rose for the complete duration of the campaign, which can be seen in Figure 10. The most common wind direction was approximate to West-Southwest (WSW) (247.50°). For all wind directions, the most common range of wind speeds was between 5.6 and 11.2 m/s.

2.3 Baseline result and scanning pattern impact

An initial exercise of interest was to compare the measurements of the cup anemometer and the LiDAR measurements themselves, i.e., without any interpolation by the Gaussian Processes. Only the point closest to the cup anemometer registered by the LiDAR is chosen for this comparison.

The LiDAR passes through this point once every scanning pattern (which takes roughly 50 seconds), meaning that for a fair comparison with the cup, the closest timestamp of this point (registered by the LiDAR) is matched with high-frequency measurements at the cup (registered by the anemometer itself). This point is taken at the gate range closest to the cup, i.e., at 300 meters. Figure 11 illustrates the point closest to the cup anemometer that can be used for the aforementioned exercise (middle elevation on the right to the dark red colored point).

Notice that the point closest registers a high CNR value when compared to the others (higher than 0 dB). This is because in the azimuthal range the LiDAR encounters the

⁴The usage of a Cartesian coordinate system (x,y,z) to identify the points where the radial wind speed is taken is an important step as the GPs require information to be encoded in this fashion.

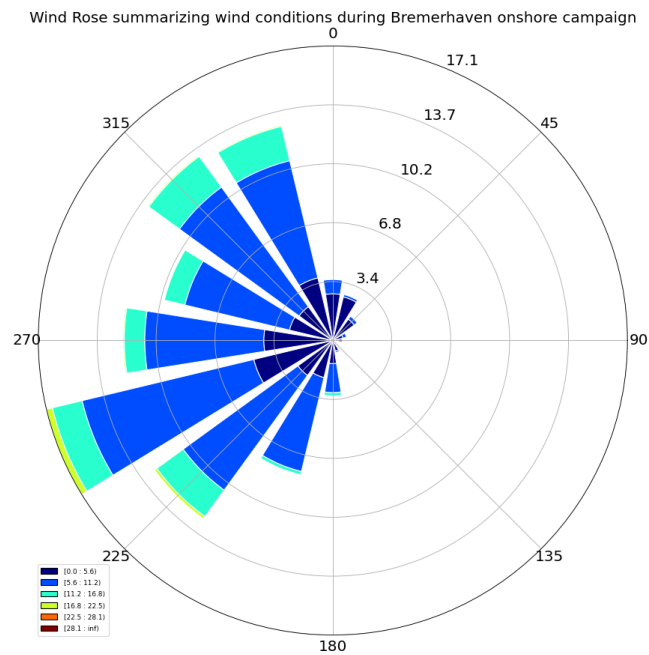


Figure 10: Wind rose representing most common wind speeds and direction measured during the onshore campaign.

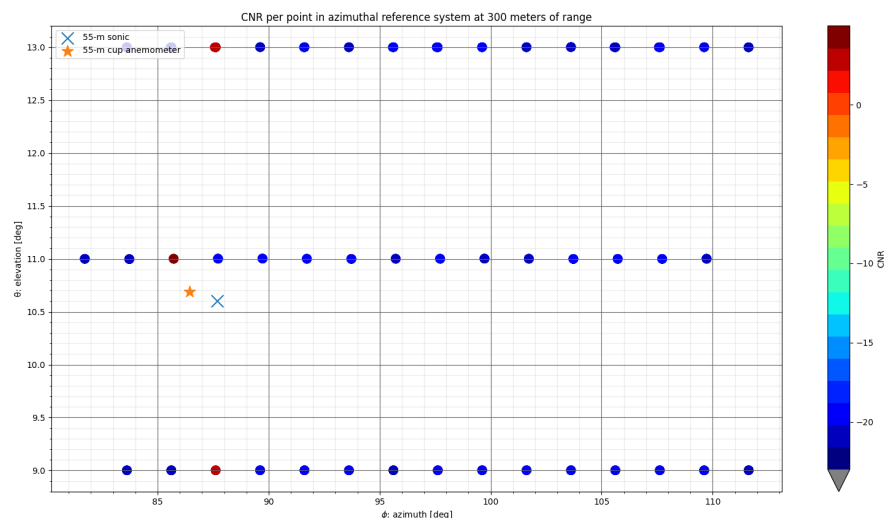


Figure 11: Carrier to Noise Ratio, in dB, measurements from the western LiDAR at a range of 300 meters represented in a grid with polar coordinates. Scattered measurements from a time period of 10 minutes.

met mast. Therefore, the point nearest to the right is used for this analysis.

Figure 12 shows a comparison between the cup anemometer resolved to the LiDAR beam (orange)⁵ and the LiDAR value at the closest time instant (blue). The result of the linear regression is a coefficient of determination of 0.951 and linear relation of $y = 0.98x - 0.25$, as seen in Figure 13. Note that the frequency of the analysed time series differs from 1Hz because only one point over the scanning pattern is analysed,

⁵This procedure is further elaborated in section 3. Given that the cup anemometer was used, the vertical component of the wind velocity is disregarded.

resolving to a time series of approximately 0,02 Hz.

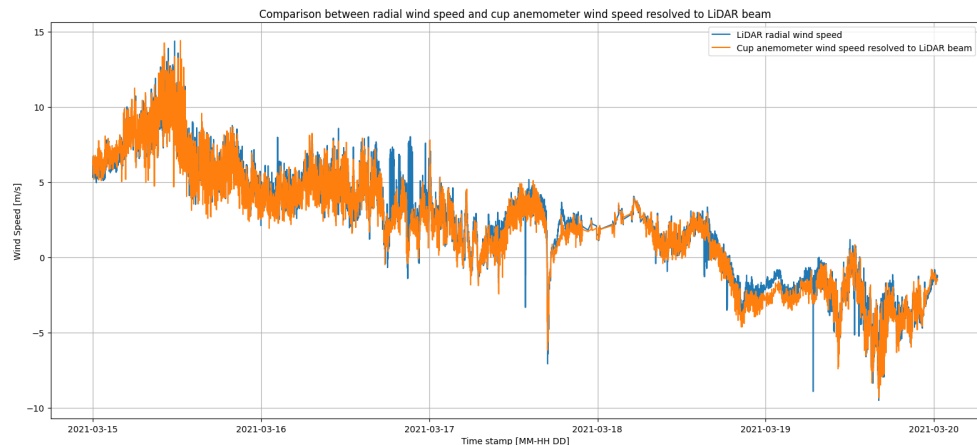


Figure 12: Comparison between LiDAR radial wind speed, taken at the point closest to the cup and sonic anemometers (blue) and cup anemometer wind speed resolved to LiDAR beam (orange).

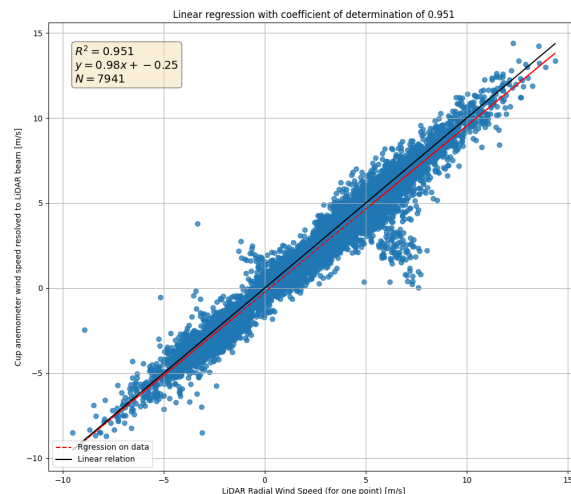


Figure 13: Linear regression between LiDAR radial wind speed, taken at the point closest to the cup and sonic anemometers (x-axis) and cup anemometer wind speed resolved to LiDAR beam (y-axis).

Impact of the scanning pattern is not a major one ^a

^aThis is valid for the exercise done where the interest lays in mean wind speed. For other timescales or higher order quantities this conclusion may not hold.

From the presented test, it can be concluded that the impact of the scanning pattern is not major, in the sense that although the radial wind speed is averaged through a range of 2 degrees (and only then registered at the last point of the latter range), the measurements of the LiDAR tend to match to the ones registered by the cup anemometer.

This value serves also as a baseline for comparison with the results from applying the GPs. In other words, as the GPs are fitted onto the LiDAR data, an exercise of interest would be to compare the value obtained with the regression with the ones that will be obtained when using the Gaussian Processes.

3 Application of Gaussian Processes

3.1 Goal and methodology

The final goal is to understand and quantify the accuracy of the GP interpolation by performing a regression analysis between the cup anemometer⁶ wind speed projected to the LiDAR's LOS and the GPs prediction of radial wind speed, for the same time stamps existing in the met mast dataset and at the cup location.

The cup anemometer does not allow to have a three-dimensional wind velocity. Therefore, the wind velocity vector $\mathbf{V}$ represents the horizontal wind velocity. The vertical component of the wind is hereby assumed to be 0. The decomposition of $\mathbf{V}$ in its u and v components translates into equation 3.1.

$$\begin{aligned} u &= -|\mathbf{V}|\sin(WD) \\ v &= -|\mathbf{V}|\cos(WD) \end{aligned} \quad (3.1)$$

where u is the velocity component aligned with the defined x axis (easterly), v is the velocity component aligned with the defined y axis (northerly), $|\mathbf{V}|$ is the wind magnitude and WD is the wind direction measured with relation to the north, positive clockwise.

In a second step, the u and v components are projected onto the LiDAR LOS according to equation 3.2. Note that by using the cup anemometer data the w component of the wind speed is not projected to the LiDARs LOS. The terrain at the test field is not a complex one, therefore although considering the w component would lead to more precise results, it is not thought to be a source of major influence for this analysis. Future work will focus on evaluating the same comparisons by making use of the sonic complete wind field time series.

$$v_{\text{resolved}} = u\sin(\phi)\cos(\theta) + v\cos(\phi)\cos(\theta) \quad (3.2)$$

where v_{resolved} is the cup anemometer velocity resolved to the LiDAR beam, ϕ is the azimuth angle, measured from the north to the location of the point being measured and θ is the elevation angle, measured from the ground to the LiDARs LOS.

The GPs are fitted according to the methodology specified in Section 1.3.

The GPs were in the first time fitted and analysed in a short period (5 days) and in a second time the regression was run for the complete time window. The results of these analyses follow in the section below.

3.2 Results

Ideally, the GPs should be able to provide accurate radial wind speed estimations for different atmospheric conditions with incoming raw data provided by the LiDAR.

With this goal in mind, minimum filtering of the datasets was first applied. The minimum filtering aims to remove radial wind speed measurements that are not accurate.

⁶The cup anemometer was first used for the analysis due to being available before the sonic anemometer data.

This can be, for example, due to internal faulty processes from the LiDAR or due to lack of signal quality. The two can be evaluated by status flags registered by the LiDAR or by the Carrier to Noise ratio (CNR).

The incoming data provided in *netcdf* files was first imported to TNO's database and only then queried to be used for the GPs fitting. Three conditions are used in the query regarding the filtering of data.

- **CNR lower threshold:** CNR values higher than -23 dB were chosen, thus (theoretically) ensuring points where the signal-to-noise ratio is sufficiently high.
- **CNR upper threshold:** CNR values lower than -3dB was chosen, thus (theoretically) eliminating all points where the LiDAR encounters solid targets (met mast).
- **Validity of information:** Radial wind speed status must be true, ensuring the measurement of radial wind speed is valid.

The results concerning the application of the GPs are organized as follows:

1. **Short time window analysis without additional filtering:** The GPs are applied to LiDAR measurements with the three filter options aforementioned. The first test is done for a short period of 5 days. The comparison between predicted and measured values is evaluated through a linear fitting of the two, and compared based on the coefficients of the linear relation and coefficient of determination R^2 . The latter is computed as $1 - \frac{RSS}{TSS}$, where RSS is the sum square of residuals and TSS is the total sum of squares. The R^2 value serves as a proxy for the strength of the linear relation. Residual analysis is performed to investigate possible biases of the model.
2. **Long time window analysis without additional filtering:** The same analysis as in (1) is performed, but now for the complete campaign.
3. **Short time window analysis with additional filtering of radial wind speeds:** A statistical filter to further filter radial wind speed measurements is tested onto a short period (the same period as in (1)), on top of the conditions already used to initially filter the data. The performance of the model is again assessed via linear regression and residual analysis.
4. **Investigation of outliers:** The outliers identified and filtered in (3) are investigated to better understand their origin, both in space and time.
5. **Investigation of filters:** A different approach to filtering is tested. This approach uses the CNR values as opposed to directly filtering the radial wind speed measurements.

3.2.1 *Short time window analysis without additional filtering*

The first time window under investigation covers a wide range of conditions, thus evaluating the GPs performance over these.

The wind speeds measured by the cup anemometer during the 5-day time window, from 15 to 25th of March, covered a wide range of values, from 0 to approximately 17.5 m/s, as shown in Figure 14. Variations in the wind speed appear higher during the first 2 days when compared to the last 3 days.

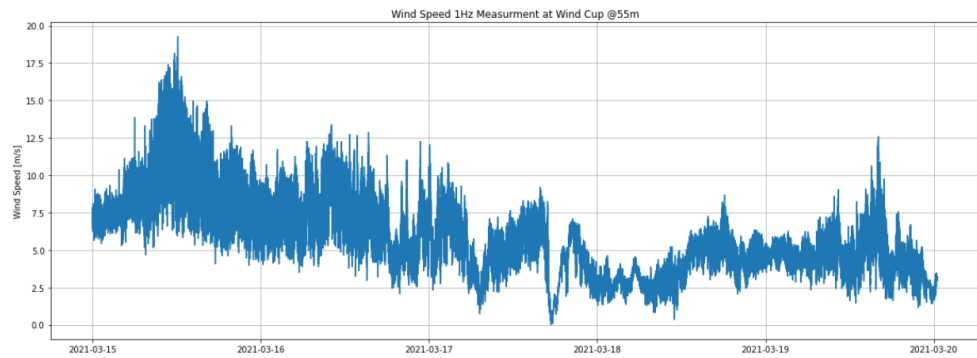


Figure 14: Wind Speed values measured by the cup anemometer throughout analysed time window.

The wind direction ranges from approximately 250 degrees to 150 degrees, as shown in Figure 15 with the majority of directions being concentrated around 300 to 50 degrees in the analysed time window.

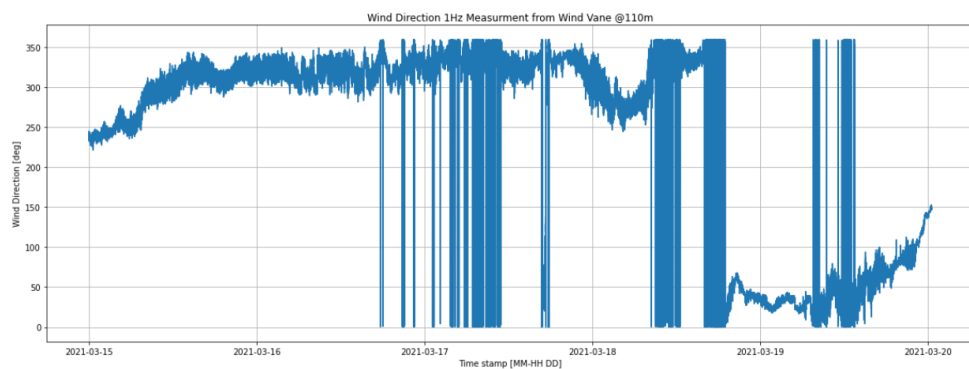


Figure 15: Wind direction values measured by the wind vane throughout analysed time window.

For the GPs, the LiDAR is used, i.e., no specific points within a certain range or azimuth are chosen, as visible in Figure 16. The radial wind speeds measured at each time instant cover a wide range of values due to the multiple elevations (meter wise), recorded at each instant.

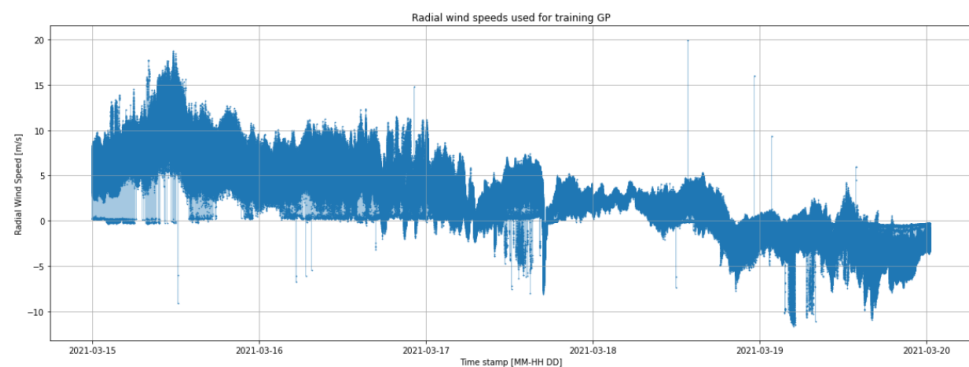


Figure 16: Radial wind speed measured by the LiDAR during the time window being analysed and used for the training of the GP. Notice that at some instants the measurements tend to deviate from the main trend and drop to very low values, such as on the 19th of March, which is referred to within the context of this work as the outliers.

The radial wind speed prediction is compared to the cup anemometer wind speed resolved to the LiDAR LOS. This can be observed in Figure 17, where the former is

plotted in blue and the latter in orange. It can be seen that, roughly, the GP predictions follow the main trend of the cup data.

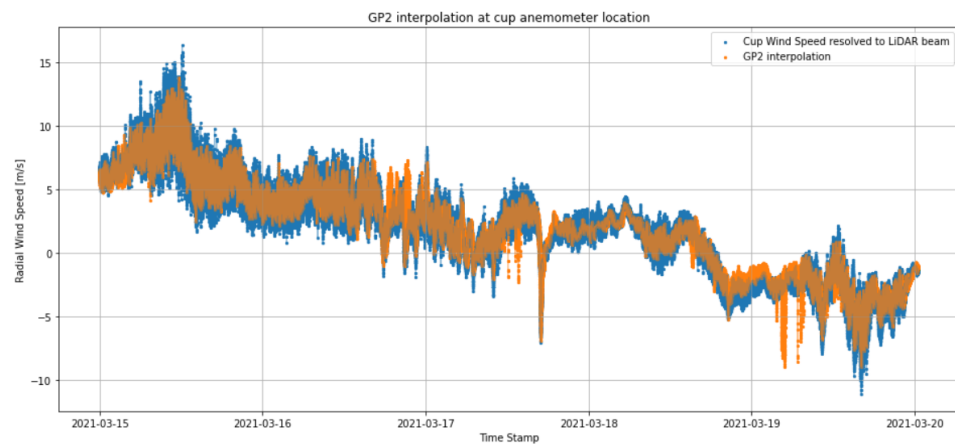


Figure 17: Comparison between cup anemometer wind speed resolved to LiDAR LOS (blue) and prediction yielded by the GP (orange).

Measured and predicted values by GP are in accordance

To further compare the two, a regression analysis is done, as shown in Figure 18. A coefficient of determination R^2 of 0.959 is found, with a linear relation of $y = 1.03x - 0.17$ indicating that the GP can provide accurate interpolations of the radial wind speed during time instants where the LiDAR is scanning further away from the point where the radial wind speed is evaluated. Note that given the high-frequency dataset of 1Hz (matched by the frequency of the predictions of the GP) the regression is evaluated for a total of approximately 400 000 points.

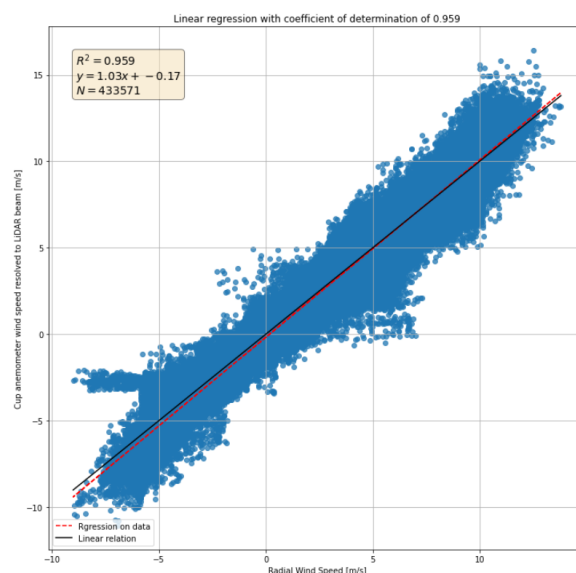


Figure 18: Linear regression analysis between cup anemometer wind speed resolved to LiDAR LOS (y axis) and radial wind speed predicted by GP (x-axis). 1 Hz time series.

When comparing 10-minute statistics, both radial wind speeds, i.e., the 10-minute statistics of the 1Hz time series interpolated by the GP and the 10-minute statistics of

the cup speed resolved to the LiDAR match, as seen in Figure 19 by evaluating the high regression coefficient of 0.99 and linear fit of $y = 1.03x - 0.19$.

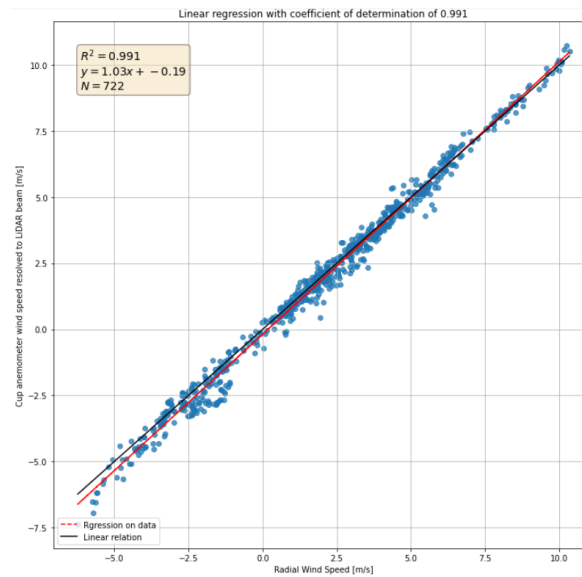


Figure 19: Linear regression analysis between cup anemometer resolved to LiDAR beam time series evaluated in 10 minute statistics and GP radial wind speed predictions time series evaluated in terms of 10 minute statistics.

The regression coefficient provides insight into the predictability of one variable as a function of another. Nevertheless, the quality of predictions yielded by a certain model can also be assessed by evaluating the residuals, i.e., the difference between the real observation (wind cup measured by the cup and resolved to the LiDAR beam) and the prediction yielded by the model (the Gaussian Process). Ideally, they would follow a normal distribution and show no particular patterns with the observations, otherwise, it can be the case that the model cannot explain some of the behavior seen in the data.

The upper plot in Figure 20 shows that the residuals do not follow a normal distribution. It would be expected that the blue points in the probabilistic plot would follow the red line if that were the case, i.e., that the relationship between the sample percentiles and the theoretical ones (of a normal distribution) would follow a linear relation. The nonlinear relation towards the beginning and end (upper and lower percentiles) points out that the residuals are not normally distributed, which can be partly explained by the existence of outliers in the predictions (this will be further discussed).

Residuals are not randomly distributed

The lower plots in Figure 20 show that the residuals are not randomly distributed across the observations. For example, when comparing the residuals and the observations, it can be slightly observed that for high and low radial wind speeds the residuals are higher when compared to radial wind speeds in the middle. A skewed clustered shape in green can be noticed, although not too pronounced.

Notice that the remaining outliers are due to the spikes in the predictions yielded by the GP for a specific time window. This is more thoroughly discussed in subsections 3.2.3 and 3.2.4. It may be possible that for periods of high and low wind speeds or certain wind directions the GP may not be able to fully encode the wind behavior,

although the trends found to require further investigation and discussion, which is out of the scope of this work.

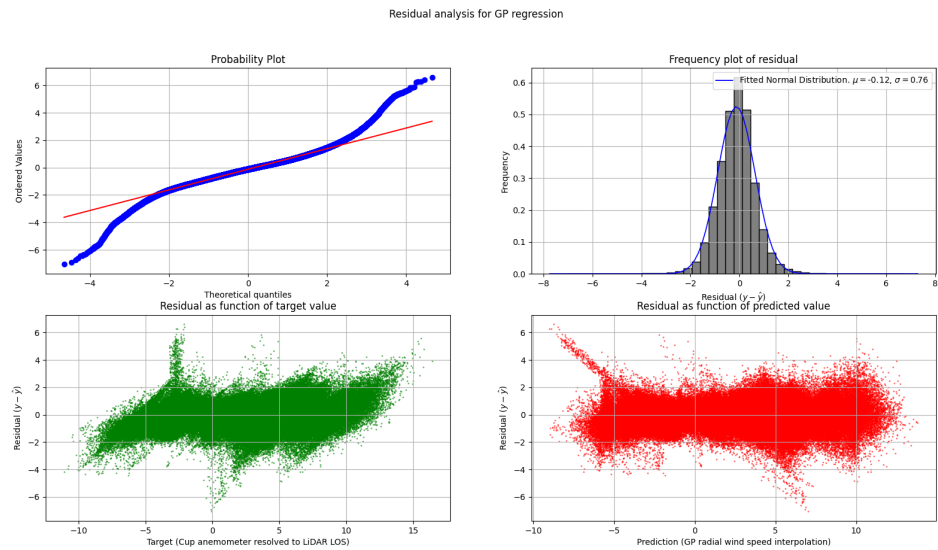


Figure 20: Residual analysis for GP predictions. Checking for normal distribution of residuals by evaluating probabilistic and histogram plot and evaluation of residuals with relation to real and predicted observations.

The plots in Figure 21 further show the relations between the residuals and the wind speed and wind direction (lower plots in orange). It can again be seen that the residuals are not randomly distributed for all wind speeds and directions.

The upper plot in black further reveals some interesting insights: it can be seen that the residuals are much lower for lower standard deviations of the predictions when compared to higher ones. Although a clear-cut linear relation does not exist, this is evidence that the uncertainty yielded by the GP embeds the actual deviation between the predicted and real radial wind speed.

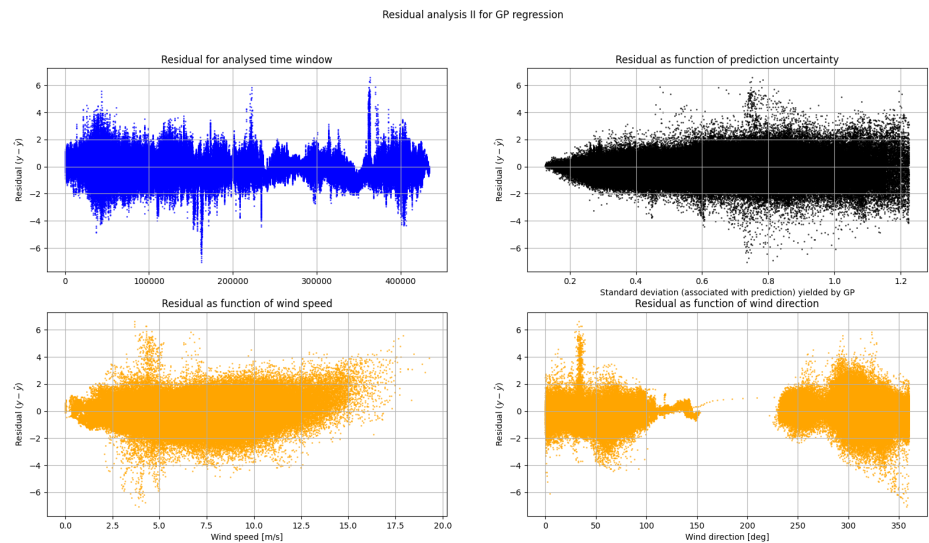


Figure 21: Residual analysis (continuation). Variation of residuals through the analysed time window (upper left). Relation between residuals and uncertainty yielded and associated by GP prediction (upper right). Analysis of the dependence of residuals with wind speed and wind direction (lower left and right, respectively).

3.2.2 Long time window analysis without additional filtering

The GPs were then applied to the complete dataset. At certain time instants, it was observed that data was missing from the LiDAR. More specifically, a gap on the 30th of March between 13:45 and 14:15 (approximately). This was then removed for the analysis, as using the GPs to obtain a prediction in a time window where the complete data is absent generates issues.

The results for the complete campaign are presented in Figures 22,23 and 24, where the comparison between the 2 usual time series is presented, along with the linear regression analysis for the high-frequency time series and 10-minute statistics, respectively.

Accurate predictions by GPs for the complete campaign

The comparison between the prediction and true value, similarly to the short time window analysis, shows that the prediction follows the main trend of the cup data (apart from spikes in the predictions for a specific period). A coefficient of determination of 0.934 is seen when comparing both high-frequency time series (prediction and cup anemometer). The linear relation is $y = 1.01x + 0.08$ indicating a good model fitness. This regression involved approximately 2 million data points. The regression involving low-frequency time series (10-minute statistics), shows a coefficient of determination of 0.979 and a linear relation of $y = 1.01x + 0.05$.

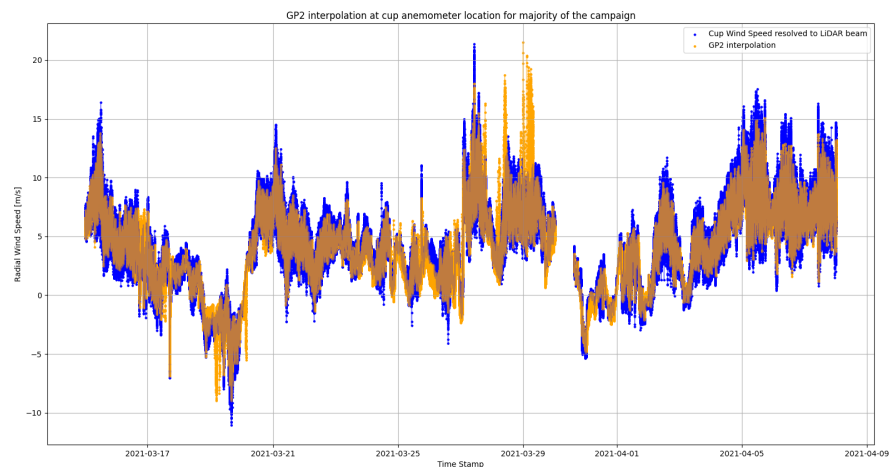


Figure 22: Comparison between cup anemometer wind speed resolved to LiDAR LOS (blue) and prediction yielded by the GP (orange) for the whole campaign.

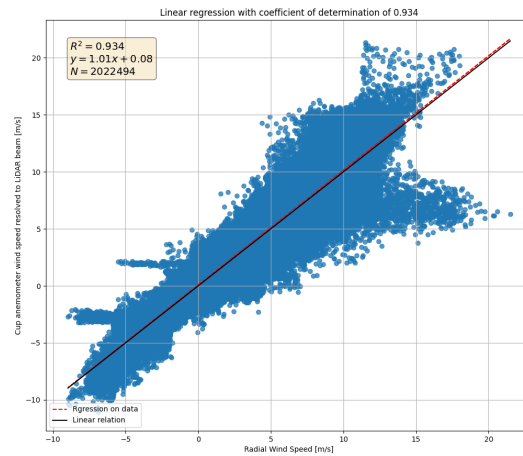


Figure 23: Linear regression analysis between cup anemometer wind speed resolved to LiDAR LOS (y axis) and radial wind speed predicted by GP (x-axis) for the complete campaign.

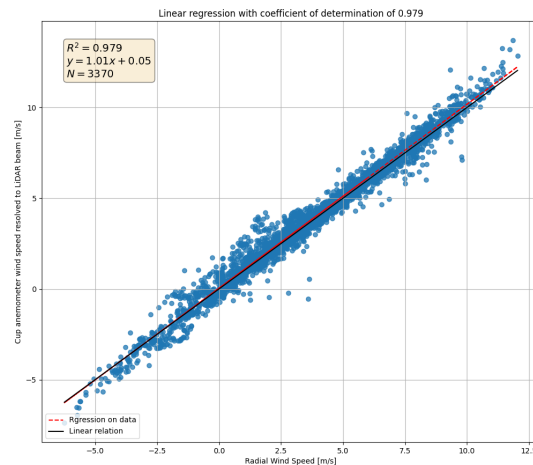


Figure 24: Linear regression analysis between cup anemometer resolved to LiDAR beam time series evaluated in 10 minute statistics and GP radial wind speed predictions time series evaluated in terms of 10 minute statistics for the complete campaign.

3.2.3 Short time window analysis with additional filtering of radial wind speeds

As noticed in the regression plots in subsection 3.2.1 there is evidence of unfiltered outliers in the LiDAR RWS measurements. In the regression analysis, these are easily observed, placed far from the linear relation diagonal. These points have been identified to originate from spikes in the predictions yielded by the GPs at certain time instants. This is especially visible on the 19th of March, where sudden drops of radial wind speed yielded by the GP are seen, as Figure 25 puts in evidence.

Efforts were put to understand and identify the nature and origin of these outliers (see section 3.2.4)

A statistical filter was used to remove such points from the radial wind speed time series, more specifically, a Hampel filter, according to the implementation in [6]. The Hampel filter detects the existence of outliers by using a sliding window of configurable size which goes over the data. Data points that are outside a certain threshold are identified as outliers. The threshold is also configurable and corresponds to a multiple of the standard deviation for a certain time window.

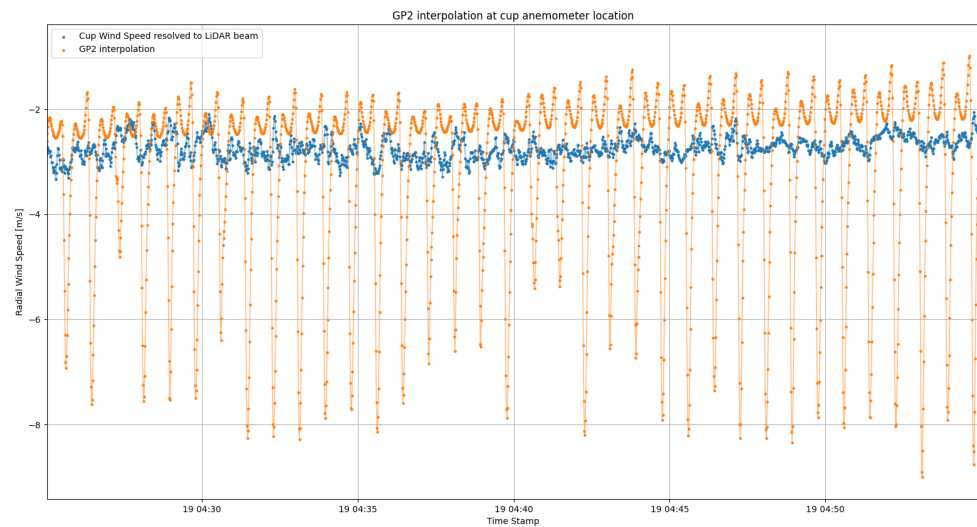


Figure 25: Detailed spikes in the radial wind speed predictions yielded by the GP, showing sudden drops for specific time instants.

The standard deviation is estimated for a certain time window by scaling the Median Absolute Deviation⁷ (MAD) by a constant factor. The constant factor is equal to 1.4826 under the assumption that the data follows a normal distribution. Other examples in the literature where the authors use some kind of statistical process to remove faulty measurements from a scanning LiDAR exist [7].

The short time window is used again, with the difference that the Hampel filter is first applied. The window's length is set to 2000 data points (which roughly corresponds to 4 scanning patterns) and a standard deviation of 2.

As a consequence, the GP is trained on a new set of radial wind speeds. Figure 26 shows the radial wind speed time series obtained from the LiDAR before and after applying the Hampel filter.

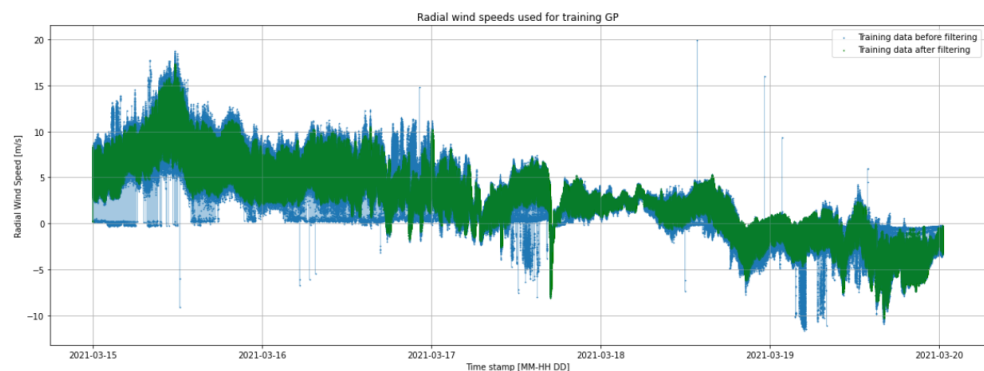


Figure 26: Comparison of radial wind speed time series used to fit the GPs before (blue) and after the application of the Hampel filter (green).

⁷The Median Absolute Deviation is defined as the median of the absolute deviations (from the existing observations within a certain data set) from the data's median.)

Small improvement of GPs prediction with filtered time series

The predictions yielded by the GPs when trained on the filtered time series differ slightly from the predictions yielded by the GPs fitted onto the unfiltered time series, as Figure 27 exhibits and Figure 28 in more detail. The R-squared value of the linear fit does not appear to change much, increasing from 0.959 to 0.962 as shown in Figure 29. This is somehow expected as the outliers observed in Figure 18 where not many and the majority of the observations laid already in the linear relation.

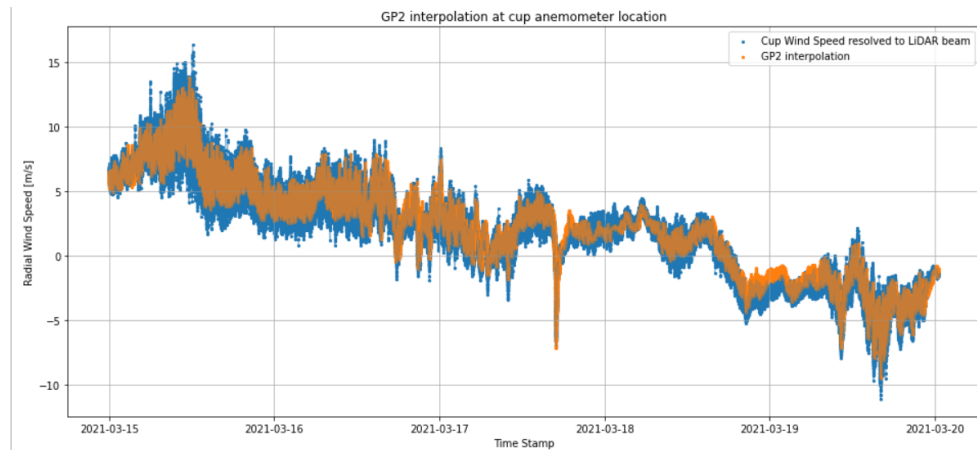


Figure 27: Comparison between cup anemometer wind speed resolved to LiDAR LOS (blue) and prediction yielded by the GP (orange) when fitted onto the radial wind speed filtered time series.

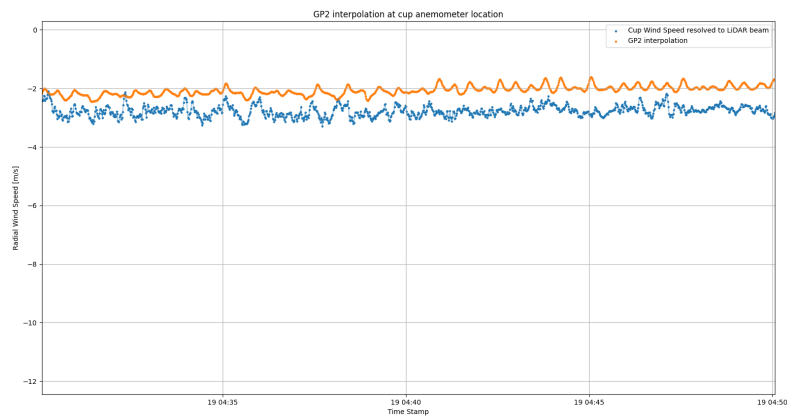


Figure 28: Detailed spikes in the radial wind speed predictions yielded by the GP, showing sudden drops for specific time instants.

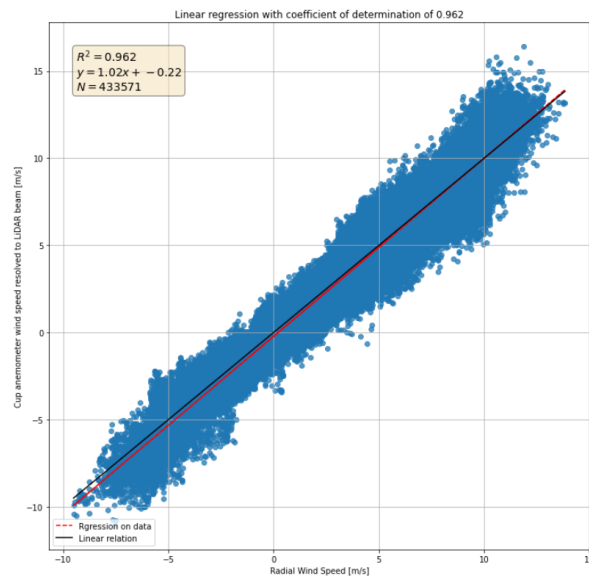


Figure 29: Linear regression analysis between cup anemometer wind speed resolved to LiDAR LOS (y axis) and radial wind speed predicted by GP (x-axis) for the GP fitted onto filtered time series.

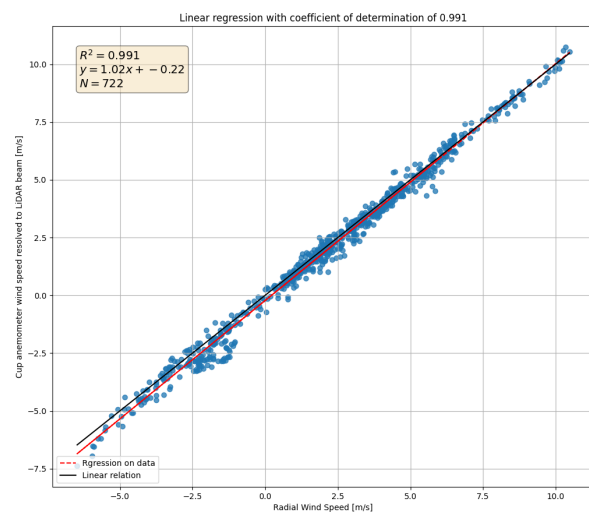


Figure 30: Linear regression analysis between cup anemometer resolved to LiDAR beam time series evaluated in 10-minute statistics and GP radial wind speed predictions time-series evaluated in terms of 10-minute statistics for the filtered time window.

The residual analysis is also presented for the filtered dataset in Figure 31. the only difference is that the patterns created by the outliers are not present anymore. The residuals appear to better follow a normal distribution by analysis of the probabilistic plot, with a slight skewness higher residuals towards high radial wind speeds, although it is fairly small.

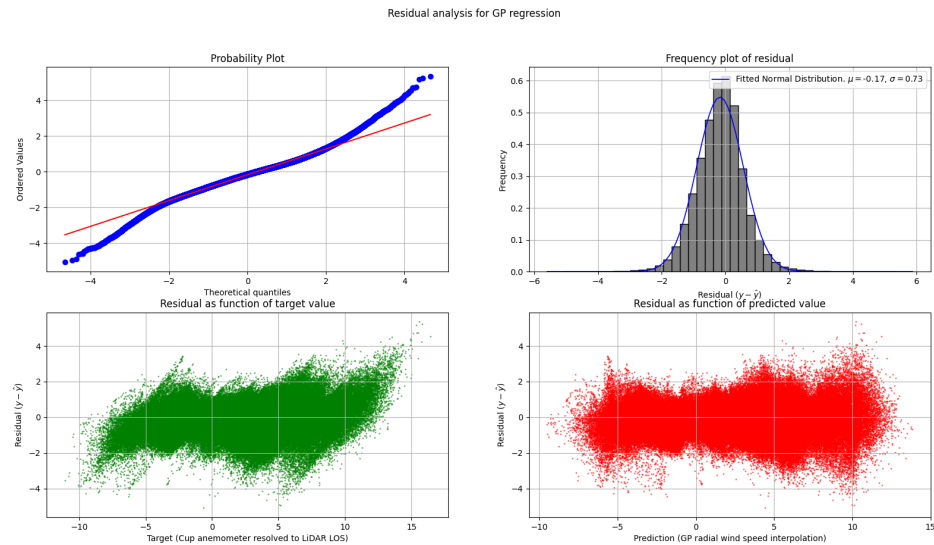


Figure 31: Residual analysis for GP predictions. Checking for normal distribution of residuals by evaluating probabilistic and histogram plot and evaluation of residuals with relation to real and predicted observations.

3.2.4 Investigation of outliers

In this section, the information that has been collected on the detected outliers coming from both the TNO and the Fraunhofer LiDAR is presented. The origin of the existing outliers is thought to be a combination of several factors: laser backscatter from solid targets not correctly filtered out through CNR levels identification, backlash from the LiDAR pattern change of direction, un-recorded atmospheric conditions, and possibly more.

Two types of outliers within the LiDARs RWS measurements, both the northern and the western ones, were observed. The first type of outliers (type A) seemed defined by sudden drops to 0 of the RWS (Figure 32(a)). The second type (type B) is characterised by either sudden drops or surges of the RWS to seemingly unpredictable values (Figure 32(b)). Type A, while still being identified and removed, did not show to have a recordable impact on the GP training and prediction. Type B, on the contrary, had a clear influence on the prediction and had to be removed (see section 3.2.3). The main difference identified between Type A and type B outliers seemed to be appearance frequency: Type B outliers were observed to appear at a frequency of $f_{TypeB} > 1 [s^{-1}]$. On the contrary, type A outliers appearance frequency was not seen to exceed $tdf_{TypeA} < 0.1 [s^{-1}]$. (see Figure 32(a and b)).

Figures 33 and 33 are visualisations of the outliers' appearance within the data coming from the western (TNO) and northerner (Fraunhofer) LiDAR. Their location (spatial), appearance frequency (count), and strength (distance from median value) can be observed. The outliers tend to appear in the vicinity of the meteorological mast, and at the outer edges of the scanning pattern. Type B outlier, for both the northern and western LiDAR, seems to be characterised by strong offsets from the local median RWS. Furthermore, the highest (13 deg) scanning row seems to be the main source of these outliers. This characteristic is shared with Type A outliers, also more densely located in the higher scanning path. Type A also tends to be located around the meteorological mast, and it is at that very location that the highest values are recorded. Finally, it is possible to observe that even for periods where there are no visibly noticeable outliers, some are still detected, and are again characterised by being in the vicinity with the meteorological mast.

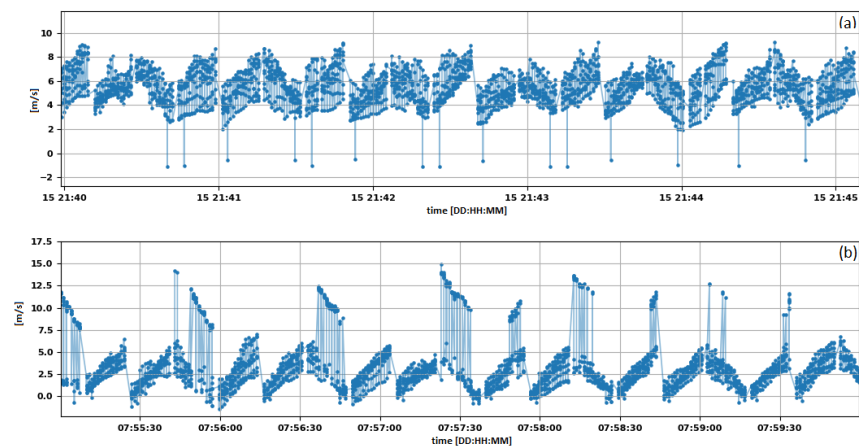


Figure 32: (a): Type A outliers. (b): Type B outliers.

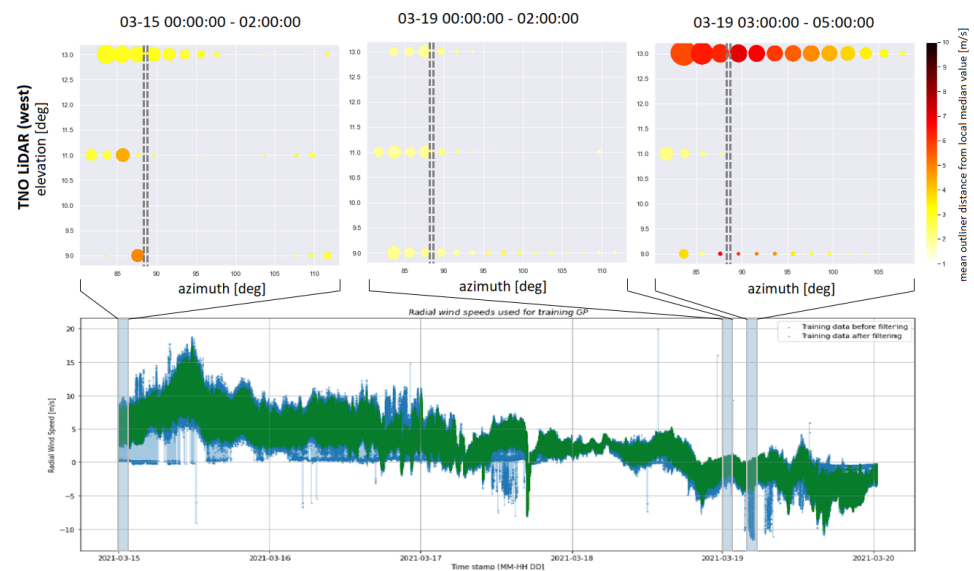


Figure 33: Spatial map of outlier appearance from the western LiDAR measurements. For each combination of azimuth (x) and elevation (y), there are 10 measurements ranges [100, 150 ... 550]m. In this analysis, the influence of the measurement range is not taken into account. The scatter size is an indicator of the number of outliers detected, (low count indicates few outliers). The scatter color is an indicator of the mean distance from the local median *RWS* value in [m/s]. For more detail see section 3.2.3.

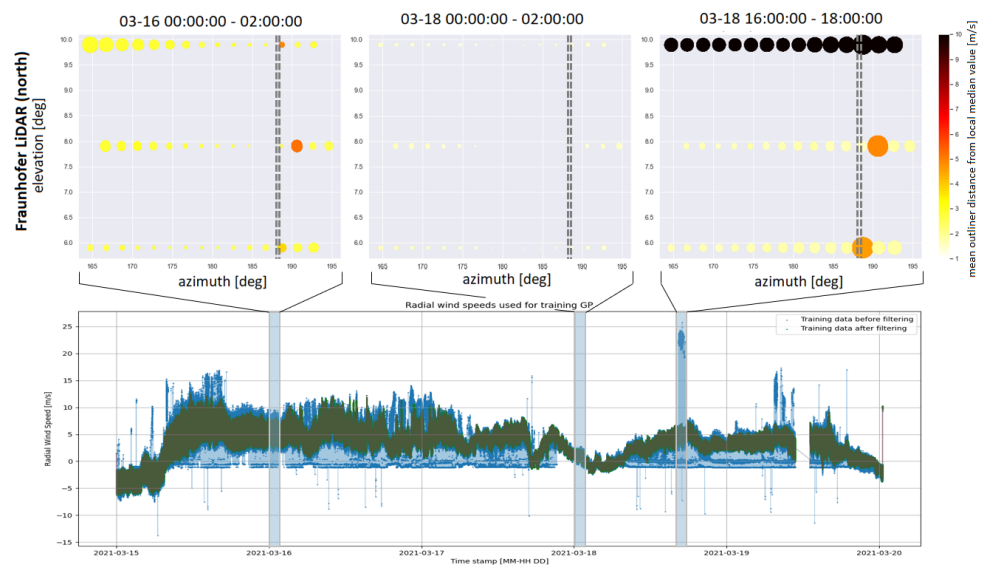


Figure 34: Spatial map of outlier appearance for the northerner LiDAR measurements. For each combination of azimuth (x) and elevation (y), there are 10 measurements ranges [100, 150 ... 550]m. In this analysis, the influence of the measurement range is not taken into account. The scatter size is an indicator of the number of outliers detected, (low count indicates few outliers). The scatter color is an indicator of the mean distance from the local median *RWS* value in [m/s]. For more detail see section 3.2.3.

3.2.5 Investigation of filters

It is understandable that directly filtering radial wind speeds may not be the most desirable method, and it may be more suited to use some other variable to truly assess the quality of the radial wind speed measurement and decide if it is suited for the GP fitting.

Therefore, the introduced Hampel filter is slightly modified: the outliers are identified based on the standard deviation estimation of the CNR measurements, meaning that observations of radial wind speed with statistically abnormal CNR are removed from the time series.

If the filter can detect and remove the existing outliers, the hypothesis that the outliers are accompanied by statistically abnormal measurements of CNR can be corroborated. The resulting comparison between filtered and unfiltered data is represented in Figure 35, and it shows that this hypothesis should be rejected. In other words, the CNR values for outliers are not statistically odd, and therefore do not provide a good way to filter the data. The best way so far, according to the experiments presented in this work, is to directly filter the radial wind speeds.

Figure 35 shows that the outliers type A persist for some time instants and that the majority of outliers of type B persist as well.

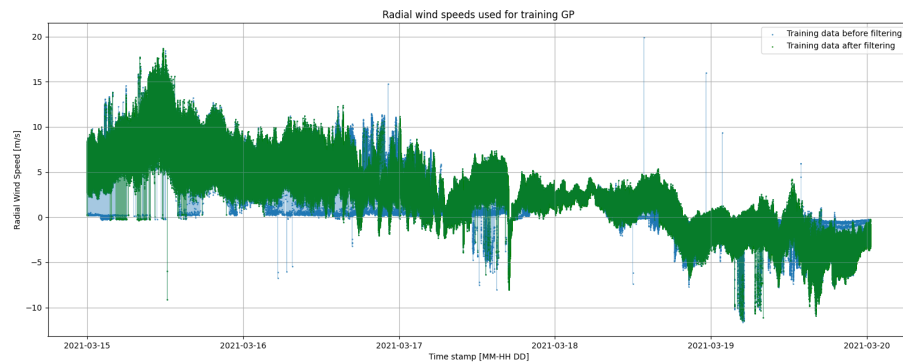


Figure 35: Non filtered and filtered data using the Hampel filter acting on CNR measurements as opposed to radial wind speeds. Window of 2000 and standard deviation threshold of 2 were used.

3.3 Conclusions and discussion

Based on the analysis carried out in this Chapter, the following is concluded:

1. **The Gaussian Processes can be used as a tool for scanning LiDARs to interpolate radial wind speeds accurately.** By applying the Gaussian Processes to one single scanning LiDAR for the majority of the onshore campaign (15th of March to 7th of April) and comparing the radial wind speed predictions for the cup location and the cup anemometer wind speed measurements resolved onto the LiDAR beam satisfactory results were found. A linear regression fitting on both time series showed an R^2 value of 0.934 and a linear relation of $1.01x + 0.08$.
2. **Filtering radial wind speed measurements is an important step before the Gaussian Processes regression fitting.** It was found that the LiDAR measurements may contain erroneous data and if the GP is fitted onto such data, the final model will yield predictions that are far off from the true value. Such points appear despite filtering for a CNR range of -27 and -3 dB and only when the flag *radial wind speed status* from the LiDAR is set to true. It was shown that using a Hampel filter that acts directly on the radial wind speed measurements can provide a solution, thus allowing one to derive accurate radial wind speed predictions (not contaminated by outliers). As for the presence of such outliers, these can be segmented into 2 categories, A and B. The A category shows sudden drops of radial wind speed measurements to 0, which may be related to being in the vicinity of the met mast, and category B with sudden drops/jumps of radial wind speed to very low/high values, for which no concrete explanation can be given at this point.

4 Combination of LiDARS

4.1 Goal and methodology

The use of Multi Lidar (ML) has clear advantages to classical profiling LiDARS relying on beam swinging. When the wind flow observed is complex, the homogeneity assumption underlying the Doppler beam swinging (DBS) or velocity azimuth display (VAD) techniques often breaks down [8]. On the contrary, combining radial measurements from intersecting LiDAR beams, allows deriving highly resolved time series of 2- and 3- dimensional wind-vector and turbulence statistics [9].

The starting point is the definition of the scalar radial wind speed ($RWS_{i,A}$) as a function of (1) the unknown true wind velocity at a given point A in space ($V_A = (u_A, v_A, w_A)$) and (2) the line of sight of LiDAR i (LOS_i):

$$\begin{aligned} RWS_{i,A} &= V_A \cdot LOS_i \\ RWS_{i,A} &= u_A \cos(\theta_i) \cos(\phi_i) + v_A \sin(\theta_i) \cos(\phi_i) + w_A \sin(\phi_i) \end{aligned} \quad (4.1)$$

From this equation, it can be observed that with 3 beams, and therefore 3 radial wind speed measurements ($RWS_{i,A}, i = 1, 2, 3$), the problem is complete. It is possible to obtain the value of the 3 components u, v, w through matrix inversion. In the presented case, only 2 LiDAs are used. To allow the resolution of the set of equations, the vertical component of the wind is assumed to be $w_A = 0$. Within the validity of this assumption, the wind components are obtained as follow:

$$\begin{pmatrix} u_A \\ v_A \end{pmatrix} = M \cdot \begin{pmatrix} RWS_{1,A} \\ RWS_{2,A} \end{pmatrix} = \begin{pmatrix} \sin(\theta_1) \cos(\phi_1) & \cos(\theta_1) \cos(\phi_1) \\ \sin(\theta_2) \cos(\phi_2) & \cos(\theta_2) \cos(\phi_2) \end{pmatrix}^{-1} \begin{pmatrix} RWS_{1,A} \\ RWS_{2,A} \end{pmatrix} \quad (4.2)$$

The reconstruction algorithm is numerically undefined when the matrix M cannot be computed, ie. when :

- $\phi_i = \phi - \frac{\pi}{2}$, ie. if the beam is pointing to the zenith ($\phi = \frac{\pi}{2}$).
- $\theta_1 - \theta_2 = \pi n$, $n \in \mathbb{Z}$, ie. if the beams 1 and 2 are parallel.

Furthermore, it is possible to derive the wind speed (WS) and wind direction (WD) scalar fields through classical geometry:

$$WS_A = \sqrt{u_A^2 + v_A^2} \quad (4.3a)$$

$$WD_A = \tan^{-1}\left(\frac{v_A}{u_A}\right) \quad (4.3b)$$

In the present situation, summarised by Figure 36, it is clear that the LiDARS need to look within the same quadrants for the beams to intersect, ensuring (2), and that the beams cannot be pointed at the zenith and still intersect, ensuring (1).

This (semi-)complete resolution of the wind field velocity comes to a price. The measurement setup requires high accuracy and precision in the control of the separate

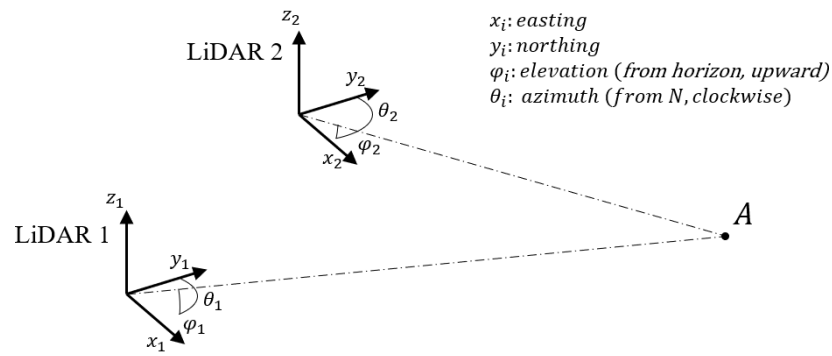


Figure 36: Schematic representation of the geometry of the problem. Point A represents the intersection of the laser beams (dashed lines) between LiDAR 1 and 2

laser beams to achieve one or multiple intersection points in time and space. Tests done by [7] illustrate the importance of this step.

Here, this limitation is tackled by leveraging the potential of Gaussian Processes to interpolate measurement data onto a predefined temporal and spatial grid G . In the results presented, the time step for G is kept to 1s to match the frequency of the cup anemometer data, while the spacial step size varies and is specified punctually.

With the prediction grid defined, it becomes possible to apply equation 4.2 to each predefined point $A_i \in G$, allowing the recreation of the horizontal wind speed at desired z levels.

4.2 Application of the Gaussian based Multi Lidar method

In this section, the accuracy of the GPbML is presented and quantified. Similar to the previous results, high frequency (1Hz) and 10-minute statistics are presented. Furthermore, the impact of using a filter to remove statistically unexpected measurements is analysed (see chapter 3.2.3). Here, the period of the analysis is from 2021-03-15 to 2021-03-19. This window was chosen as it allowed to analyse of four days without interruption of neither the TNO's LiDAR nor the Fraunhofer LiDAR measurements. This was done as measurement data analysed had gaps of a few minutes up to multiple days. While these gaps do not infer with the Gaussian Process interpolation, they do have an impact on the accuracy of the results. Here, for simplicity, these gaps have been omitted.

From Figure 37 it can be seen that there are important improvements between the 1Hz and 10min averaged regressions. The effect of outliers for the studied time window is not as noticeable as for the tests presented in section 3.2.3. The regression shows high fidelity, with regression factors $R^2 > 0.95$ and a linear relation $y = ax + b$ with $a \simeq 1$ and $b < 0.2$.

Figure 38 and Figure 39 illustrate and quantify the accuracy of the horizontal wind field reconstruction compared against the measurements from the cup anemometer. Through the decomposition of the wind velocity in its horizontal components u and v , information on the accuracy of the reconstruction of the wind direction becomes visible.

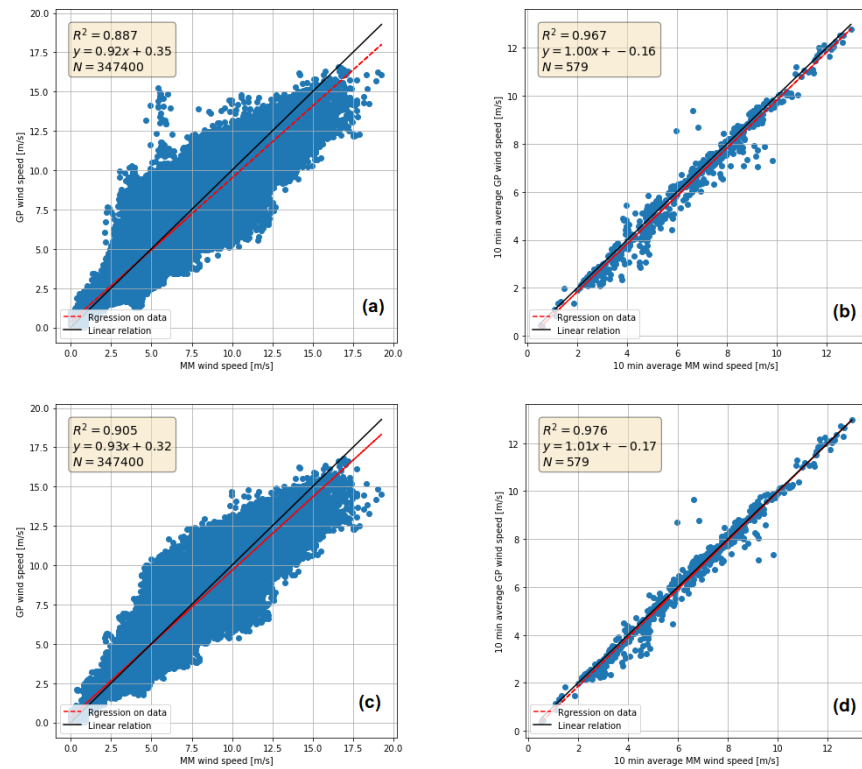


Figure 37: (a): linear regression between the cup anemometer wind speed measurement on the meteorological mast (MM) and the prediction of the GPbML at the anemometer location - 1Hz (b): same as (a) but for 10 min statistics. (c): same as (a) but with outlier filtering. (d): same as (b) but with outlier filtering

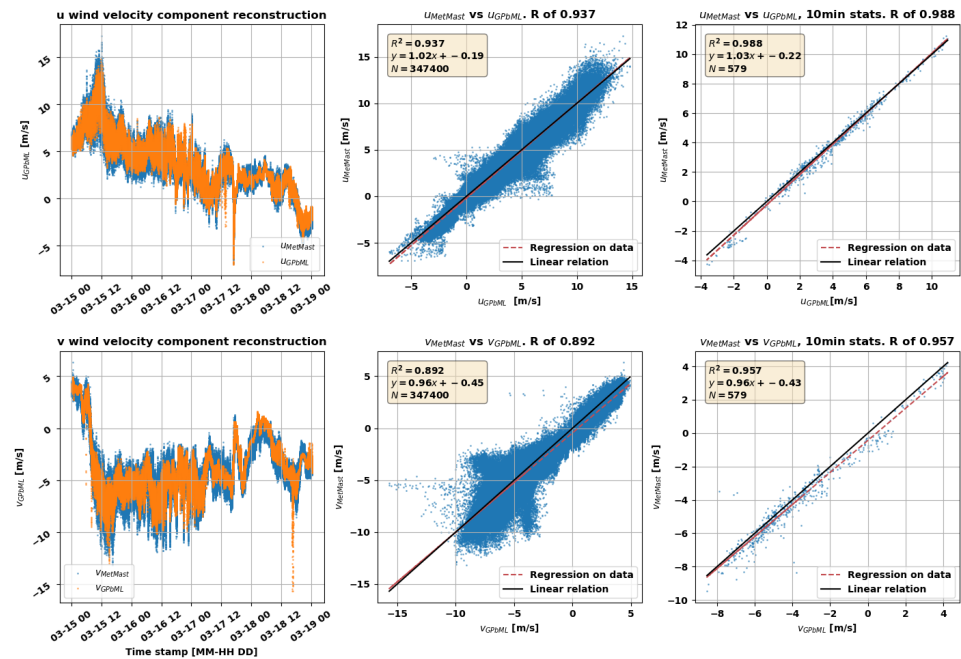


Figure 38: top: (left to right) The first figure depicts the time series of u component as obtained from the GPbML prediction (orange) and measurements from the cup anemometer and the wind vane (blue). The second and third figures are 1-to-1 plots of the 1Hz data and the 10 minutes statistics data from both the GPbML prediction and the cup anemometer measurements. bottom: The figures depict identical analysis as the top ones but for the v component.

From Figure 38 shows that both the high frequency and the 10 min statistics between the GPbML predictions and the measurements from the cup anemometer match accurately ($R^2 > 0.95$). As for the linear relation, the results show a higher level of mismatch than what was observed with the wind speed comparison. Here the parameters are $a \simeq 1$ but $0.2 < b < 0.5$. The divergence between the 1-to-1 fit and the results is particularly marked for the v component. Clusters of mismatched values appear. The nature of these mismatches is unclear. A couple of hypotheses can be made:

- The reconstruction encounters singularities that inflate errors. From Figure 37, where only the wind speed is depicted, these errors do not seem to appear. Therefore, it is plausible that these errors come from the computation of the wind direction using equation b.
- The second LiDAR (north), which data has been less thoroughly analysed, has measurement issues. This is plausible as errors are appearing mostly for the v component and the northern LiDAR is almost aligned with the y-axis of the reconstruction. In the analysed scenario, the wind direction is N-W (see Figure 40) and therefore there should not be physical reasons why one component should be harder to reconstruct than the other.

Future work outside the scope of this report will attempt to tackle these un-explained phenomena.

Figure 39 show that the removal of outliers shows as expected a slight increase in the accuracy of the predictions. Nonetheless, the clusters of mismatched values between the GPbML and the data from the cup anemometer remain.

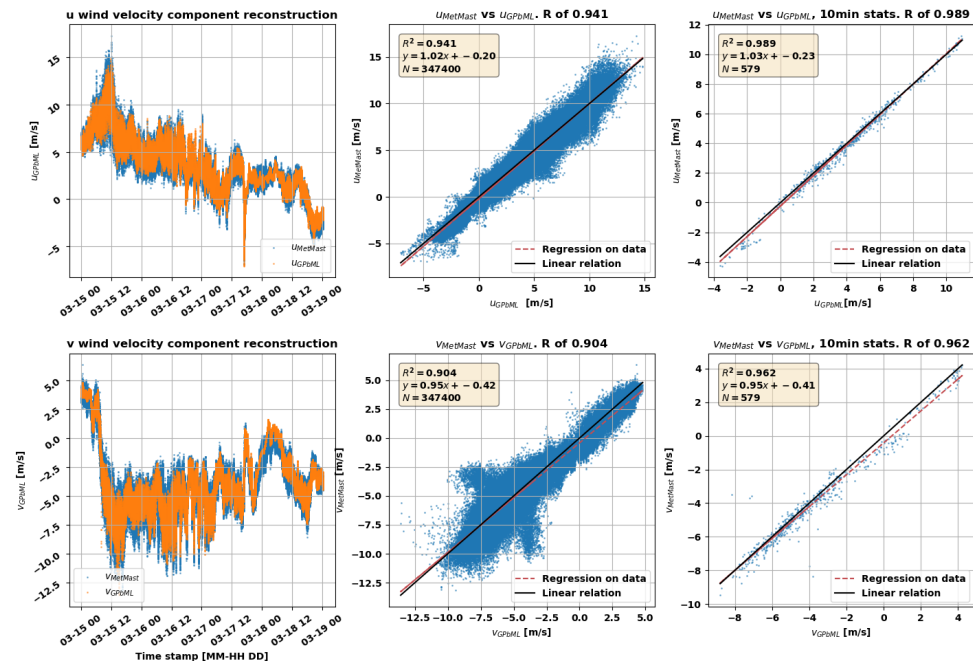


Figure 39: top: (left to right) The first figure depicts the time series of u component as obtained from the GPbML prediction (in orange) and the cup anemometer measurements. All input data was filtered using a Hampel filter to limit the number of outliers, identified as statistically abnormal measurements. The second and third figures are 1-to-1 plots of the 1Hz data and the 10 minutes statics data from both the GPbML prediction and the cup anemometer measurements. bottom: The figures depict identical analysis as the top ones but for the v component.

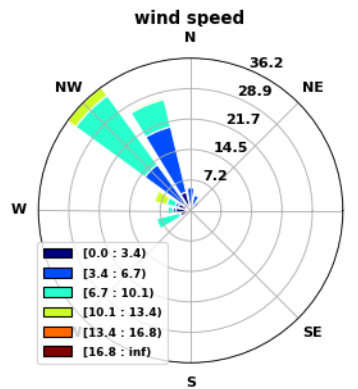


Figure 40: Wind rose during the four days time window 15.03.2021-19.03.2021

For 2-dimensional representation, vertical and planar slices are also presented. Figure 41 illustrates the 3 dimensional reconstructed field through planar slices along x, y and z . In the presented case, the wind is coming from SWW (235 deg). Structures of the wind field could be visualised. Unfortunately, not enough measurement data is available to validate 2 and 3D reconstructions of the wind field.

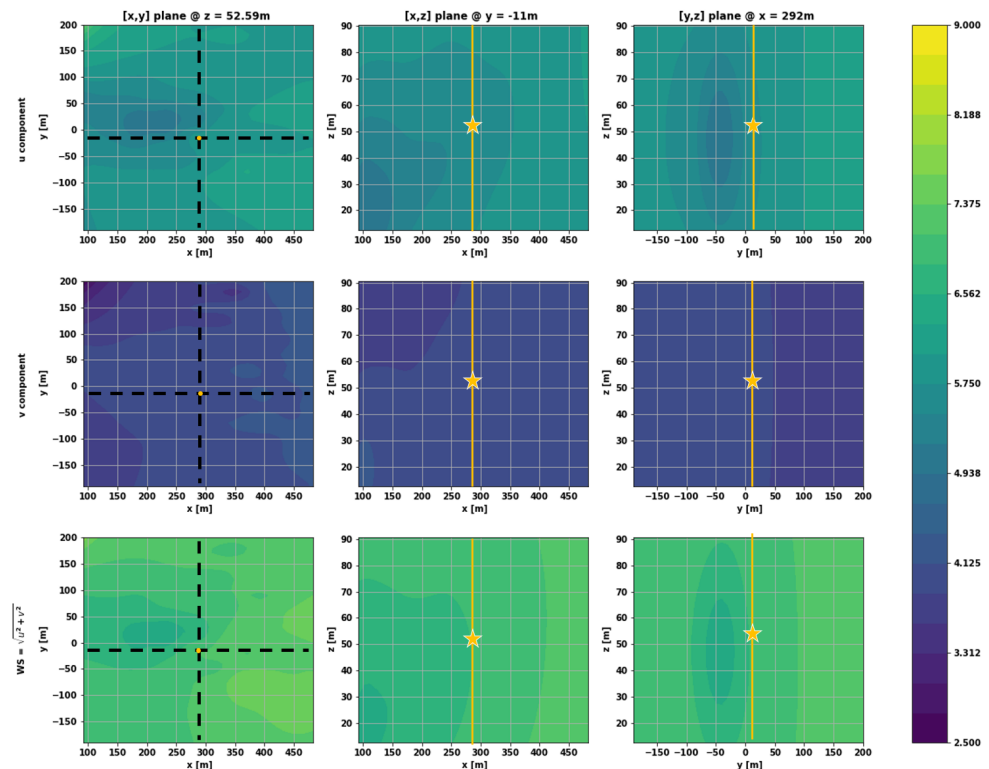


Figure 41: Rows indicate the variable (ex. 1st row: u component), columns indicate the slice geometry (ex. 1st column, horizontal plane). The color code indicates the intensity of both the wind speed and its planar components. The dashed line in the left column indicates the location at which the 3d field is sliced vertically. The orange line and dot indicate the position of the meteorological mast. The star indicates the position of the cup anemometer. The $[x,y]$ plane was obtained using a prediction grid G of size 40×40 (10m resolution). The $[x,y,z]$ planes were computed using a prediction grid G also of size 40×40 (10m along x/y and 2m resolution along z)

4.3 Conclusions and discussions

From the presented results, it can be seen that the proposed method combining Multi Lidar reconstruction and Gaussian Processes wind field prediction is capable of accurately reconstructing the wind field. In the presented case study, two LiDARs are used, allowing for a horizontal resolution of the wind field, as long as the vertical component w can be safely assumed to be 0. A third LiDAR would allow testing a full resolution of the wind field. The accuracy of the method was verified through both high and low frequency (1s and 10 min statistics respectively) linear regression analysis against cup anemometer and wind vane measurements. In both cases, the analysis has shown a close fit ($R_{1s}^2 \simeq 0.9$, $R_{10min}^2 > 0.95$ and regression coefficients $(a, b) \simeq (1, 0)$ respectively). Furthermore, spatial visualisations have been presented to give an outlook onto the capabilities of the method.

5 Uncertainty evaluation

5.1 Goal and methodology

Gaussian Process regression uses a prior and input training data to provide a statistical distribution as a prediction for the interpolated data. This distribution has a mean (μ) which is to be interpreted as the "most likely" value. The same distribution also has a variance (σ^2), which is to be interpreted as the confidence of the GP in the estimated mean value. The square root of this variance (here called standard deviation, or σ) is hereafter used to quantify the uncertainty of each estimate provided by the GPs. For the GPyDAR model, RWS predictions are associated with an RWS uncertainty. In this chapter, uncertainty predictions from the GPyDAR tool and the derived GPbML tool are briefly presented and analysed.

While for the RWS analysis, the uncertainty value is readily available, a transformation is applied onto the latter through equations 4.2 to propagate it to sub signals ($\sigma_u, \sigma_v, \sigma_{WS} \dots$).

In engineering applications, a simple formula assuming linear characteristics of the gradient of f can be derived and used as propagation error formula for $X = f(u, v, \dots)$ [10]:

$$\sigma_X^2 \simeq \sigma_u^2 \left(\frac{\partial X}{\partial u} \right)^2 + \sigma_v^2 \left(\frac{\partial X}{\partial v} \right)^2 + \dots \quad (5.1)$$

Equation 5.1 is applied to 4.2 and 4.2 to obtain the following set of equations:

$$\sigma_u^2 \simeq \sigma_{RWS_1}^2 M_{(1,1)}^2 + \sigma_{RWS_2}^2 M_{(1,2)}^2 \quad (5.2a)$$

$$\sigma_v^2 \simeq \sigma_{RWS_1}^2 M_{(2,1)}^2 + \sigma_{RWS_2}^2 M_{(2,2)}^2 \quad (5.2b)$$

$$\sigma_{WS}^2 \simeq \sigma_u^2 \frac{u^2}{u^2+v^2} + \sigma_v^2 \frac{v^2}{u^2+v^2} \quad (5.2c)$$

$$\sigma_{WD}^2 \simeq \sigma_u^2 \frac{v^2}{(u^2+v^2)^2} + \sigma_v^2 \frac{u^2}{(u^2+v^2)^2} \quad (5.2d)$$

It is now possible to resolve in every point in space the wind field, given the assumption of $w = 0$, and associate this result with an uncertainty hereby measured through the standard deviation σ_X from the mean value $\bar{X}$.

5.2 Results

5.2.1 Uncertainty evaluation for cup location

The interpolations analysed in section 3 provided by the Gaussian Process are accompanied by a value of variance that encodes the uncertainty of that same prediction. The later uncertainty is dependent, for example, on the spatial/temporal interval between the data the GP was fitted onto and the point where information is being extrapolated.

Figures 42 and 43 represent the predictions yielded by the GP in orange, the wind speed resolved to the LiDAR beam in blue, and in red the confidence intervals limited by two standard deviations of the predicted radial wind speeds. It is of interest to see that for the zoomed-in represented time windows, the true value, i.e., the wind speed

resolved to the cup, is within the confidence intervals of the GP prediction. The results illustrated in Figures 42 and 43 were computed using the filtered data and short time window (5 days) in subsection 3.2.3.

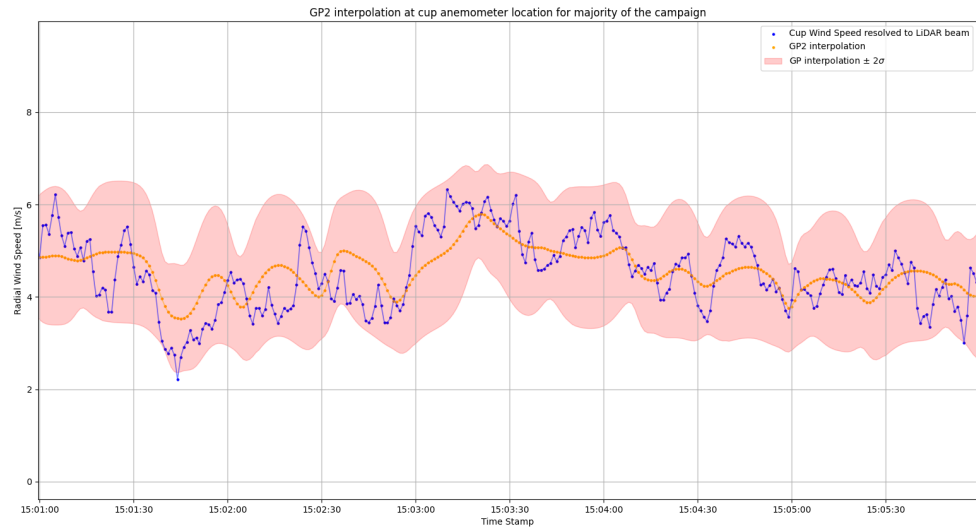


Figure 42: Evaluation of GPs predictions (orange) with corresponding uncertainty (red). Detailed results for 5 minutes during time window on the 15th of March

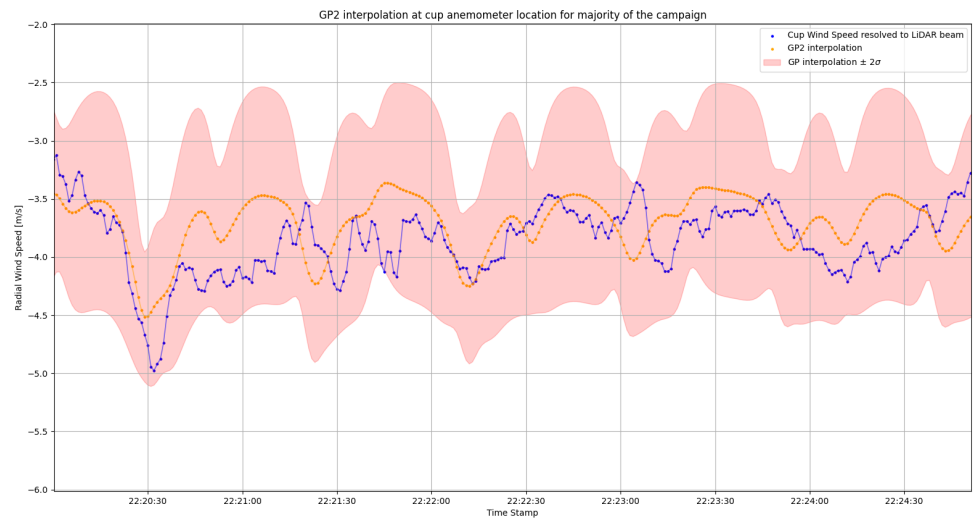


Figure 43: Evaluation of GPs predictions (orange) with corresponding uncertainty (red). Detailed results for 5 minutes during time window on the 22nd of March

It is important to note that the uncertainty values vary depending on the time window being analysed. Figure 44 shows the variation of the standard deviation throughout the complete campaign, based on the results of subsection 3.2.2. For example, the uncertainty can reach very high values of 3 m/s, as was the case in the vicinity of the 29th of March (during this interval the GP was fitted onto corrupted data as the spikes in the predictions in Figure 22 show).

The 95% confidence intervals obtained from the predicted uncertainty ($\pm 2 * \sigma_i$, where $i = RWS, u, v, \dots$) were tested. This was done by evaluating the fraction of total

measurements coming from the meteorological mast (cup anemometer, wind vane) that are encompassed within the said confidence interval, and how many escape this prediction.

It was found that when applying the GPs to one single scanning LiDAR, for the majority of the campaign (15th of March to 7th of April) the cup anemometer wind speed measurements resolved to the LiDAR LOS were 87% (of all time instants recorded) within the above-stated confidence interval.

For the GPbML, over 4 days, it was observed that, for u and v , 82% and 80% of the met mast measurements were found within the confidence intervals, respectively.

It should be emphasized at this point that the standard deviations yielded by the GP and often referred to within the context of this work as uncertainty does not translate the uncertainty derived from the equipment and measurements themselves, but merely with the confidence of the Gaussian Process when predicting a certain value of radial wind speed.

Overall, the standard deviation varied between the minimum value of 0.065 and the maximum value of 3.239, with a mean value of 0.737 throughout the majority of the onshore campaign, as seen in Figure 44.

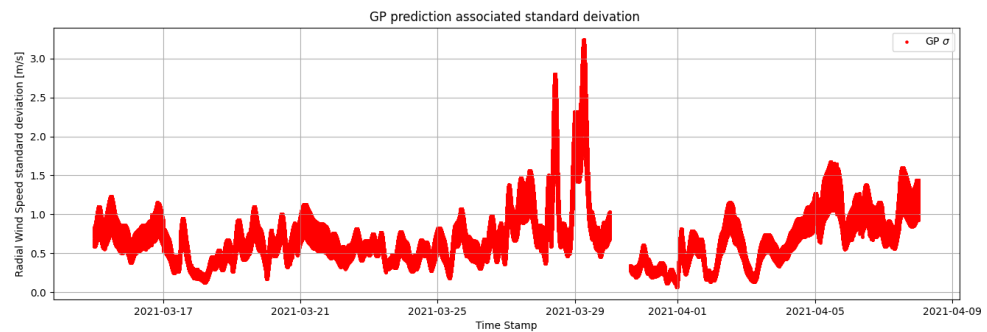


Figure 44: Variation of the standard deviation value throughout the majority of the onshore campaign.

For shorter intervals, it is also possible to analyse the variation of the standard deviation. In Figure 45 it can be seen that it presents a very clear pattern of $T = 50s$, which may be associated with the LiDAR full scanning pattern, that has a period of similar length. This is further discussed in subsection 5.2.2.

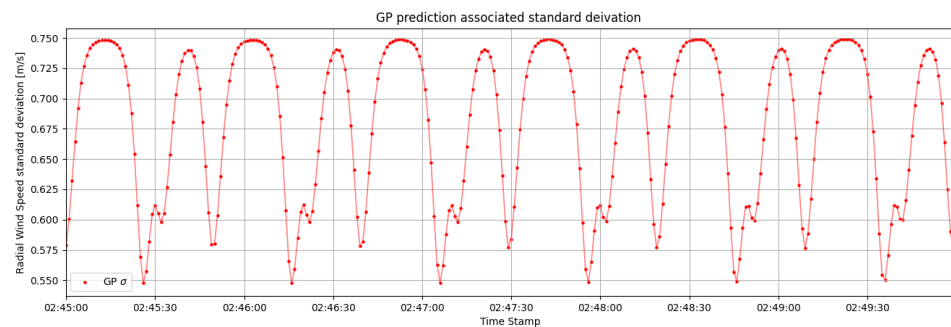


Figure 45: Minute based variations of standard deviation yielded by the GP.

5.2.2 Evolution over time and space of the uncertainty prediction

In this section, several visualisations of the uncertainty prediction produced by the GPbML are presented.

Figure 46 depicts the vertical spatial mapping of the scalar fields σ_u , σ_v and σ_{WS} . In these tests, the resolution is 10m along x/y and of 2m along z for a 1x40x40 grid. The slices are used as support to visualise how the standard deviation of each variable follows closely the position of the beam ie. the closest measurement in time.

Figure 47 depicts the evolution of the 3d standard deviation scalar fields. The time window used as support begins on 15-03-2021.

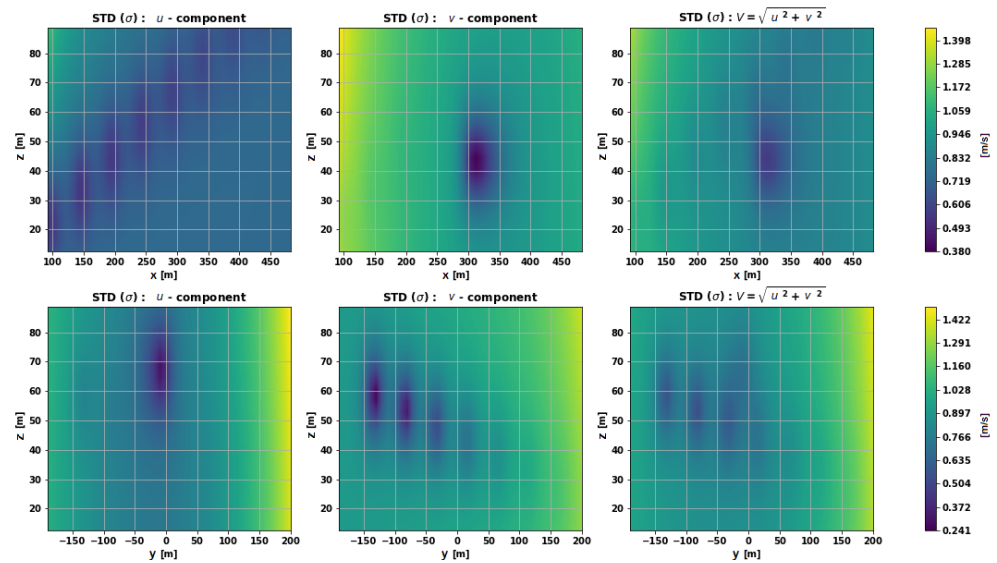


Figure 46: Vertical slices along x with $y = -50\text{m}$ (top) and along y with $x = 350$ (bottom) of the standard deviation for (from left to right) the u component, the v component.

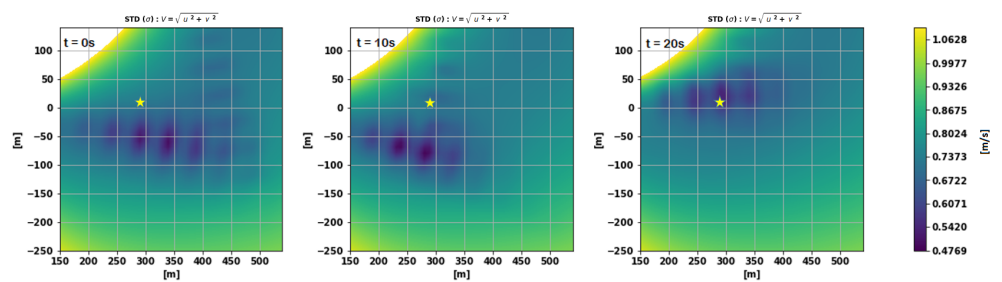


Figure 47: Three time frames ($t = t_0, t_0 + 10\text{s}$ and $t_0 + 20\text{s}$) of the horizontal wind speed standard deviation scalar field. The Met Mast is represented by the yellow star.

Figure 48 presents insight into the variation in time of the vertical profiles of the model uncertainty. In these tests, the grid has been defined as a vertical cluster of points with 1m of spacing for a final 1x1x80 grid. From the latter, it can be seen that the values of the σ_u and σ_v vary from a lowest of 0.5 to a highest of 0.8. The variation of the total wind speed is more correlated to the u component as at the time instant presented the wind direction is SWW ($\simeq 235[\text{deg}]$). Furthermore, from the vertical profiles of the u component, it is possible to identify the western LiDAR scanning patterns at 9deg

(lower mode, $\tan^{-1}(46.5/292) \simeq 9deg$), 11deg (middle mode, $\tan^{-1}(57/292) \simeq 11deg$) and 13deg (upper mode, $\tan^{-1}(67.5/292) \simeq 13deg$).

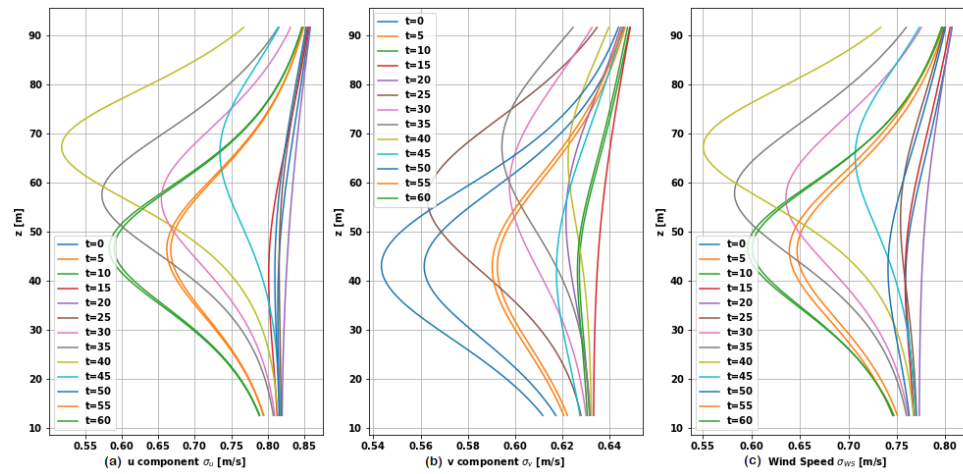


Figure 48: Vertical profiles at $[x, y]_{MetMast}$ over 80m in height for (a) the u component, (b) the v component and (c) the total wind speed. The visualisation is done over 60s during which the scanning pattern is complete and starts over (note some overlapping due to the restart of the scanning pattern, which is expected to last 45s).

Finally, Figure 49 shows the time series of the wind speed uncertainty as predicted by the GPbML at the location of the anemometer. The signal is periodic. The frequency analysis indicates that the observed modes are harmonics of 0.02 Hz or 50s. Understanding the origin of this fundamental frequency is set for future work and does not fall within the scope of this report.

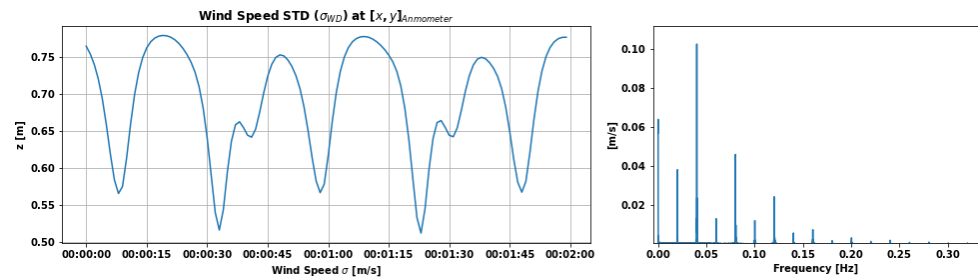


Figure 49: *left:* Evolution in time of $\sigma_{SD}W$ at $[x, y, z]_{anem}$ over a period of 120s. It can be seen that there are multiple modes appearing. *right:* Fast Fourier Transform of σ_{WS} highlighting the harmonic components of the signal ($f_{fund} = 0.02Hz$)

5.2.3 GP uncertainty compared against vertical wind profile extrapolation

As one of the key values of the presented method is the possibility to derive a "virtual measurement" anywhere in space, a useful application to it is to use it when errors are made during the set up of instruments. Here, we approach the possibility of aiming our scanning pattern at the wrong elevation. In this scenario, the GPbML allows using the obtained data to query virtual measurements at the original point of interest.

To test this, two wind profile laws are used. These are commonly used to extrapolate wind speed to higher or lower positions, from a point measurement in space. The two are similar and both rely on one parameter each. The Log Law (equation 5.3) requires the "roughness length" and the Power Law (equation 5.4) requires the shear

factor [11].

$$v(z) = v_h \frac{\log(\frac{z}{z_0})}{\log(\frac{h}{z_0})} \quad (5.3)$$

$$v(z) = v_h \left(\frac{z}{h}\right)^\alpha \quad (5.4)$$

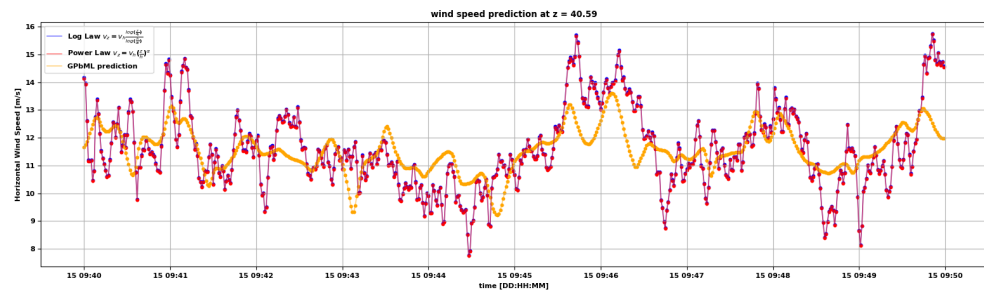


Figure 50: Evolution over time of (1) a Log Law - $z_0 = 0.05$ from tables -, (2) a Power Law - $\alpha = 0.1$ and (3) the GPbML predictions of the wind speed 12 m below the cup anemometer (40.59).

The main drawback of these laws is the need to manually select the mentioned parameters. Here, the nonparametric nature of the GP method becomes a clear advantage.

From Figure 50, it can be observed the comparison of (1) the Log Law, (2) the power law, and (3) the GPbML prediction of the wind speed. Vertical profile laws allow maintaining the high frequency (turbulent) dynamics of the signal, something that the GP omits. Nonetheless, these dynamics are not necessarily representative of the reality as they are translated from the extrapolation origin at 52.59m. Furthermore, the GPbML extrapolation⁸ does not require inputs such as the shear factor or roughness length and gives a prediction of comparable accuracy. In addition, an uncertainty estimation comes along with the mean estimate acting as a turbulent prediction.

5.3 Conclusions and discussions

The GPyDAR tool gives an estimate of the uncertainty around the most likely RWS value estimated or so-called "prediction". A method allowing to propagate this uncertainty to RWS's pseudo-signals such as wind direction and wind speed was proposed.

The uncertainty of the GP when applied to a single scanning LiDAR was firstly assessed. It was seen that the cup measurements are at the majority of times within the confidence interval defined by 2 standard deviations yielded by the GP. Throughout the campaign, it was seen that 87% of measurements are within the confidence interval. Furthermore, it was also shown that the standard deviations yielded by the GP are not constant throughout the campaign and may present variations (a minimum value of 0.065 and a maximum value of 3.239 were registered, with a mean value throughout the campaign of 3.239).

Similarly, a test on the u and v estimates given by the GPbML over 4 days showed that confidence intervals encompass 80% and 82% of the respective measurements coming from the cup anemometer combined with the wind vane. This may indicate

⁸Note that GP regression is made for interpolation and not extrapolation. As we attempt to extrapolate, the uncertainty of the estimated value will increase rapidly as we move away (in our case either in time and/or space) from the input data [12]

that the uncertainty prediction for the GPbML should be less confident, and have larger bounds.

6 Conclusions and future work

6.1 Conclusions

In the presented document, an algorithm was validated to reconstruct the wind field from scanning LiDAR data. The algorithm used Gaussian Processes regression to interpolate radial wind speed measurements, and an algorithm to resolve the real horizontal wind field.

The first step presented has been the demonstration of the developed tool to estimate high-frequency virtual RWS measurements. These estimates have been compared and validated against measurements obtained through a cup anemometer. The estimated punctual RWS showed high fidelity with the anemometer data (10min statistics: $R^2 = 0.979$, $y = 1.01x + 0.05$).

The tool was further developed into a Multi LiDAR, allowing for two-dimensional reconstruction of the wind field (ie. wind speed and wind direction, assuming no vertical component). Similar to the here-fore presented validation, the method was compared against data coming from both a cup anemometer and a wind vane. The wind speed analysis showed high accuracy in the estimates (10min statistics: $R^2 = 0.967$, $y = 1.01x + 0.17$). The wind direction analysis was presented through the decomposition of the wind in u and v components. The u component regression showed high fidelity (10min statistics: $R^2 = 0.988$, $y = 1.03x + 0.22$). The results for the v component reconstruction were of lower accuracy (10min statistics: $R^2 = 0.957$, $y = 0.96x + 0.43$). The origins of this discrepancy could not be investigated in this work.

The appearance of statistical outliers within the LiDAR measurement time series was also detected. These outliers were identified and categorised in two types (A, B), characterised by their distance from the local median value, their frequency of appearance, and most importantly their impact on the GPyDAR tool training. Outliers of Type B were seen to locally mislead the Gaussian Processes, leading to incorrect radial wind predictions. A statistically-based Hampel Filter was used as an approach to successfully remove these outliers.

A methodology to propagate estimation uncertainty yielded by the Gaussian Processes was presented. The evolution in time and space of this uncertainty output was analysed. The spatial variation of the latter was seen, as expected, to be correlated with the LiDAR(s) beam motion. The temporal variability was also observed to be influenced by the beam position, but unexpected periodic patterns were observed. The origin of these patterns remains to be clarified.

Overall, the work presented has given evidence of the capabilities of using Gaussian Processes regression to add value in the use of scanning LiDARs within the scope of reconstructing two(three)-dimensional wind field at high frequency. Using spatially sparse data coming from scanning patterns, it was possible to both increase the accuracy of point RWS measurement and develop a Multi LiDAR method that does not require high precision beam control, nor relies on homogeneous field assumptions. The potential of this method is large, allowing the creation of virtual meteorological masts at nearly any point in space. Furthermore, it was shown that beam position correction can be easily done without relying on shear models, an important advantage of the GP model. Nonetheless, the preprocessing of data before the use of the model remains, for now, a work in progress.

The method tested on this onshore campaign has furnished results of quality. It is with

confidence that the method can move to the next step of the GloBE project. Given measurements from numerous scanning LiDARs, it is expected that the quality of the wind field analysis done using the GbML achieves similar accuracy as the traditional Dual Doppler method, although further work is necessary to evaluate such claim. This will be further evaluated in the context of the GLOBE project. The additional qualities of the GbML such as (1) measurement correction without the need for empirical shear models and (2) fast interpolation between large and sparse data sets (common to scanning LiDARs), arguably make it a valuable tool to estimate the large scale dynamics of the wind field in large areas such as the surroundings of offshore wind farms.

Table 3: Summary of results from GP application to data from the onshore campaign.

Analysis	Duration	Frequency	Filtered	Measured	Calculated	Linear fit		R ²
						a	b	
Single LiDAR	26 days	1 Hz	No	Cup anemometer wind speed resolved to LiDAR beam	GP radial wind speed prediction at cup location	1.01	0.08	0.934
	26 days	10 min	No	Cup anemometer wind speed resolved to LiDAR beam	GP radial wind speed prediction at cup location	1.01	0.05	0.979
Single LiDAR	5 days	1 Hz	Yes	Cup anemometer wind speed resolved to LiDAR beam	GP radial wind speed prediction at cup location	1.02	-0.22	0.962
	5 days	10 min	Yes	Cup anemometer wind speed resolved to LiDAR beam	GP radial wind speed prediction at cup location	1.02	-0.22	0.991
Dual Doppler	4 days	10 min	No	u velocity component (using cup and vane)	GPs radial wind speed prediction at cup location and reconstruction	1.02	-0.19	0.937
	4 days	1 Hz	No	v velocity component (using cup and vane)	GPs radial wind speed prediction at cup location and reconstruction	0.96	-0.45	0.892
	4 days	10 min	No	u velocity component (using cup and vane)	GPs radial wind speed prediction at cup location and reconstruction	1.03	-0.22	0.988
	4 days	10 min	No	v velocity component (using cup and vane)	GPs radial wind speed prediction at cup location and reconstruction	0.96	-0.43	0.957
Dual Doppler	4 days	1 Hz	Yes	u velocity component (using cup and vane)	GPs radial wind speed prediction at cup location and reconstruction	1.02	-0.2	0.941
	4 days	1 Hz	Yes	v velocity component (using cup and vane)	GPs radial wind speed prediction at cup location and reconstruction	0.95	-0.42	0.904
	4 days	10 min	Yes	u velocity component (using cup and vane)	GPs radial wind speed prediction at cup location and reconstruction	1.03	-0.23	0.989
	4 days	10 min	Yes	v velocity component (using cup and vane)	GPs radial wind speed prediction at cup location and reconstruction	0.95	-0.41	0.962

6.2 Future work

Future work concerning the development of the Gaussian Processes will most likely address the following points:

1. **Increasing GPs robustness to input data:** it was seen that in certain time windows data from the LiDARs was missing. The time windows were short, not spanning more than a few minutes. However frequent gaps as the one observed can create limitations to the GP regression accuracy. Currently, the solution found was to simply execute two different fittings on the adjacent time series where data exists. In the future, such a solution is not desirable. This will require lower-level coding work to be tackled in the future.
2. **Standardizing filtering of LiDAR data:** it was seen that erroneous data measurements may exist, despite the initial filtering of data based on CNR thresholds and logical quality flags given by the LiDAR (radial wind speed status). It was further concluded that such erroneous data measurements will lead to a GP which will provide abnormal predictions. Therefore more thorough filtering of the data has been observed to be an important step in the workflow. The filtering approach chosen is of course dependent on certain parameters that the user can choose. This for now is not a generalized solution and may have different results for different datasets themselves. A standard approach to the tuning of the filtering parameters for future projects is seen as a possibly important improvement.
3. **Understanding limitations of the GPs:** although throughout the work here presented a satisfactory comparison between the measurements taken on the test site and the GP predictions was found, further understanding of this comparison is deemed necessary. A preliminary residual analysis was done to check for the distribution of residuals and the dependence of the latter on other variables (wind speed and wind direction, for example). However, further investigation should be carried out to properly evaluate the limitations of GPs in topics such as turbulence estimation and uncertainty determination.
4. **Validation of GPs reconstruction with simulation data:** the Gaussian Process-based Multi-LiDAR method presented allows to derive wind field reconstructions over a spatial grid. Such reconstruction has been validated against a single-point wind speed measurement and single-point wind direction measurement. Its precision for multi-dimensional reconstruction is currently unknown. It is of interest to further validate such wind field reconstruction. Theoretically, many measurements points would be needed, which may become cumbersome for a real field experiment. Therefore, an interesting alternative would be to validate such wind field reconstruction using "numerical LiDARs" measurements obtained from high-fidelity CFD simulations (DNS, LES).

7 Bibliography

- [1] Clym Stock-Williams. *Wind field reconstruction from Lidar using Gaussian Processes: Memorandum to OWA Globe project partners*. Tech. rep. TNO, 2020.
- [2] Clym Stock-Williams, Paul Mazoyer, and Sébastien Combrexelle. "Wind field reconstruction from lidar measurements at high-frequency using machine learning." In: *Journal of Physics: Conference Series* 1102 (2018), p. 012003. DOI: 10.1088/1742-6596/1102/1/012003. URL: <https://doi.org/10.1088/1742-6596/1102/1/012003>.
- [3] *MS Windows NT Introduction to Gaussian Process Regression*. https://colab.research.google.com/drive/102urxK7Fr1l122Xh-f1NhkvbE-HfhE_K?usp=sharing. Accessed: 02-09-2021.
- [4] J. Duncan C.F.W. Stock-Williams G. Bergman. *Validation at High Frequency of Wind Field Reconstruction from Scanning Lidar using Gaussian Processes*. Tech. rep. TNO, 2018.
- [5] J.W. Wagenaar C. Stock-Williams E.K. Fritz. *Scanning LiDAR measurements in Rotterdam harbou, TNO 2019 R11982r*. Tech. rep. TNO, 2020.
- [6] *MS Windows NT Outlier Detection with Hampel Filter*. <https://towardsdatascience.com/outlier-detection-with-hampel-filter-85ddf523c73d>. Accessed: 01-10-2021.
- [7] R. Menke et al. "Multi-lidar wind resource mapping in complex terrain." In: *Wind Energy Science* 5.3 (2020), pp. 1059–1073. DOI: 10.5194/wes-5-1059-2020. URL: <https://wes.copernicus.org/articles/5/1059/2020/>.
- [8] Lukas Pauscher et al. "An Inter-Comparison Study of Multi- and DBS Lidar Measurements in Complex Terrain." In: *Remote Sensing* 8.9 (2016). ISSN: 2072-4292. DOI: 10.3390/rs8090782. URL: <https://www.mdpi.com/2072-4292/8/9/782>.
- [9] Jakob Mann et al. "Comparison of 3D turbulence measurements using three staring wind lidars and a sonic anemometer." In: *Meteorologische Zeitschrift* 18.2 (May 2009), pp. 135–140. DOI: 10.1127/0941-2948/2009/0370. URL: <http://dx.doi.org/10.1127/0941-2948/2009/0370>.
- [10] H. H. Ku. "Notes on the Use of Propagation of Error Formulas." In: (1966).
- [11] Stefan Emeis and Matthias Turk. "Comparison of Logarithmic Wind Profiles and Power Law Wind Profiles and their Applicability for Offshore Wind Profiles." In: Feb. 2007, pp. 61–64. ISBN: 978-3-540-33865-9. DOI: 10.1007/978-3-540-33866-6_11.
- [12] C. K. I. Williams. "Prediction with Gaussian Processes: From Linear Regression to Linear Prediction and Beyond." In: *Learning in Graphical Models*. Ed. by Michael I. Jordan. Dordrecht: Springer Netherlands, 1998, pp. 599–621. ISBN: 978-94-011-5014-9. DOI: 10.1007/978-94-011-5014-9_23. URL: https://doi.org/10.1007/978-94-011-5014-9_23.