Correcting turbulence measurements from continuous-wave, ground-based lidars

Department of Wind Energy E-Report

Alfredo Peña, Ginka G. Yankova and Vasiliki Mallini

DTU Wind Energy Report-E-0249

June 2024

DTU Wind Energy
Department of Wind Energy

Document Information

Project	Co-financed research between Energinet and DTU Wind and Energy Systems
Document	Correcting turbulence measurements from continuous-wave, ground-based lidars
Type	External Report
Number	DTU Wind Energy Report-E-0249
Year	2024
Contractor	Technical University of Denmark (DTU)
Authors	Alfredo Peña, Ginka G. Yankova and Vasiliki Mallini
Contact **ISBN**:	Mail: giyo@dtu.dk [978-87-87335-77-5]

Published by:
DTU, Department of Wind and Energy Systems
Frederiksborgvej 399, Building 118, 4000 Roskilde Denmark
www.vindenergi.dtu.dk

Preface

As part of the cooperation agreement on co-financed research between Energinet Eltransmission A/S (Energinet) and DTU Wind and Energy Systems (DTU Wind), whose objective is the study of correction methods to lidar wind turbulence measurements so that they agree with industry-accepted measurements of turbulence, DTU Wind's second task is to suggest a correction method for turbulence measurements from continous-wave, ground-based lidars. These lidars are of the same type as those used by Energinet to determine wind siting conditions at several positions in the North and Baltic Seas for Energinet as part of the Danish government plans with regards to offshore wind farms and energy islands.

Roskilde, 28-06-2024

Ginka G. Yankova (Senior Development Engineer), Vasiliki Mallini (Development Engineer), and Alfredo Peña (Senior Scientist), DTU Wind

1 Acronyms

2 Definitions

v_r	radial velocity
$F_{v_r}(k_1)$	radial velocity spectrum
$F_w^l(k_1)$	vertical velocity spectrum
$\hat{\varphi}$	Fourier transform of the weighting function φ
$n_{i,j}$	components of the unit vector of the lidar beams
k_1, k_2, k_3	wave numbers in the along-wind, transverse and vertical direction
$\Phi_{ij}(k)$	three-dimensional spectral velocity tensor
u	longitudinal component of the wind field in the along-wind direction
v	transversal component of the wind field perpedicular to u component
w	vertical component of the wind field
$\overline{u'w'}$	covariance between the u and w components
$\overline{v'w'}$	covariance between the v and w components
u_*	friction velocity
σ_{u_l}	standard deviation of u component of the lidar
σ_{u_s}	standard deviation of u component of the sonic
σ_{v_l}	standard deviation of v component of the lidar
σ_{v_s}	standard deviation of v component of the sonic
$\sigma^2_{u_l}$	variance of u component of the lidar
$\sigma^2_{u_s}$	variance of u component of the sonic
$\sigma^2_{v_l}$	variance of v component of the lidar
$\sigma^2_{v_s}$	variance of v component of the sonic
dU/dz	vertical wind speed gradient
X_i, Y_i, Z_i	sonic wind speed components along North-South, East-West, and vertical direction
T_i	temperature
L_O	Obukhov length
k	von-Karman constant
g	gravitational acceleration
Dir_{avg_l}	average of lidar wind direction
Dir_{std_l}	standard deviation of lidar wind direction
Dir_{min_l}	minimum of lidar wind direction
Dir_{max_l}	maximum of lidar wind direction
Dir_{avg_s}	average of sonic wind direction
spd_{sonic}	average sonic wind speed
spd_{lidar}	average lidar wind speed
spd_{cup}	average cup wind speed
$N_{minutes}$	number of minutes with at least one valid sample
N_{lidar}	number of valid lidar samples per 10min
N_{sonic}	number of valid sonic samples per 10min
z	height above the ground
L	turbulence length scale
α, β, γ	spectral weighting functions
ϕ	lidar half-opening angle
$\hat{T}_f$	spectral transfer function accounting for the low-pass filter effect
$\mathbf{n}(\phi, \theta)$	unit vector describing lidar scanning pattern
θ	azimuthal positions
d_f	lidar focused distance
z_R	Rayleigh length
λ	laser wavelength
α_0	effective beam radius at the output lens
$\alpha\epsilon^{2/3}$	dissipation rate of turbulence
Γ	turbulence anisotropy

Contents

3 Introduction

As part of the co-financed research between Energinet and DTU Wind, we established a work package (WP), aimed to provide a methodology to correct turbulence measurements performed from fixed continuous wave continuous wave (CW) wind profiling lidars so that by using these measurements we get closer to those of standard anemometers, such as sonic and cup anemometers. In particular, we are interested in analyzing measurements from a ZX 300 wind lidar, a fixed CW wind profiler, which is the same type of lidar used by Energinet at several positions within the North and Baltic seas to assess the siting conditions for the future wind farms and energy islands.

Corrections for lidar-derived turbulence measurements have been investigated for decades. The major difficulties for the correction are related to two points. First, most lidars used in the wind industry reconstruct the wind velocity components by scanning the atmosphere at different positions in space, since in most cases only one unit is available. This complicates turbulence measurements because with a single lidar, we can only measure the radial velocity; in most cases scanning at several positions leads to what is commonly referred to as 'contamination' of the target velocity variances because they become dependent (contaminated) on other components of the turbulence covariance tensor. Second, lidars measure within a probe volume, and depending on the lidar (and on the measurement range for the case of CW lidars), the length of this probe determines the size of the atmospheric turbulent eddies that we can measure. The probe volume therefore acts as a turbulence filter, also known as probe-volume averaging effect. We could thus perform 'perfect' measurements of turbulence using lidars if we had three units and their probe volume was small enough to measure the eddies with most energy.

An interesting part of the problem is what we here refer to as the lidar turbulence 'paradox'. To understand the reasoning behind the paradox, we can start by looking at the expression for the radial velocity (v_r) spectrum of a wind lidar, which is given as [Mann et al., 2009]:

$$F_{v_r}(k_1) = n_i n_j \iint |\hat{\varphi}(\boldsymbol{k}\cdot\boldsymbol{n})|^2 \, \Phi_{ij}(\boldsymbol{k}) dk_2 dk_3, \tag{1}$$

where $\hat{\varphi}$ is the Fourier transform of the weighting function φ that describes the probe-volume characteristics, $n_{i,j}$ are the components of the unit vector of the lidar beams, $k_{2,3}$ the wave numbers in the transverse and vertical directions (k_1 is on the along-wind direction), and Φ_{ij} is the spectral velocity tensor. For a close-to-ideal anemometer, e.g., a sonic anemometer, we usually assume $\hat{\varphi} \approx 1$, so the one-point spectra of the velocity components are simply:

$$F_{ij}(k_1) = \iint \Phi_{ij}(\boldsymbol{k}) dk_2 dk_3, \tag{2}$$

and the one-point velocity variances and covariances are then given as

$$\langle u_i' u_j' \rangle = \int F_{ij}(k_1) dk_1. \tag{3}$$

The lidar turbulence paradox appears already in Eqn. (1): we measure turbulence with a lidar and in the best of the cases (e.g., when we coincidentally measure parallel to the along-wind velocity component, i.e., $i, j = 1$), we just need to determine the effect of $\hat{\varphi}$ on Φ_{ij}. But Φ_{ij} is the turbulence spectral tensor, whose components depend on the turbulence characteristics. Therefore, for this 'simple' lidar case, to determine the effect of the probe volume on $F_{v_r}(k_1)$, we need to know in advance the characteristics of turbulence, which we in the first place are trying to determine with lidar-based turbulence measurements. One can now imagine that when the lidar beam is not perfectly aligned with none of the three velocity components, as in most scanning configurations, the complexity of the problem increases because all components of the spectral tensor might contribute to the radial velocity spectrum and 'interplay' with the weighting function.

There are ways to avoid the paradox. If we assume that the Doppler spectrum of radial velocities measured by a lidar contains turbulence information only, we can use this Doppler spectrum to determine the so-called unfiltered radial velocity variance, i.e., a radial velocity variance that does not suffer from probe-volume averaging effects. By using unfiltered radial velocity variances, and depending on the scanning geometry, one can determine unfiltered velocity-component covariances. Using a ZephIR wind profiler, a predecessor of the ZX 300 lidar, Mann et al. [2010] computed a portion of the momentum flux, the $\overline{u'w'}$-covariance, using the unfiltered radial velocity variance of the two sets of lidar beams that were aligned with the mean wind direction. Also, if unfiltered radial velocity variances are available, we can in principle compute all six Reynolds stresses. A procedure to compute all stresses is explained in Sathe et al. [2014] for a ground-based lidar and in Fu et al. [2023] for a nacelle-based lidar.

Unfortunately, a number of ZX 300 units do not store information with regards to the Doppler spectrum of radial velocities for each of the lidar beams and, therefore, we cannot determine the unfiltered radial velocity variances. Also, for a number of ZX 300 units (offshore and onshore versions), we cannot access information of the radial velocity estimate, which is internally performed by the unit from the Doppler spectrum of radial velocities. The rawest level of information available of these units, among them those utilized by Energinet, concerns the time series of reconstructed horizontal velocity and direction within each 10-min interval.

In summary, we are dealing with reconstructed horizontal velocities, whose second-order moments are both 'contaminated' from contributions from the other components of the spectral tensor and 'filtered' due to probe-volume averaging effects, and as we will show it is complex to understand such effects. Here, we also explore an alternative method to correct turbulence measurements, which is based on training a neural network (NN) with either physics-based datasets or directly with measurements from lidars. The NN can further be trained with different sets of inputs and we can study their level of importance; we a priori expect that the most important inputs to predict the close-to-ideal turbulence measures from lidar turbulence estimates are related to inputs we know are most dominant from theory.

The work is organized as follows. In Sect. 4, we present the site, and lidar and meteorological mast measurements, which we use to evaluate different methods to correct lidar turbulence measurements. Analysis of these measurements is presented in Sect. 5. Section 6 introduces the methodologies that we use to correct lidar turbulence measurements. Section 7 shows the results of the different methodologies and Section 9 provides bullet-point conclusions on the work.

4 Site and measurements

All measurements used in this report come from DTU's test station for large wind turbines in Østerild and cover a period of seven months in 2018 and 2019. At that time, part of the facility included a 250m tall meteorological mast in the South end of the row of turbine stands. It was heavily instrumented with cup and sonic anemometers, temperature, pressure, humidity and rain sensors. The metmast was often used for commercial lidar calibrations. Data from two of those lidars - units 590 and unit 542 - is used in this report as the basis for developing the method for correcting lidar turbulence measurements.

4.1 Site description

The test station for large wind turbines is situated in Northern Jutland as indicated in Figure 1. Back in 2018, the site comprised of seven test pads, each with its own meteorological mast; all turbines were situated in one row - from North to South. Additionally, two 250m meteorological masts were installed - one in each end of the test station. The site was characterized by grassland and forests in the Southern part; the canopy height was around 10-20m. For most of the site, the terrain variations are within 5m, whereas in the North part of the site the variations can be up to 10m. Around 4km away from the North end of the station is the North Sea coast. The Limfjord is approximately 6km away from the South end of Østerild.

Figure 1: Location of Østerild test station

4.2 Measurement setup

Light mast south in Østerild is a triangular mast with a side of 1.2m. The wind speed measurements come from either Windsensor P2546 cups or 3-D METEK sonics. All sonics and cup anemometers are side-mounted, on booms with a free boom length of 4.8m, as indicated in Figure 2. In this report, we have used data from four cup anemometers mounted at 40m, 106m, 178m and 244m. The sonic anemometers we used were mounted on the same side of the mast as the cups, just 3m below - so at 37m, 103m, 175m and 241m. The sampling frequency for the cup anemometers was 10Hz, while for the sonics it was 20Hz. As mentioned, the metmast was used for commercial calibrations, so all data from the metmast was subject to regular data quality checks and all instruments had valid calibrations and were exchanged periodically.

The two units used for the data analysis were installed close to light mast South, approximately 11m to the West of the metmast center. Unit 542 was on site from 3rd of July, 2018 until 2nd of January, 2019. Unit 590 was on site from 24th of August, 2018 until 31st of March, 2019. It should be noted that, unit 590 had special configuration settings in the period from 3rd of November until 18th of January, 2019, therefore, this period was not included in the analysis. Both lidars were configured to measure at 12 different heights in the range from 38m to 800m. The heights we are interested in are 39m, 105m, 177m and 243m. A full list of the lidars' measurement heights is presented in Table 1.

	10m	38m	39m	102m	105m	139m	174m	177m	209m	240m	243m	289m	800m
590	X	X	X	X	X	X	X	X	X	X	X	X	X
542		X	X	X	X	X	X	X	X	X	X	X	X

Table 1: Height settings for unit 590 and unit 542

Time synchronisation was ensured during the campaign - both for the metmast and the lidar measurements. It is expected that any time delay between the two systems does not exceed 10s.

It is important to mention that, for both units, the met station was installed too close to the lidar window, much closer than the recommended minimum of 1m. Additionally, the lidars were surrounded by a fence. These two factors contributed to occasionally erroneous measurements of the wind direction.

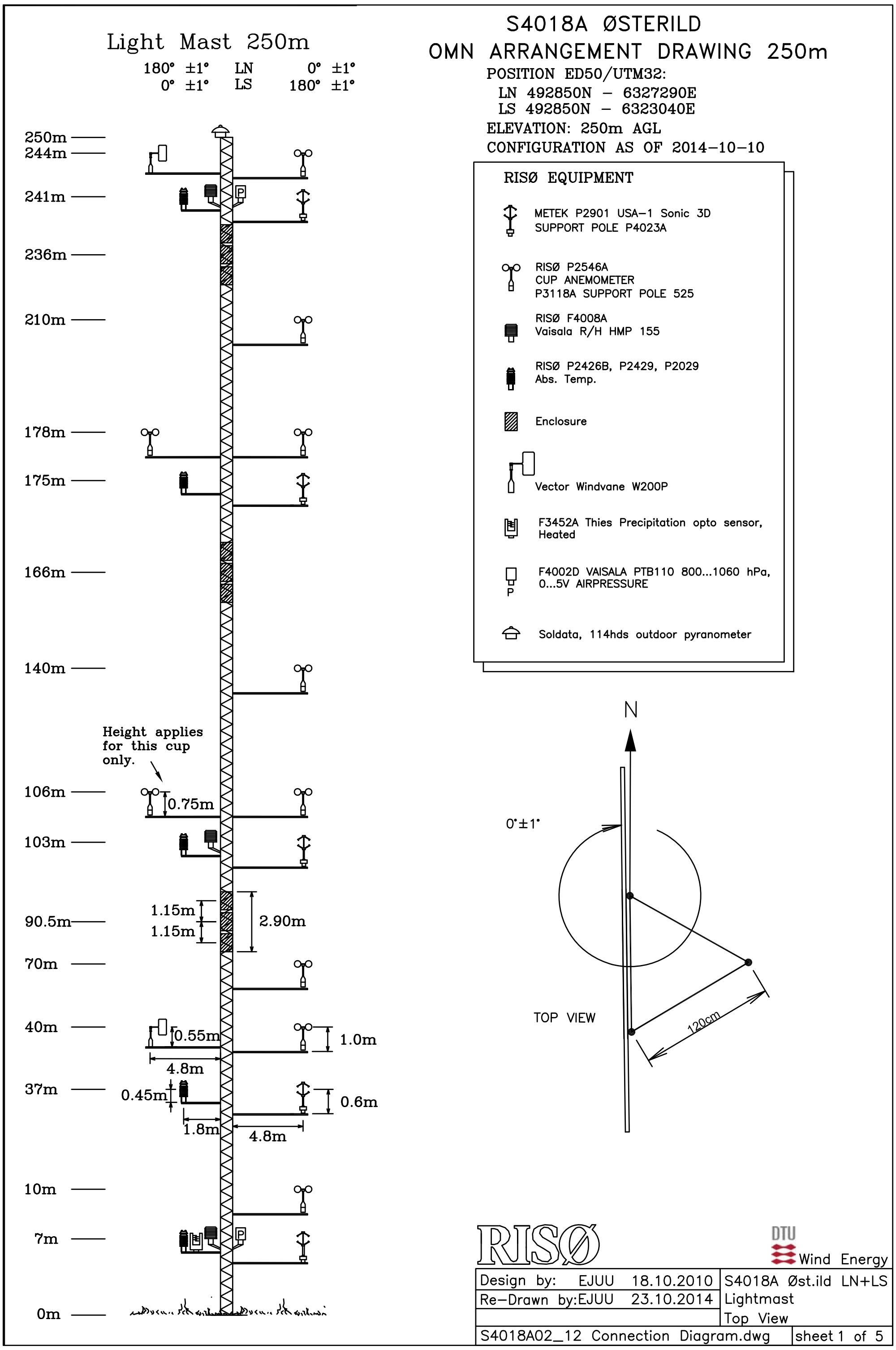

Figure 2: Drawing of light mast South

5 Data analysis

The following section focuses on two main topics - filtering of the data from the lidar units 590 and 542, filtering of the data from the sonic anemometers, and deriving additional statistics that were used later on to develop the methods for correction of lidar turbulence intensity (TI) measurements. The sonics are used for deriving turbulence parameters that would not be possible or not so easily obtained from a cup anemometer only. The sonic measurements are therefore the benchmark for validating the TI correction models. The cup anemometers are used only in the data filtering process for the two lidars.

5.1 Sonic data - extra computations and data filtering

The 20hz data from the sonics was post-processed to obtain additional statistics related to turbulence, as described in Peña [2019]. The sonic anemometers on the met mast provide the three wind components X_i, Y_i, and Z_i along the North-South, East-West and vertical direction, respectively, and the temperature T_i as well. The index "i" denotes a sample in the ten-minute period. The measurements from the sonic have a 2D correction applied; however for the analysis in this report we removed the 2D correction and applied the 3D correction suggested in Metek GmbH [2004]. Afterwards, by applying pitch and yaw rotations to the sonic measurements, we obtained the three components of the wind speed vector u, v and w. Thus, the u component is always aligned with the mean wind. Additionally, we obtained the variances and covariances of all three components of the wind speed vector. Finally, we calculated the friction velocity and the Obukhov length, as shown below:

$$u_* = (\overline{u'w'}^2 + \overline{v'w'}^2)^{1/4} \tag{4}$$

$$L_O = -T_i u_*^3 / (kg\overline{w'T_i'}) \tag{5}$$

In Eqn. 4 and Eqn.5, the prime $'$ denotes fluctuations, while the overline denotes a mean; k and g are the von-Karman constant and the gravitational acceleration, respectively.

It should be noted that not all 10min periods had all 12000 samples as we would expect when sampling at 20Hz. Therefore, we chose a filter that discards any 10min periods where the amount of samples are below 10000.

5.2 Lidar data - extra computations and data filtering

Even though the lidar provides ten-minute statistics for the horizontal wind speed, we produced the statistics used in the analysis. For the calculations, we applied two filters to the 1Hz data, namely we removed all samples where the parameter *datavalid* was not 0, and we removed all samples that were showing values of the horizontal wind speed higher than 100. After filtering, for each 10min period we calculated the following statistics for each of the four heights of interest (39m, 105, 177m and 243m):

- wind direction statistics - average, standard deviation, minimum, and maximum: Dir_{avg_l}, Dir_{std_l}, Dir_{min_l}, and Dir_{max_l}
- number of minutes with at least one valid sample (a valid sample is a sample that passes the filter for data validity and wind speed lower than 100), $N_{minutes}$
- the variance of the wind speed along the mean wind direction $\sigma^2_{u_l}$; this calculation is done in a similar way as we did with the sonic data - however, this time we used the wind direction measurement from the lidar itself.
- the variance of the wind speed component, transverse to the mean wind direction $\sigma^2_{v_l}$
- number of valid samples per 10min, N_{lidar}
- for unit 590, we had to calculate the average values for backscatter, cloud backscatter ratio and fog backscatter ratio as these were not provided in the statistical files. It should be noted that in these calculations we did not apply the aforementioned filters on the 1Hz data. For unit 542, the cloud backscatter ratio and fog backscatter ratios were not available in the files.

According to the recommendations in the lidar manual, the lidar average data is quality controlled i.e. the lidar can identify atmospheric conditions which affect adversely the measurements and rejects the data, if necessary. However, when we compared the lidar wind speed measurements to the cup anemometer measurements, we found that at different heights the agreement between lidar-mast is different and generally tends to deteriorate with height. Therefore, we did some investigation to find suitable filters based on the lidar internal parameters. In order to find proper quality filters for the ten-minute data, we relied on a comparison between the wind speed measurements of the lidar and the cup anemometers at 40m, 106m, 178m and 244m. We used data only from the West sector (wind directions between 240 and 300 deg) where the temperature was above 2 degC (in order to avoid icing on the cups) and the wind speeds were equal to or above 4m/s. For each ten-minute period we calculated the wind speed difference between lidar and cup (lidar deviations). First, we wanted to see the influence of the number of samples on the lidar deviations. Figure 3 shows, surprisingly, that the amount of data has very little influence. We chose to show the results only for two heights, but the general trend is that we get fewer valid samples for higher heights. In order to ensure reliable statistics without losing too much data, we chose to filter out data with N_{lidar} less than 20. This filter removes about 16% of the data at 244m. If we chose a limit of at least 30 valid samples, then we would lose 30% of the data at 244m. The value of 20 for the packets in average parameter is also a little bit above half of the total amount of samples we would get with this particular lidar configuration (it was configured to measure at 12 heights).

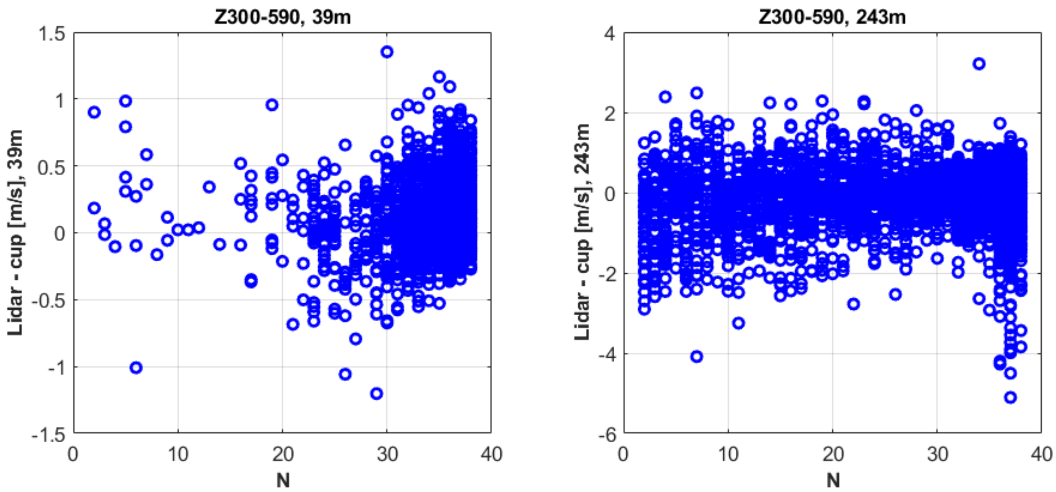

Figure 3: Lidar deviations as a function of N_{lidar}

After applying a filter of $N_{lidar} \geq 20$, it was clear that some of the deviations were still very high. In order to dive in a bit more, we plotted the lidar deviations as a function of the backscatter, the cloud backscatter ratio, and the fog backscatter ratio. Based on the results in Figure 4, 5, and 6 it is difficult to define a filter, or a combination of filters, which will allow us to remove only wrong data without discarding correct measurements as well. For example, it is obvious that most of the large deviations at 244m occur at low backscatter ratios (and cloud and fog backscatter ratios). However, it is not true that low backscatter ratios will always lead to a wrong measurement of the horizontal wind speed. Based on the data in Figure 6, we could potentially apply a filter of fog backscatter ratio higher than 0.0001; in case we apply it, we will be left with only 60% of the data.

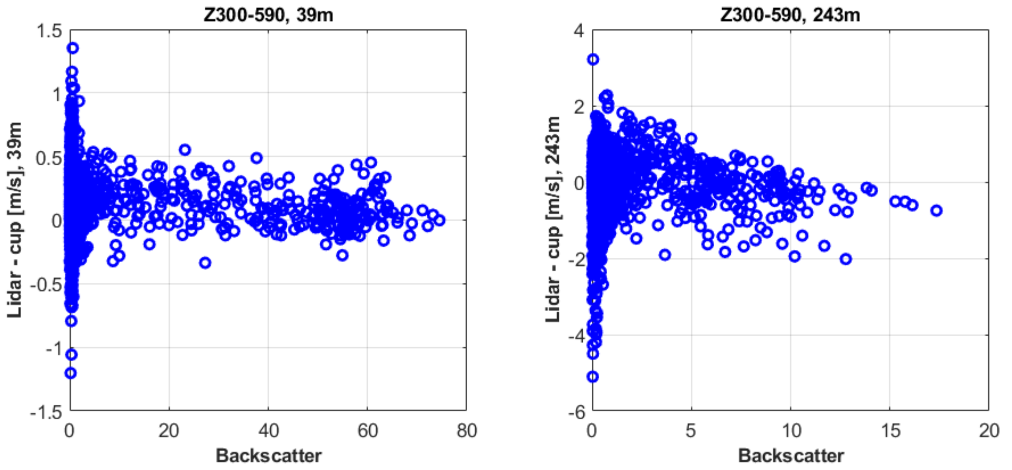

Figure 4: Lidar deviations as a function of backscatter

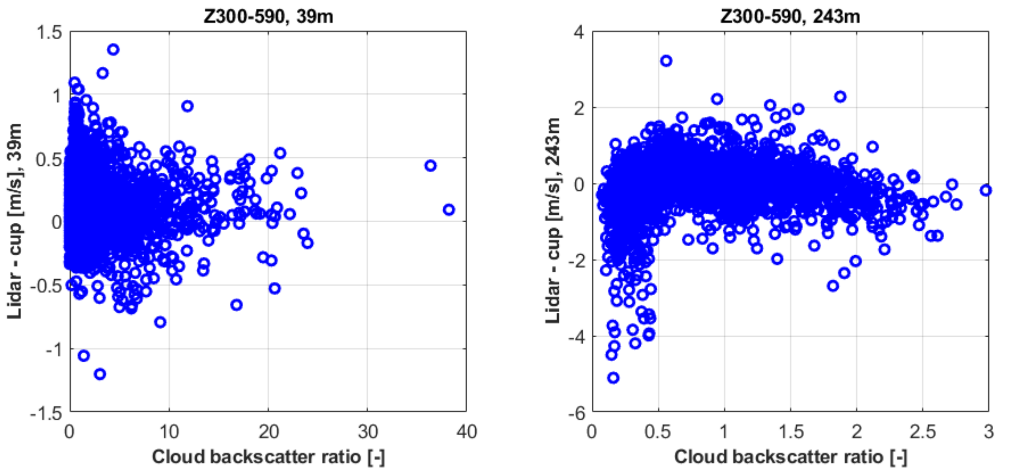

Figure 5: Lidar deviations as a function of cloud backscatter ratio

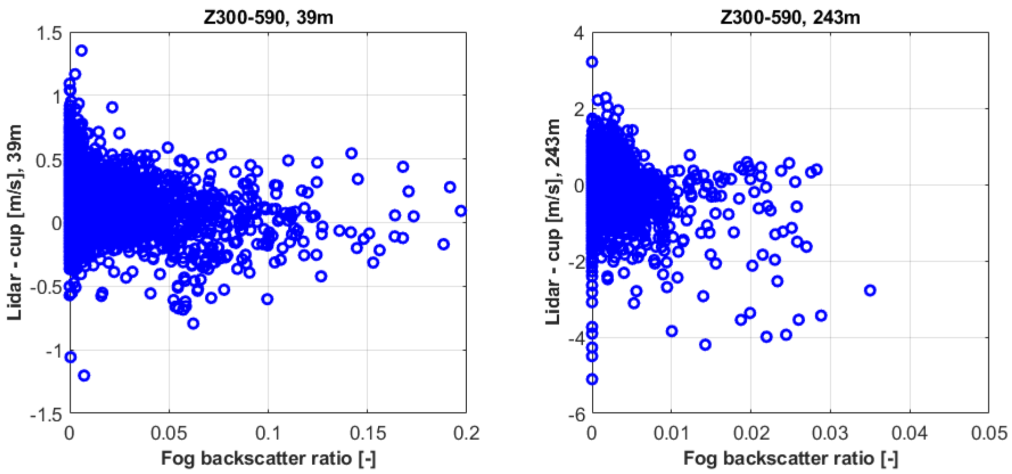

Figure 6: Lidar deviations as a function of fog backscatter ratio

Therefore, we decided not to use the lidar's own internal parameters for quality filtering. The final choice of filters for the lidar data is presented below:

1. Filters based on lidar data
   (a) $N_{lidar} > 20$
   (b) $N_{minutes} = 10$ - this filter ensures that there are no gaps of data in any 10min period, where the lidar is not measuring for more than a minute
   (c) $Dir_{std_l} < 40$deg - this filter ensures that there is no ambiguity in the wind direction in any of the samples
   (d) Dir_{max_l} - $Dir_{min_l} < 150$deg or Dir_{max_l} - $Dir_{min_l} > 210$deg - this filter ensures that there is no ambiguity in the wind direction in any of the samples
2. Filters based on metmast data
   (a) $N_{sonic} > 10000$
   (b) Dir_{avg_s} between [60,120]deg or [240, 300]deg

   The following two filters were applied in order to avoid negative shears and as a consequence - unrealistic turbulence length scales:

   (c) $spd_{sonic_{7m}} < spd_{sonic_{37m}} < spd_{sonic_{103m}} < spd_{sonic_{175m}} < spd_{sonic_{241m}}$
   (d) $\delta U/\delta z > 0$, for all heights
3. Filters based on both lidar and metmast data
   (a) $|spd_{lidar} - spd_{cup}| < 1$ - this filter is based on DTU's experience with lidar calibrations; it is expected that the difference in wind speed between lidar and mast does not exceed 1m/s at any height
   (b) Dir_{avg_l} - $Dir_{avg_s} < 20$deg - this filter is applied in order to avoid wind direction ambiguity in any of the lidar measurements.
4. Other filters
   (a) the period between 13-Mar-2018 18:40 and 13-03-2018 21:40 was removed because of faulty sonic measurements

After applying these filters, we inspected visually all high frequency data from both sonics and lidars that passed the filters. In the case of unit 542, we found that there were periods where one or a few of the lidar wind speed measurements were much different from the sonic wind speed measurements within the same period. One such example can be seen in Figure 7 around 19:30, where the sonic wind speed is illustrated with blue color and the lidar wind speed is illustrated with red color, for the four different heights.

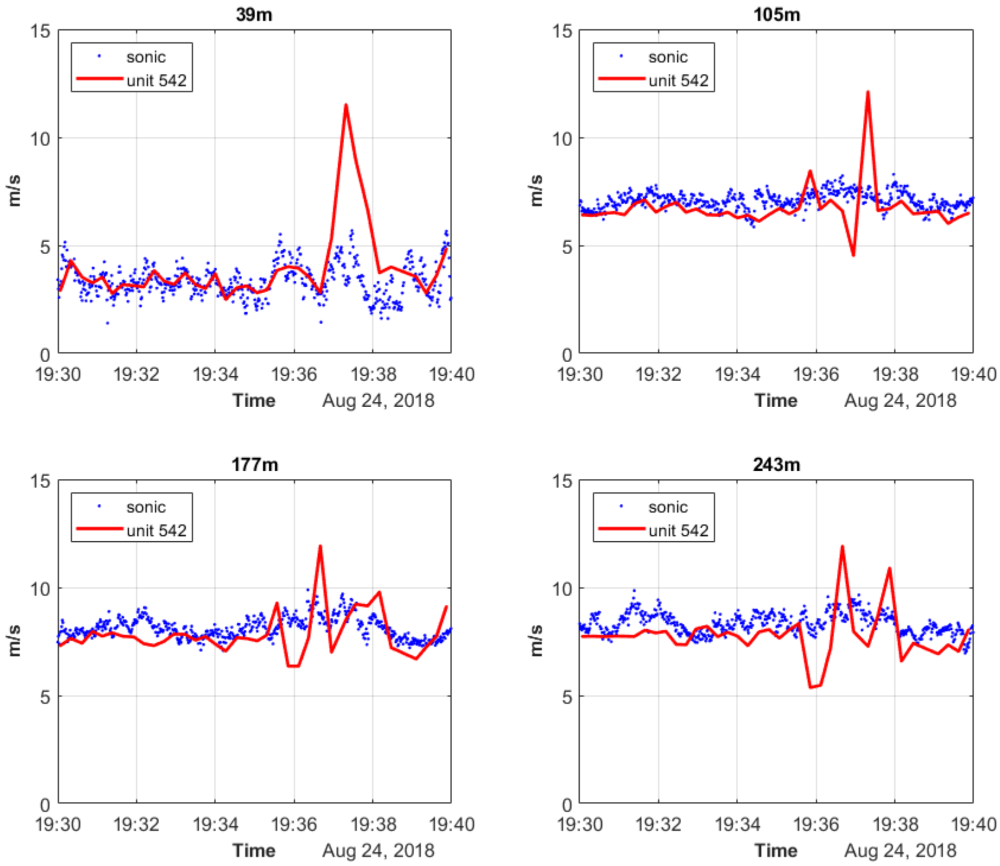

Figure 7: Time series plot of wind speeds from unit 542 and the sonics

5. Additional time filters for unit 542:

   24-Aug-2018 19:30

   25-Aug-2018 20:40

   between 26-Aug-2018 04:40 and 26-Aug-2018 04:50

   07-Sep-2018 15:50

   07-Sep-2018 16:10

   13-Sep-2018 17:00

   14-Sep-2018 07:20

   between 14-Sep-2018 00:30 and 14-Sep-2018 00:40

   22-Sep-2018 15:50

   28-Dec-2018 19:10

The five filters mentioned above removed a significant part of all available data. The total result is presented in Table 2.

Amount of data	unit 590	unit 542
Before filtering	18743	26139
After filtering	3844	3519

Table 2: Amount of data available from the two lidar units before and after filtering

5.3 Cup anemometer and sonic turbulence intensity

Even though the sonic anemometer can capture higher frequencies in the wind speed fluctuations than the cup anemometers, both sensors measure very similar turbulence intensity levels. Figure 8 shows a comparison of the TI measured by the two instruments at 40m, 106m, 178m, and 241m. The cup anemometer TI is calculated as the standard deviation divided by the mean value for each ten minute period. The sonic TI is obtained by dividing the σ_{u_s} by the scalar average of the horizontal wind speed for every ten minutes - in other words this is the turbulence intensity of the longitudinal component of the wind speed. As we can see from Figure 8 the TI measured by the cup and sonic correlate very well. The mean difference of the entire dataset between the two sensors is low - at 40m it is 0.7% and for the other three heights it is close to 0%.

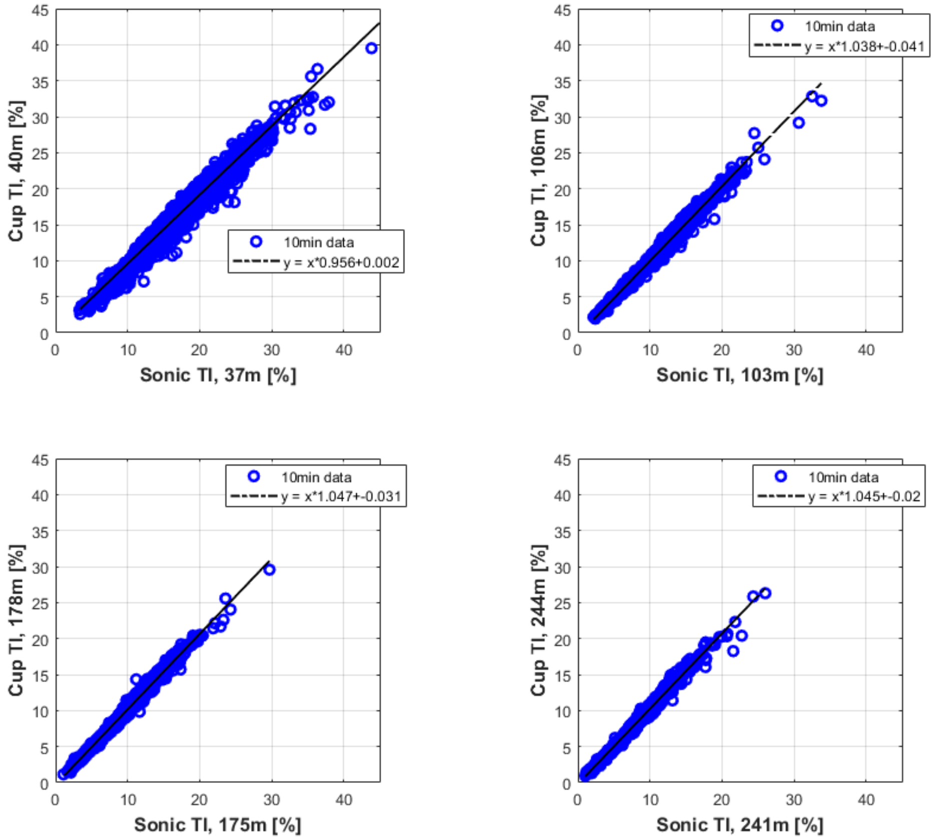

Figure 8: Comparison of TI measured by a cup anemometer and a sonic anemometer

5.4 Stability conditions

For the analysis presented in this section and in Sect. 7.1, we classify the unit 590 and sonic-anemometer data within three atmospheric stability conditions (stable, neutral, and unstable) based on the dimensionless stability parameter z/L_O computed from the sonic anemometer at 37 m. And within each stability class, we further classify the data into three categories (small, medium, and large length-scale ranges) based on an approximation of the Mann turbulence length scale, i.e., $L = \sigma_{u_s}/\left(dU/dz\right)$ using the sonic-based turbulence and mean vertical wind shear measures. The different ranges used for classification into atmospheric stability and length-scale categories are presented in Table 3, together with the amount of 10-min samples for each case.

Scatter plots of 10-min means of horizontal velocity of unit 590 and the sonic anemometers at different heights are shown in Figure 9 to illustrate the good agreement between the horizontal wind speed measured by the lidar and the sonics. Figure 10 shows scatter plots of the 10-min along-velocity variance between the lidar and the sonic anemometer measurements at the matching vertical levels. We use logarithmic scales, since a large amount of 10-min values concentrate close to zero. As illustrated, stable atmospheric conditions generally show less along-wind variance compared to neutral and unstable atmospheric conditions at all matching vertical levels. The highest variance values appear under neutral atmospheric conditions, where winds are also generally the highest. Also, under stable atmospheric conditions the lidar variance measure is clearly lower than the sonic measure within the first two vertical levels in particular, whereas the lidar variance seems to be closer to the sonic measure under unstable atmospheric conditions.

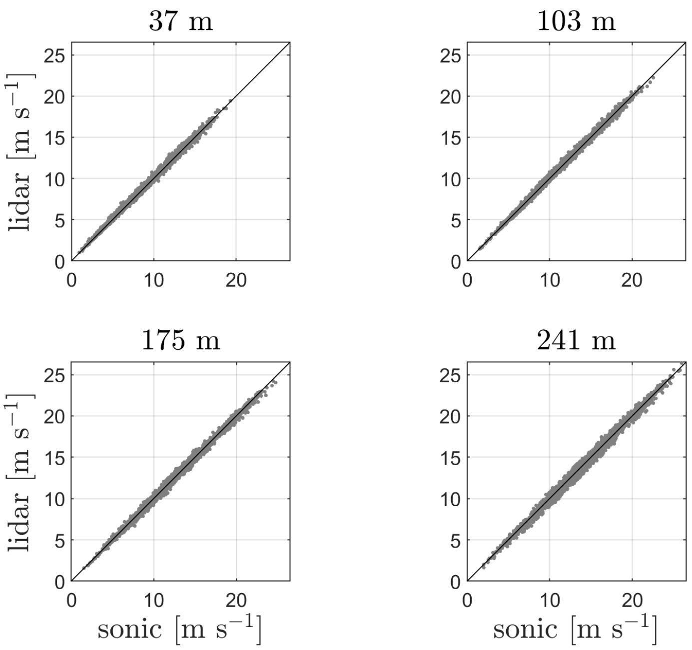

Figure 9: 10-min means of horizontal velocity measured by the sonic anemometers and unit 590 for each of the matching vertical levels

Table 3: Median value of the Mann turbulent length scale L in meters within the atmospheric-stability and length-scale categories for the four sonic-lidar matching heights (in bold). Number of 10-min samples within each category is also given

stability range	stable $z/L_O \geq 0.10$			neutral $-0.05 < z/L_O < 0.10$			unstable $z/L_O \leq 0.05$		
	sta 1	sta 2	sta 3	neu 1	neu 2	neu 3	uns 1	uns 2	uns 3
length scale	small	medium	large	small	medium	large	small	medium	large
range [m]	$L \leq 10$	$10 < L < 16$	$L \geq 16$	$L \leq 23$	$23 < L < 30$	$L \geq 30$	$L \leq 35$	$35 < L < 45$	$L \geq 45$
# 10-min	488	525	518	555	598	643	196	163	158
39	6.77	13.14	19.39	20.12	26.12	35.58	27.76	39.68	52.82
105	12.91	22.86	35.96	34.55	48.30	76.01	54.12	95.05	140.24
177	17.20	28.09	47.09	41.22	62.65	106.67	73.10	149.49	212.95
243	20.22	30.67	52.08	44.06	72.39	130.17	86.34	195.27	283.46

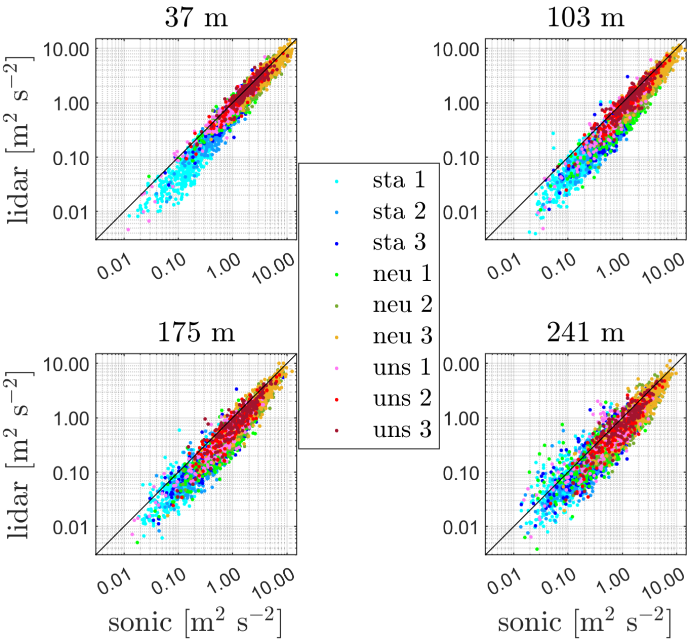

Figure 10: 10-min along-velocity variances measured by the sonic anemometers and unit 590 for each of the matching vertical levels. The different categories are color-coded and refer to those in Table 3

6 Methodologies for correcting lidar turbulence measurements

The correction of TI measurements from lidars implies providing an answer for the so-called lidar turbulence paradox. Here, we do not attempt to answer the paradox as we do not have the resources to do so within this project. Instead, our approach is the following:

1. We use measurements from a ground-based CW lidar to evaluate the goodness of a physics-based lidar-turbulence model. We use lidar and mast measurements analysed and classified in Section 3 to serve as the basis of the evaluation of the model.

2. If we find that the physics-based lidar-turbulence model performs well against the measurements in point 1), then we use this physics-based lidar-turbulence model to develop a physics-based neural network (PBNN) model and we can continue with point 3. Otherwise, we can go directly to step 4).

3. The PBNN model is then evaluated first against simulated turbulence data from the physics-based lidar-turbulence model, which served as a basis of the PBNN model. The PBNN model can be seen as a sophisticated look-up table (LUT). In this step, the assessment of the model is performed by using K-fold cross-validation so that we do not assess the model performance using the same data. We test multiple PBNN models, derived from different combinations of inputs to the NN, and in this fashion we can study the importance of the predictors (inputs).

4. Whether the physics-based lidar-turbulence model performs well or not in point 1), we also develop a data-driven neural network (DDNN) model, which is based on the lidar measurements analysed in Section 3 only.

5. As in point 3), we can evaluate the DDNN model against the lidar measurements themselves using a K-fold cross-validation. We can also test multiple DDNN models, derived from different combination of inputs to the NN and study the importance of the predictors.

6. We can then evaluate the DDNN model against another set of measurements from an independent lidar, which have not been used to develop the DDNN model.

7. If the PBNN model was developed, we can also evaluate it against the two independent sets of lidar measurements.

6.1 The physics-based lidar-turbulence model

The physics-based lidar-turbulence model has been already described by Sathe et al. [2011]. They derived expressions for the velocity spectra measured by a velocity azimuth display (VAD) lidar. The ZX 300 lidar we study here performs conical scanning of winds in a VAD mode and therefore the vertical velocity spectra can theoretically be described by:

$$F_w^l(k_1) = \left(\frac{1}{\cos^2\phi}\right)\widehat{T_f}(k_1)\iint \Phi_{ij}(\boldsymbol{k})\alpha_i(\boldsymbol{k})\alpha_j^*(\boldsymbol{k})dk_2dk_3, \tag{6}$$

where α is a spectral weighting function (* means complex conjugation), ϕ the lidar half-opening angle (30° for a ZX 300 lidar), $\widehat{T_f}$ a spectral transfer function accounting for the low-pass filter effect due to the time the lidar takes to scan the cone. Since the ZX 300 takes 1 s to scan the 50 azimuthal positions, we will assume the effect of the low-pass filter is negligible.

For the u and v velocities, the spectrum is similar to that in Eqn. (6) but the term $\cos^2\phi$ in the denominator should be replaced by $\sin^2\phi$ and the weighting function α by β and γ, respectively. For a CW lidar, these weighting functions are

$$\begin{aligned}
\alpha_i(\boldsymbol{k}) &= \frac{1}{2\pi}\int_0^{2\pi} n_i(\theta)\exp\left(1id_f\boldsymbol{k}\cdot\boldsymbol{n}(\theta)\right)\exp\left(-z_R|\boldsymbol{k}\cdot\boldsymbol{n}(\theta)|\right)d\theta, && (7)\\
\beta_i(\boldsymbol{k}) &= \frac{1}{\pi}\int_0^{2\pi} \cos\theta n_i(\theta)\exp\left(1id_f\boldsymbol{k}\cdot\boldsymbol{n}(\theta)\right)\exp\left(-z_R|\boldsymbol{k}\cdot\boldsymbol{n}(\theta)|\right)d\theta, && (8)\\
\gamma_i(\boldsymbol{k}) &= \frac{1}{\pi}\int_0^{2\pi} \sin\theta n_i(\theta)\exp\left(1id_f\boldsymbol{k}\cdot\boldsymbol{n}(\theta)\right)\exp\left(-z_R|\boldsymbol{k}\cdot\boldsymbol{n}(\theta)|\right)d\theta, && (9)
\end{aligned}$$

where $\boldsymbol{n}(\phi, \theta) = (\cos\theta \sin\phi, \sin\theta \sin\phi, \cos\phi)$ is the unit vector describing the lidar scanning pattern with θ being the azimuthal positions, d_f the focused distance, and z_R the Rayleigh length [Sonnenschein and Horrigan, 1971] that characterizes the length of the probe volume and is estimated as

$$z_R = \frac{\lambda d_f^2}{\pi a_0^2}, \tag{10}$$

where λ is the laser wavelength and a_0 the effective beam radius at the output lens.

By combining Eqns. (6)–(10), we can compute the velocity spectra. By integrating the spectra (see Eqn. 3), we can obtain the velocity variances. The ZX 300 lidars used by Energinet provide horizontal velocity standard deviations for each 10-min interval and therefore we can in principle compare the modelled lidar turbulence with the lidar measurements. We only need to know how to model Φ_{ij}. Here, we use the three-dimensional spectral tensor turbulence model by Mann [1994] (hereafter Mann model) to describe Φ_{ij}. The Mann model contains three parameters besides the wavenumber vector $\boldsymbol{k}$; the dissipation rate of turbulence $\alpha\epsilon^{2/3}$, the turbulence length scale L, and the turbulence anisotropy Γ. For a detailed analysis of the behavior of the Mann parameters in the atmosphere under a range of turbulence, atmospheric stability, and wind speed conditions, we refer to Peña et al. [2010], Kelly [2018], and Peña [2019].

Figure 11-left shows the lidar to sonic standard deviation ratios for the three velocity components and for four different heights. The lidar standard deviations are computed via using Eqns. (6)–(10) (using the configuration and specifications of the ZX 300 lidar) and the sonic/ideal standard deviations via Eqn. (2). In both cases Φ_{ij} is described with the Mann model using $L = 50$ m and $\Gamma = 3$. The value of $\alpha\epsilon^{2/3}$ is irrelevant as it only scales turbulence. In Figure11-right, the ratio of the v- to the u-velocity standard deviation is shown for both the sonic and the lidar case.

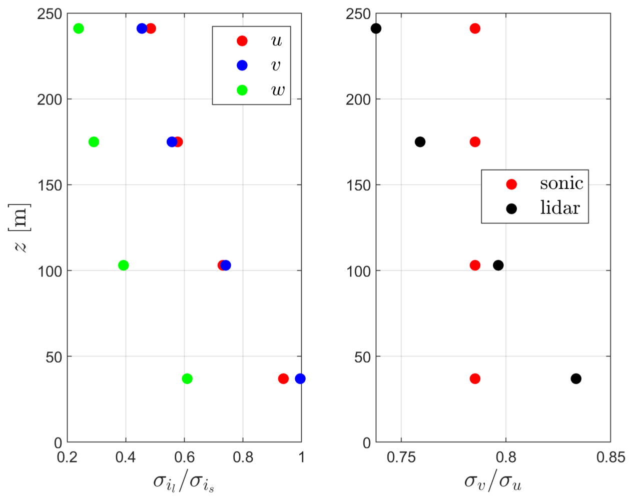

Figure 11: Lidar to sonic ratios of the velocities' standard deviation as function of height (left frame) and v- to u-velocity standard deviation ratio (right frame)

As illustrated, all ratios of velocity components' standard deviations are lower than one, so all lidar derived turbulence values are lower than the sonic values, with the v-ratio being the closest to one at 37 m. The ratios decrease with height partly because of probe-volume averaging: the length of the probe volume increases with height and so the degree of filtering. Since we use the same Mann parameters at all heights, the ratio σ_v/σ_u is the same at all heights for the sonic, whereas it decreases with height for the lidar and this is due to the complex interaction of the components of the spectral tensor with the lidar weighting function.

6.2 The physics-based neural network model

The physics-based turbulence model in Sect. 6.1 is computationally expensive and paradoxical as it needs information about turbulence to understand the behaviour of the lidar-to-sonic turbulence ratios at the different heights. We, therefore, construct a dataset of model-based outputs from the physics-based turbulence model. We want to construct this dataset so that the output of the physics-based turbulence model covers the range of lidar turbulence measurements, which are used to predict the sonic turbulence measurements. Thus, we use information from the measurements as proxy to establish a range of Mann parameters that show a similar turbulence 'climatology'.

We take the different heights separately as ZX lidar turbulence is height dependent. We use measurements of the unit 590 lidar to determine approximately the observed turbulence ranges. Specifically, we use the lidar horizontal velocity standard deviation recorded each 10-min period to derive a range of $\alpha\epsilon^{2/3}$ values assuming $\alpha\epsilon^{2/3} \approx 7.5\sigma^2_{u_l}/z^{2/3}$. We use a similar approximation for the turbulence scale of the Mann model by Kelly [2018], i.e., $L \approx \sigma_{u_l}/dU/dz$. As shown in Sect. 5, the mean wind speeds of the lidar agree very well with those of the sonic anemometers at the different matching heights and so we assume that the vertical wind speed gradient dU/dz from lidar measurements is as good as that based on the sonic measurements. However, we use the sonic-based gradient because we can use the lowest sonic height on the mast (7m) to properly determine the gradient at 37m; the first lidar measurement is at 39m and, thus, estimating the gradient at the first level can result in high inaccuracies. For future work, we suggest to always measure below the first height needed for turbulence estimation.

3844 samples (corresponding to the entire dataset of unit 590 10-min measurements) of estimated L and $\alpha\epsilon^{2/3}$ values are then ready to determine turbulence variances. We only need information about Γ, which is somehow a more complex parameter. We can now vary Γ within the range 1.5–4.0 from the spectral velocity analysis performed by Peña [2019] at Østerild. We then construct a preliminary dataset of 3844×6[1], which we use to generate sonic-based velocity variances. This is a much faster procedure than the computation of lidar-derived velocity variances as we can use the pre-computed LUT of Mann-based turbulence, i.e., Eqns. (2) and (3) with Φ_{ij} described by the Mann model. We randomly select 1000 out of the 21882 possible atmospheric states and compare the u-variance from the computation and the measured u-variance from the sonic anemometer. Figure 12 shows this random selection for the dataset derived for the height 37 m. We also show the original sonic-based histogram of observations of σ_{u_s} and, as illustrated, both show a much higher frequency of samples in the low turbulence range with decreasing samples the larger the turbulence level.

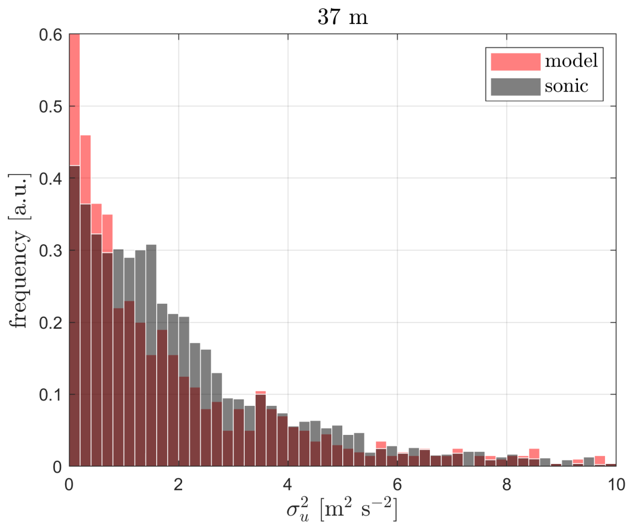

Figure 12: Histograms of the u-variance from the measurements at a height of 37 m from the sonic anemometer and from the model. Histograms are normalized

We only select 1000 states as the lidar-based turbulence calculations, i.e., Eqns. (6)–(10), are rather computationally expensive. The performance for a higher number of states was not evaluated - we assume that 1000 states are enough to cover the turbulence climatology at the site, for each height and for that specific period of time. We generate a new dataset with 1000 samples that includes the lidar-based velocity

[1]6 from varying $\Gamma = (1.5 : 0.5 : 4.0)$

variances. A similar procedure is performed at all lidar measurement heights and so we have four datasets of Mann parameters, each corresponding to a particular observed vertical level.

We are now ready to construct the PBNN model. It is out of the scope of the work to test different NNs and the different characteristics of these (e.g., the number of hidden layers or neurons). However, first trials, which were performed on the physics-based datasets only, revealed that NNs were superior than a number of tested decision trees using different algorithms for classification and regression. From these first trials, the performance of the NNs did not increase using more than 20 neurons or using more complex layer structures. Also, out of two algorithms tested, Levenberg-Marquardt and Bayesian regularization, it appeared that the later was superior. We therefore use the 'default' NN `fitnet` for function fitting problems of Matlab [The MathWorks Inc., 2022]. This is a feedforward NN and we use the architecture shown in Figure 13, i.e., one hidden layer with 20 neurons and one output neuron because we only need one response value: the lidar u-variance.

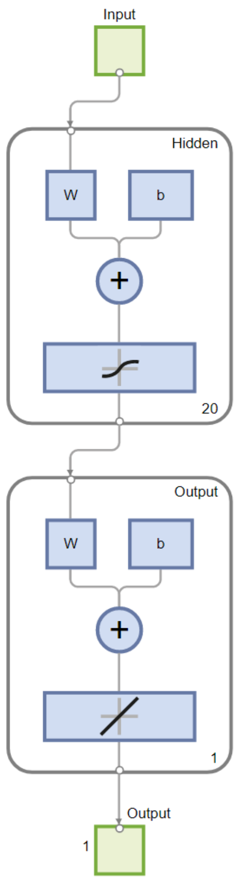

Figure 13: Shallow neural network architecture used to estimate lidar turbulence with one hidden layer including 20 neurons

6.2.1 Cross-validation

The PBNN model is cross-validated using simulated turbulence data from the physics-based lidar-turbulence model, which served as a basis of the PBNN model. The assessment of the model is performed by using K-fold cross-validation. We test multiple PBNN models, whose solely difference is the number of inputs to the NN in Figure 13. In this way we can study the importance of the inputs (predictors).

We use a type of a K-fold cross-validation. The 1000-samples dataset is randomly permuted and subdivided into a training and testing subset, with 90% of the dataset going to training and 10% to testing. We train the NN with the training subset, generate predictions using different inputs of the testing subset, and, finally, compare the model predictions with the known value of the prediction, in this case the u-velocity variance. We compute the root mean square error (rmse), mean absolute percentage error (mape), and the slope of a linear fit through origin between the prediction and the 'real' value to evaluate the model perfomance. This process is repeated 100 times to build statistics. Figure 14 illustrates the result of one of the hundred cross-validations.

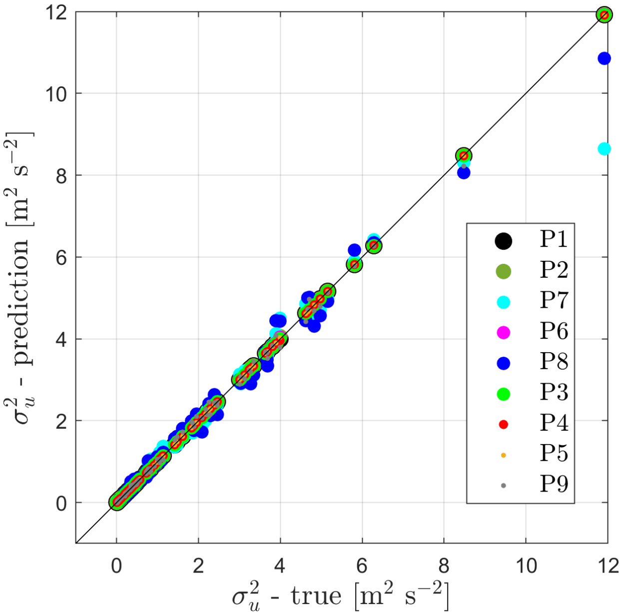

Figure 14: An example of a cross-validation of the predictions and true values of the u-variance using the PBNN model. Different inputs (P) are used and described in Table 4

As illustrated, different inputs (P) are tested and described in Table 4. The inputs are basically a combination of all possible input variables one could eventually provide to the physics-based lidar-turbulence model, i.e., Γ, L, and $\alpha\epsilon^{2/3}$, together with the computations of the lidar turbulence values $\sigma_{u_l}^2$, and $\sigma_{v_l}^2$. The results of the model performance for each of the tested inputs are shown as box plots in Figure 15. As shown in the latter figure, mape and rmse are much higher for the last three sets of inputs (P7–P9), which only use two or less input parameters: P9 median mape, rmse, and slope is 10.44%, 0.15 m^2 s^{-2}, and 0.995, which are fairly good statistics (the slope median for all sets of inputs is always very close to one). P6 is an interesting case, since we could have means to compute the three input parameters with the ZX 300 lidar and as shown its statistics are within the data inputs with median mapes far below 10% and median rmses below 0.01 m^2 s^{-2}. P1 and P5 show the lowest mape and rmse, with P5 being the only input set with mape below 1%. In principle, if we know the lidar probe volume characteristics and the scanning height, the three inputs used by P5 are sufficient for the computation of the lidar u-variance and so it is not surprising that this shows the best statistics. When adding the lidar turbulence to these three inputs (as we do in P1), the model might become overfitted degrading its statistics.

Table 4: Description of the different training input datasets for the PBNN model

Input short name	Input variables
P1	Γ, L, $\alpha\epsilon^{2/3}$, $\sigma^2_{u_l}$
P2	Γ, L, $\sigma^2_{u_l}$
P3	L, $\alpha\epsilon^{2/3}$, $\sigma^2_{u_l}$
P4	Γ, $\alpha\epsilon^{2/3}$, $\sigma^2_{u_l}$
P5	Γ, L, $\alpha\epsilon^{2/3}$
P6	L, $\sigma^2_{u_l}$, $\sigma^2_{v_l}$
P7	$\sigma^2_{u_l}$, $\sigma^2_{v_l}$
P8	$\sigma^2_{u_l}$
P9	L, $\sigma^2_{u_l}$

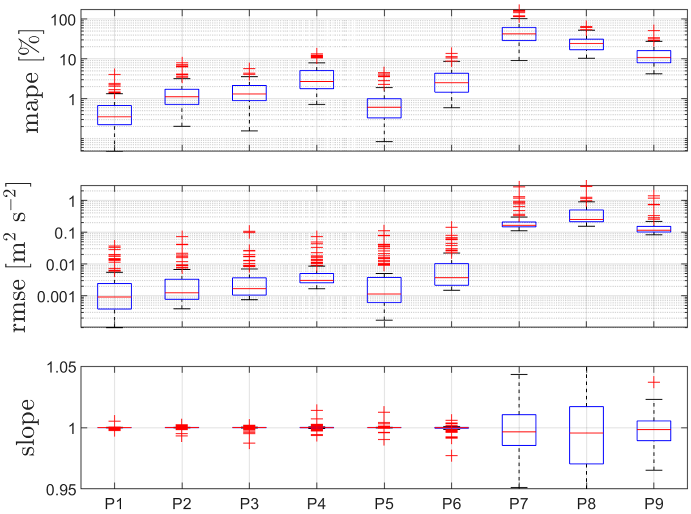

Figure 15: Performance of the PBNN model for a different set of inputs (see Tab. 4). Red solid line represents the median of 100 randomly tested datasets

When applying the neural network to actual lidar measurements, we can only use the combination of inputs described in P6, P7, P8, and P9. The Mann model parameters Γ and $\alpha\epsilon^{2/3}$ cannot be derived from the lidar data.

6.3 The data-driven neural network model

A DDNN model is also constructed in a similar way as the PBNN model in Sect. 6.2. However, instead of using 1000 possible atmospheric states, we use the 3844 samples corresponding to the entire dataset of unit 590 10-min measurements to construct four NNs, each corresponding to a different measurement height. The structure and the type of NN is the same as that used for the PBNN model.

6.3.1 Cross-validation

The DDNN model is cross-validated using the same methodology as that we employed for validating the PBNN model in Sect. 6.2.1. Here, each of datasets of measurements (one per height), each with 3844 10-min statistics is randomly permuted and subdivided in a training and testing subsets, the NN is trained, and the generated predictions are compared with the known 'true' value, i.e., the sonic-anemometer u-variance measurements matching the height of the lidar measurement. This process is also repeated 100 times. Figure 16 illustrates the result of one of the hundred cross-validations, where we show that no matter what we choose as inputs, the DDNN-based models predict turbulence with some degree of uncertainty.

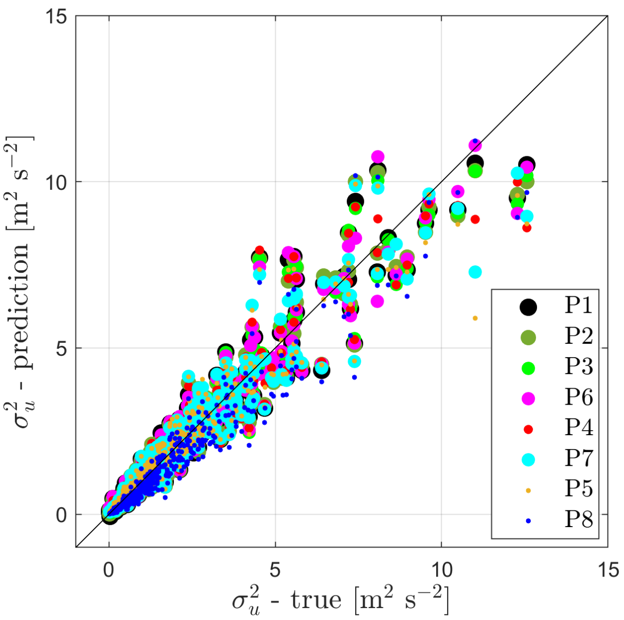

Figure 16: An example of a cross-validation of the predictions and true values of the u-variance using the DDNN model. Different inputs (P) are used and described in Table 5

We use different inputs combining possible variables that we can measure with a lidar: the lidar mean horizontal velocity, the lidar-derived along-wind and transverse velocity variances, the mean vertical wind shear, and a lidar-based turbulent length scale, i.e., U_l, $\sigma_{u_l}^2$, $\sigma_{v_l}^2$, dU/dz, and $\sigma_{u_l}/dU/dz$, respectively (see Table 5). The results of the model performance for each of the tested inputs are shown as box plots in Figure 17. Note that we provide also results (P8) for the case in which the prediction is the same as the input variable, i.e., we assume the true u-variance is the same lidar u-variance (persistence).

Table 5: Description of the different training input datasets for the DDNN model

Input short name	Input variables
P1	U_l, $\sigma_{u_l}^2$, $\sigma_{v_l}^2$, dU/dz, $\sigma_{u_l}/dU/dz$
P2	U_l, $\sigma_{u_l}^2$, $\sigma_{v_l}^2$
P3	$\sigma_{u_l}^2$, $\sigma_{v_l}^2$, dU/dz, $\sigma_{u_l}/dU/dz$
P4	$\sigma_{u_l}^2$, $\sigma_{v_l}^2$
P5	$\sigma_{u_l}^2$
P6	$\sigma_{u_l}^2$, $\sigma_{v_l}^2$, dU/dz
P7	U_l, $\sigma_{u_l}^2$
P8	persistence (not a NN)

As illustrated, persistence has the highest median mape and rmse and the lowest mean slope, which basically means that assuming the lidar u-variance as a proxy for the sonic u-variance is not the best choice at all. Interestingly, when including the lidar v-variance (P2 vs P7 or P4 vs P5), the statistics worsen and adding the mean wind speed does not seem to help either (P5 vs P7). Also note that all the median slopes are below one so all the DDNN models predict a sonic u-variance very close but lower than the measured value.

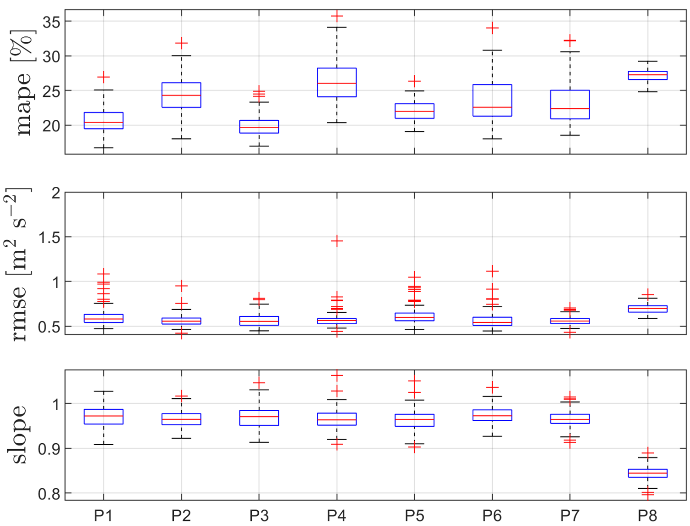

Figure 17: Performance of the DDNN model for a different set of inputs P (see Tab. 5). Red solid line represents the median of 100 randomly tested datasets

7 Results

7.1 Physics-based lidar-turbulence model

Using the physics-based lidar-turbulence model in Sect. 6.1, we predict the ratios of the lidar-to-sonic horizontal-velocities' variances for the four lidar-sonic matching heights. We compare these predictions with the medians of the measurements analyzed in Sect. 5 for unit 590 within the different atmospheric-stability and length-scale categories. Based on the turbulence model, these ratios are independent on the Mann parameter $\alpha\epsilon^{2/3}$ as this only scales the turbulence level. Since we only want to get an idea of the goodness of the physics-based lidar-turbulence model, we use the best estimates of L values, in this case the medians of the sonic-derived L values (see Table 3), and we assume $\Gamma = 3.0$ for all atmospheric-stability and length-scale cases, and heights. The assumed value is a good approximation for the site of Østerild.

7.1.1 Neutral atmospheric conditions

Figure 18 illustrates the results of the model predictions and the measurements for neutral atmospheric conditions and the three length-scale categories. On Figure 18-left, which portraits the findings for the u-variance, the general behavior with height of lidar-to-sonic turbulence is shown for both model results (in triangles) and measurements (in circles): the lidar-based turbulence is lower than the sonic-based turbulence, with the model always predicting a decrease of the lidar-to-sonic variance ratio with height. Model results are close to the observations at the first two heights, and the variances' ratios from the observations do not change much between the two highest heights. At these two heights, the model results show the largest differences with the measurements. As expected, both observations and model results show larger ratios (closer to one) within the category with largest length scales (cat 3).

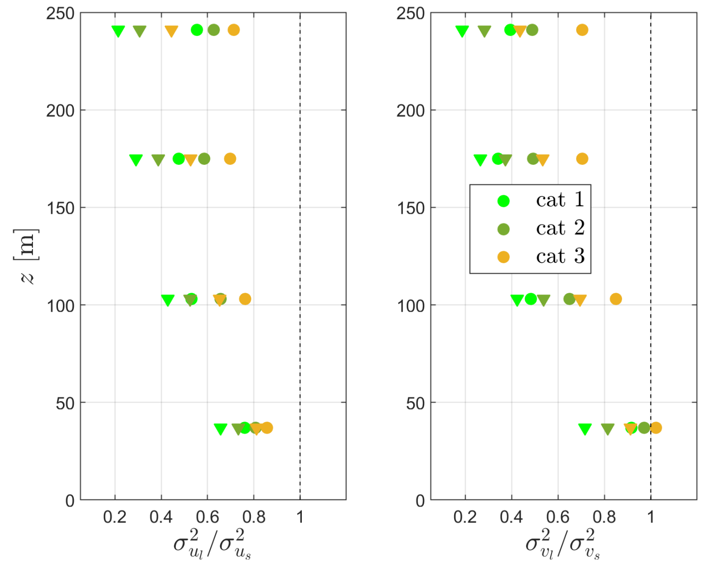

Figure 18: Lidar-to-sonic ratio of the u-variance (left) and v-variance (right) as function of height under neutral atmospheric conditions for the three length-scale categories (Table 3, cat 1 - small, cat 2 - medium, cat 3 - large length scales). Model results and measurements are shown in triangles and circles, respectively. Colors are in agreement with those of Fig. 10

The comparison between model results and measurements for the case of the v-variance (Figure 18-right) shows similarities with respect to that of the u-variance. However, within the first two vertical levels, both model results and measurements clearly show larger variance ratios, which includes a case with a variance ratio larger than one that corresponds to measurements under the larger length-scale category. For these type of lidars and within the first tens of meters from the ground, the lidar-to-sonic v-variance ratio tends to be larger than the u-variance ratio [Sathe et al., 2011].

7.1.2 Stable atmospheric conditions

Figure 19 illustrates the results of the model predictions and the measurements for stable atmospheric conditions for the three length-scale ranges. As illustrated, both model results and measurements generally show lower lidar-to-sonic variance ratios under all stable length-scale categories when compared to the results under neutral atmospheric conditions in Figure 18. The accuracy of the model prediction is generally worse for all stable categories compared to the neutral cases, with better model predictions for the cases with larger length scales compared to those with lower length-scale vales. At 37 m, both measurements and model results show that within the lower length-scale range, the lidar only measures about 40% of the sonic-based velocity variances, whereas both measurements and model results are always above ≈65% at 37 m for all length-scale categories under neutral conditions.

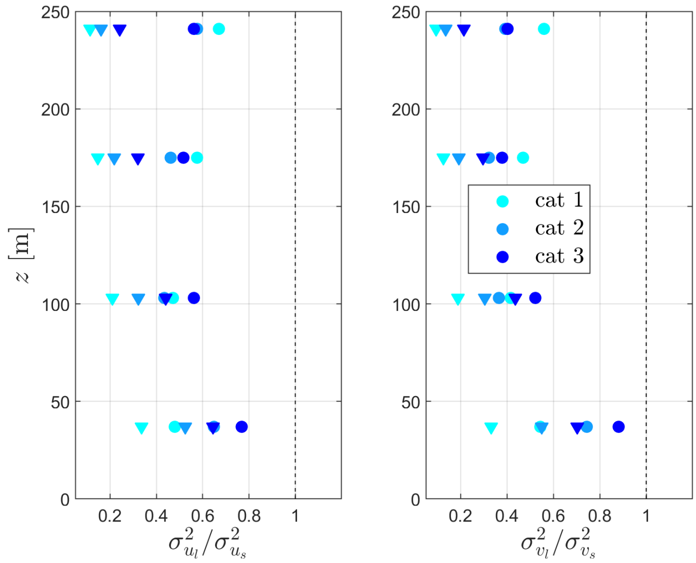

Figure 19: Similar to Fig. 18 but for stable atmospheric conditions

Within stable atmospheric conditions, the differences between model results and measurements are larger at the two highest levels, particularly of the u-variance ratios. Also, within stable conditions at these two heights, we clearly notice a 'recovery' of the lidar velocity variances on the measurements, especially for the case with the smallest length scale, which the model fails to predict.

7.1.3 Unstable atmospheric conditions

Figure 20 illustrates the results of the model predictions and the measurements for unstable atmospheric conditions for the three length-scale ranges. As illustrated, both model results and measurements generally show the highest lidar-to-sonic variance ratios under all unstable length-scale categories when compared to the results under neutral and stable atmospheric conditions. The accuracy of the model predictions under unstable conditions is also generally the highest when compared to the other two atmospheric stability classes with regards to the u-variance.

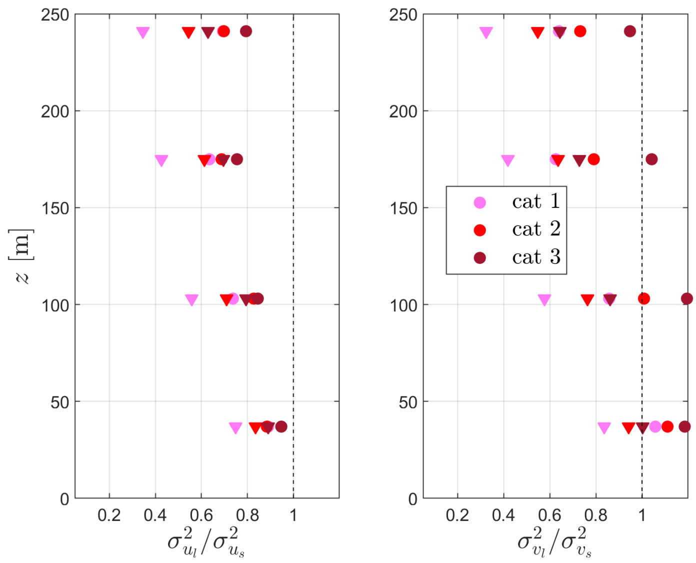

Figure 20: Similar to Fig. 18 but for unstable atmospheric conditions

Model predictions and measurements show considerable differences for the lidar-to-sonic v-variance ratios. At 37 and 103 m, the lidar often measures larger v-variances than the sonic anemometer, up to 20% for the larger length-scale range.

7.2 The physics-based neural network

Using the same type of PBNN as that described in Sect. 6.2, we train four sets of NNs, each NN set corresponding to a combination of the three possible parameters, which we can measure or derive from the observations of unit 590, i.e., L, $\sigma^2_{u_l}$, and $\sigma^2_{v_l}$ (see Table 6). Here we use the full set of 1000-samples of physics-based lidar-turbulence model inputs and output. The testing is performed with the unit 590 dataset from which we also know the true value of the u-variance: the sonic-anemometer value at the given height. The predictions are therefore performed on the 3844 10-min measured samples. We perform the analysis per measurement height. Note that the physics-based lidar-turbulence model uses the parameter L: we cannot measure L with a lidar but we approximate it by $L_l = \sigma_{u_l}/dU/dz$. Also, we use the dU/dz values from the sonic measurements, since we do not have lidar measurements below 37 m. Since the NNs are trained with initial random weights, the predictions change each iteration. The training is therefore performed 100 times over the 1000-samples input dataset and predictions are performed on the 3844 10-min measurement dataset; the result is a set of 100 3844-samples of predictions of sonic-equivalent σ^2_u values.

Table 6: Description of the different training inputs for the PBNN model for the full validation

Input short name	Input variables
P1	L_l, $\sigma^2_{u_l}$
P2	L_l, $\sigma^2_{u_l}$, $\sigma^2_{v_l}$
P3	$\sigma^2_{u_l}$, $\sigma^2_{v_l}$
P4	$\sigma^2_{u_l}$
P5	persistence (not a NN)

7.2.1 37-m height

Results of the performance of each PBNN for the 37-m measurements are illustrated in Figure 21. P5, i.e., persistence, does not vary, since no model is involved. It does not have poor mape or rmse (we expected this from the 1:1 comparisons in Figure 10), but the slope indicates a lidar-to-sonic mean bias of 16% for the

u-variance. From the results, P4, which only uses the lidar-based u-variance input, shows the lowest mape, the lowest median rmse, and the closest slope to one of all NNs.

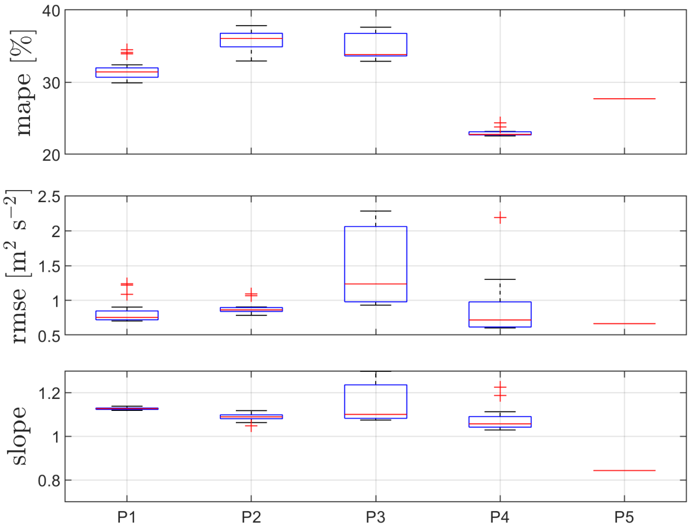

Figure 21: Performance of the PBNNs for different set of inputs P (see Table 6) at 37 m. Box plots of 100 trained datasets

We use the predictions from the P4-based PBNNs to construct a final prediction of lidar-based estimations of sonic-equivalent turbulence. For each of the 3844 10-min samples, we take the median of the 100 predictions of the P4-based NNs and compare the sonic-based TI values with those of the lidar and the lidar-based predictions. The results are shown in Figure. 22 for a number of mean wind speed ranges.

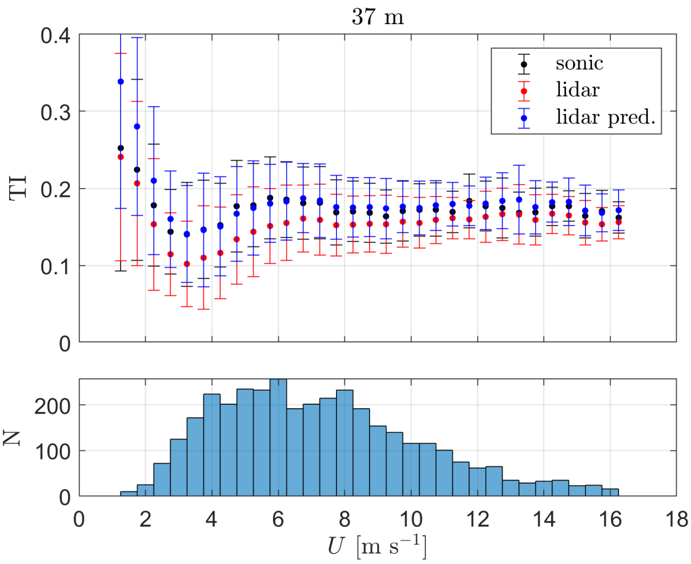

Figure 22: Turbulence intensity as a function of the mean wind speed from both sonic and lidar measurements, and lidar-based predictions (lidar pred.) at 37 m. The markers show the mean and the error bars ±one standard deviation within each mean wind speed bin. N represents number of 10-min on each mean wind speed bin

As illustrated, for most mean wind speed bins, the lidar predictions of TI are closer to the sonic-based TI values when compared to those from the original lidar measurements; this is always the case for the wind speed ranges with most measurements (3–10 m s^{-1}) and for most of the larger wind speed ranges (> 10 m s^{-1}). For more than half of the mean wind speed ranges, lidar-predicted TI values are slightly larger than the sonic-based TI values as the mean bias (slope) of the PBNNs is larger than one (see P4 in Figure 21).

7.2.2 103-m height

Similarly to the 37-m height case above, we show results for each of the PBNNs in Figure 23 for the 103-m measurements. The models' performance is similar to that of the 37-m case, but with deteriorated statistics (mape, rmse, and slope). P4 is also the PBNN that performs the best and it has a median slope of 1.02, which results in a lower mean bias than persistence (0.74).

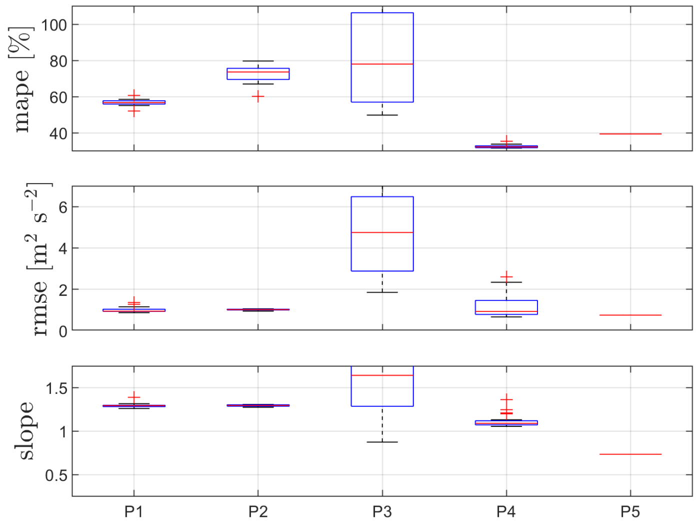

Figure 23: Similar to Fig. 21 but for 103 m

The final prediction is also constructed with the P4-based PBNN. As carried out for the 37-m height, we take the median of the 100 predictions of the P4 model and compare the sonic-based TI values with those of the lidar and the lidar-based predictions (Figure 24).

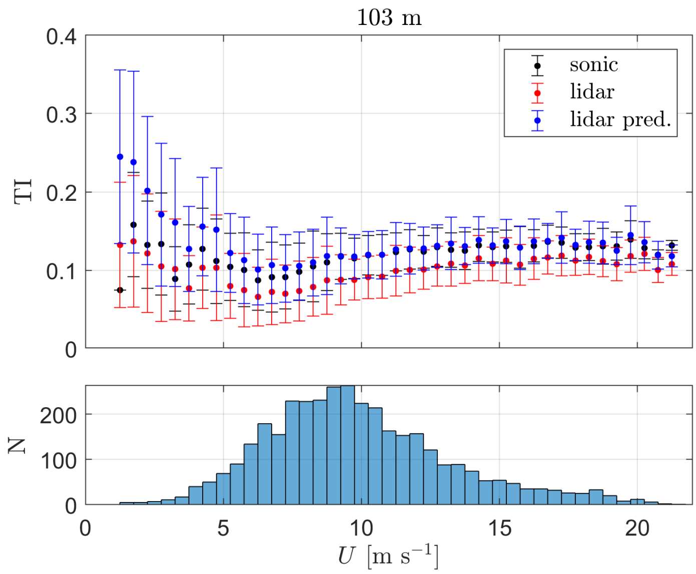

Figure 24: Similar to Fig. 22 but for 103 m

As illustrated, for most mean wind speed bins, the lidar-based predictions of TI are closer to the sonic-derived TI values when compared to the original lidar-based TI measurements. However, the lidar predictions match better the sonic values from the 5.25–5.50 m s $^{-1}$ bin onwards, where most occurrences lie, whereas the lidar TI predictions highly overestimate the sonic-based TI within the lower wind speed range.

7.2.3 175-m height

Similarly to the two previous cases, which are the two lowest measurement heights, we show results for each of the PBNNs in Figure 25 for the 175-m measurements. The performance of the models follows that of two lowest heights, but with further deteriorated statistics (mape, rmse, and slope). Once again, it is P4 the PBNN that performs the best and it has a median slope of 1.08, which results in a lower mean bias of this prediction compared to persistence (0.69).

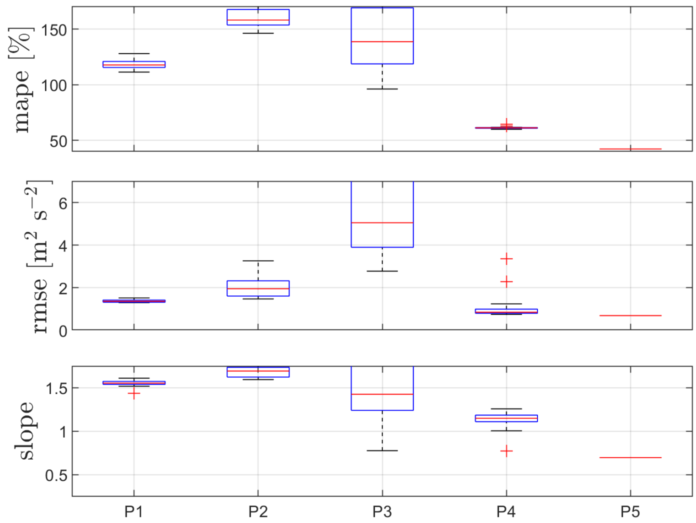

Figure 25: Similar to Fig. 21 but for 175 m

A final prediction is constructed using the median of the 100 predictions of the P4-based PBNN result, which we compare against both the sonic-based and lidar-based TI values (Figure 26).

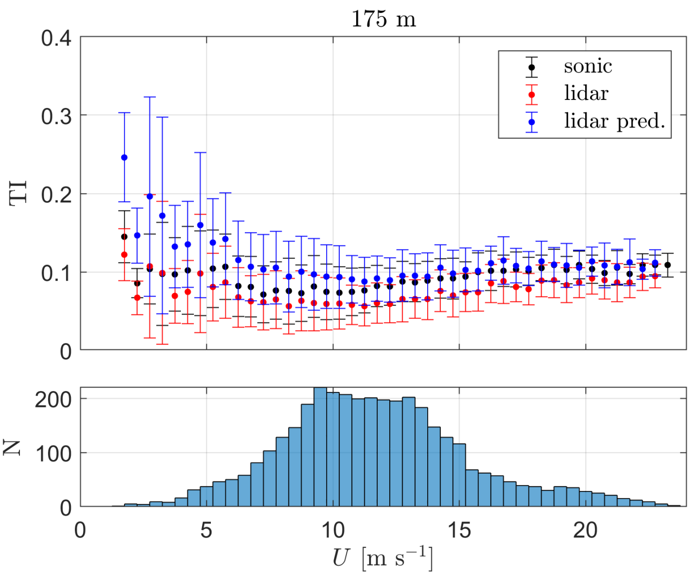

Figure 26: Similar to Fig. 22 but for 175 m

For nearly half of the mean wind speed bins (the highest half, i.e., $U > 9.25$ m s $^{-1}$), the lidar-based TI predictions are closer to the sonic-based TI values than the original lidar TI measurements. The lidar-based TI predictions highly overestimate the sonic TI within the lower wind speed ranges.

7.2.4 241-m height

Lastly, Figure 27 shows results for each of the PBNNs for the 241-m measurements. As expected, the models' performance follows that of the three lowest heights but with further deteriorated statistics. Once again, it is P4 the PBNN generally performing the best but with a large mape. P1's and P4's median rmses are close (about 2 m^2 s^{-2}) and P4 has a median slope of 1.00, which results in a lower mean bias of this prediction compared to persistence (0.74).

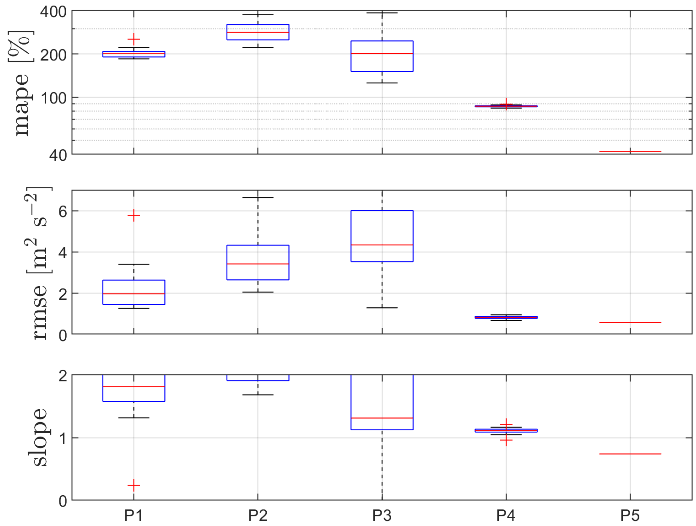

Figure 27: Similar to Fig. 21 but for 241 m

The final prediction with the P4-based PBNN results is similarly constructed as with the previous three heights and compared to the sonic-based and lidar-based TI values (Figure 28).

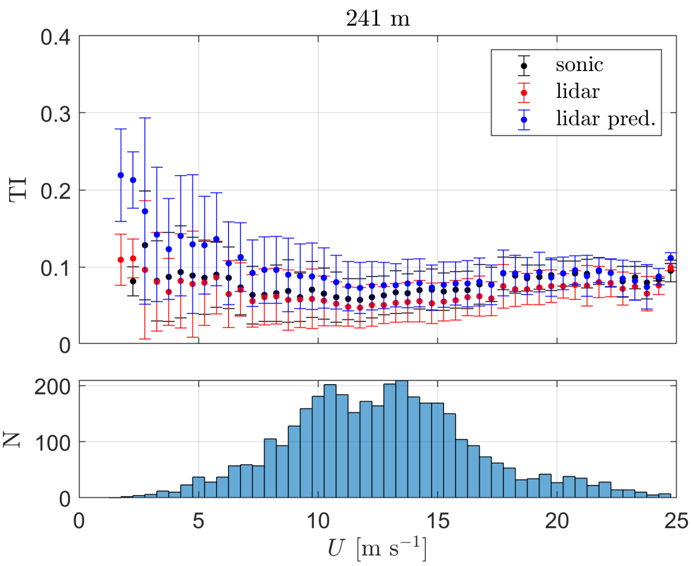

Figure 28: Similar to Fig. 22 but for 241 m

At this height, it becomes clearer that the ability of the lidar-based TI predictions to match the sonic-based TI measurements is wind speed dependent. The neural network overpredicts significantly the TI for wind speeds up to around 10 m s^{-1} and it is only the TI values of the highest wind speed ranges that are better predicted by the PBNN.

7.2.5 Summary

Overall, it seems that P4, namely $\sigma^2_{u_l}$, performs the best for all heights, as it has the lowest mape, the lowest rmse, and slope closer to one compared to the rest NNs. However, the performance above 103m deteriorates in height and in wind speed.

7.3 Data-driven neural network model

We now use the full set of 3844 10-min samples corresponding to the measurements from unit 590 to construct four height-dependent NNs and evaluate them with the measurements from unit 542 and the sonic anemometers. The training is performed 100 times over the 3844 10-min dataset from unit 590 and predictions are performed on the 3519 10-min dataset from unit 542; the result is a set of 100 3519 samples of predictions of sonic-equivalent σ^2_u values.

7.3.1 37-m height

Results of the performance of the DDNNs for the 37-m height using the unit 542 measurements are illustrated in Figure 29. In this case, P7 (U, $\sigma^2_{u_l}$) shows the best performance (lowest median mape and rmse, and closer to one slope). Similar but deteriorated performance is achieved by P5 ($\sigma^2_{u_l}$), which might be related to better predictions when including the mean wind speed due to the turbulence dependency on wind speed. P4, which only adds the v-variance to the P5 input, shows the highest mape. Using the lidar u-variance as proxy of the sonic-equivalent value (P8) results in nearly 10% lower u-variance (mean bias) compared to the sonic-based value.

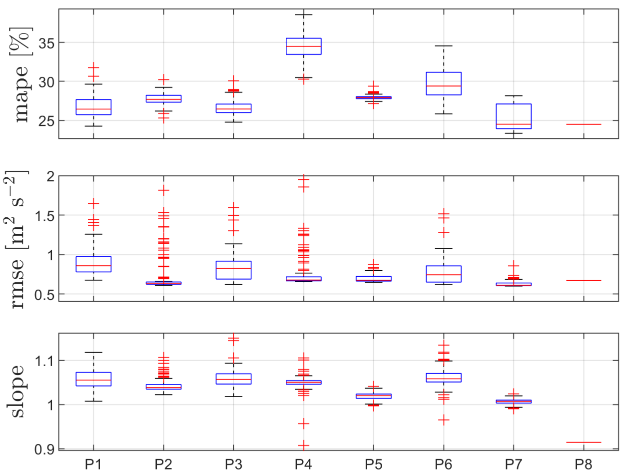

Figure 29: Performance of the DDNNs for different set of inputs P (see Table 5) at 37 m. Box plots of 100 trained datasets

For the DDNNs, we use the predictions from the P7 model to construct a final prediction of lidar-based estimations of sonic-equivalent turbulence. For each of the 3519 10-min samples, we take the median of the 100 predictions of the P7-based NNs and compare the sonic-based TI values with those of the unit 542 and unit 542-based predictions. The results are shown in Figure 30 for a number of mean wind speed ranges.

As illustrated, for nearly all mean wind speed bins, the unit 542-based TI predictions are in better agreement with the sonic-based TI measurements when compared to the original unit 542-based TI measurements.

For most wind speeds, the lidar-predictions of TI are higher than the sonic-based TI measurements, as expected since the P7-model, as well as all other NN-based predictions, show a mean bias higher than unity (see Figure 29).

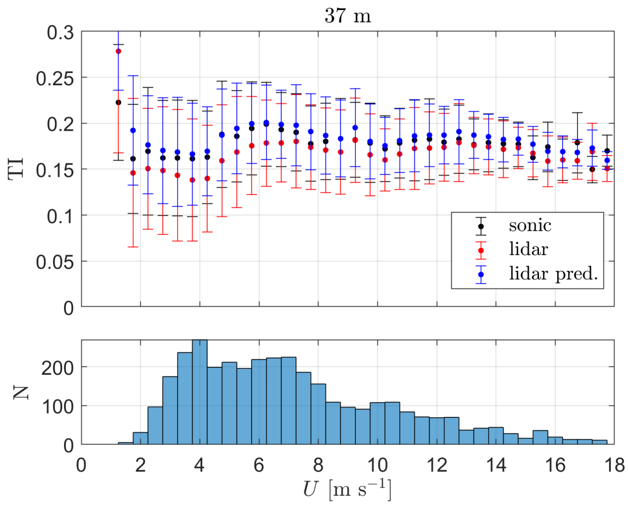

Figure 30: Turbulence intensity as a function of the mean wind speed from both sonic-anemometer (sonic) and unit 542 (lidar) measurements, as well as that from unit 542 predictions (lidar pred.) based on NNs constructed with unit 590 measurements at 37 m. The markers show the mean and the error bars ±one standard deviation within each mean wind speed bin. N represents number of 10-min on each mean wind speed bin

7.3.2 103-m height

Results of the performance of the DDNNs for the 103-m height using the unit 542 measurements are illustrated in Figure 31. Similarly to the 37-m case, P7 shows the best performance; generally, when compared to the first matching height, mapes and mean biases (slopes) are deteriorated for the NN models but P7's rmse does not seem increase. Using the unit 542 u-variance as proxy of the sonic-equivalent value (P8) results in nearly 15% lower u-variance (mean bias) compared to the sonic-based value.

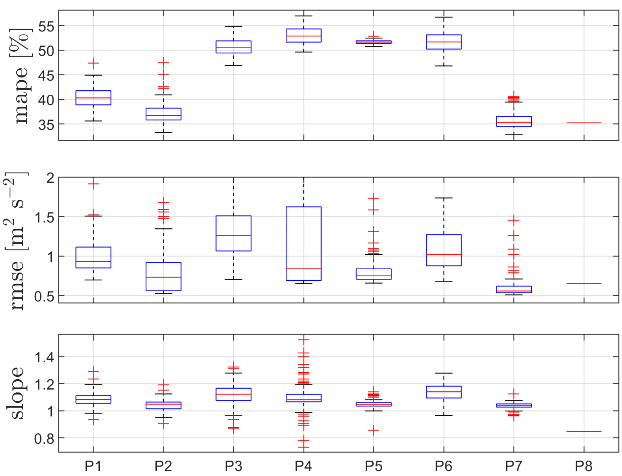

Figure 31: Similar to Fig. 29 but for the 103-m height

The final TI predictions for this height are also constructed with the P7-based DDNN, in a similar fashion as performed for the 37-m height and the results shown in Figure 32. As illustrated, for about half of the mean wind speed bins, the unit 542-based predictions of TI are closer to the sonic-derived TI values when compared to the original lidar-2 based TI measurements. However, the predictions do not better perform under a particular wind speed range and, under all wind speed ranges, they predict larger TI values than the original unit 542 measurements, systematically showing larger TI predictions compared to the sonic-based TI values within nearly all mean wind speed ranges.

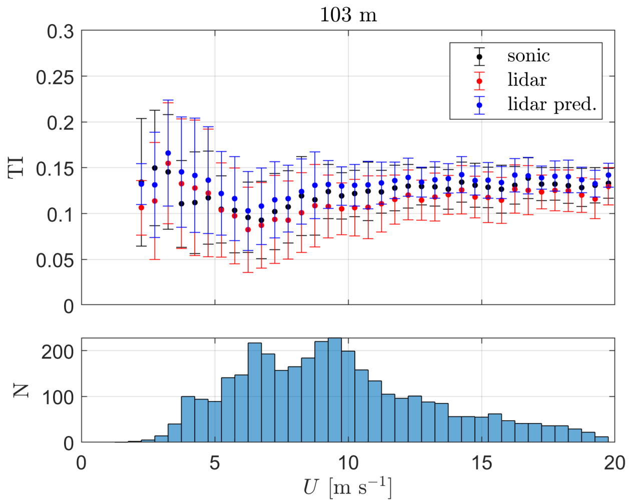

Figure 32: Similar to Fig. 30 but for the 103-m height

7.3.3 175-m and 241-m height

We combine the results for the two highest matching vertical levels, as they show similar behavior compared to that of the lowest two matching vertical levels but with deteriorated statistics, particularly for the mape values; however, all models' rmse values are rather similar between these two heights (see Figure 33). For both heights, using the unit 542 measurements as a proxy for the sonic-equivalent u-variance results in mean biases of about 16% and 12%, whereas the median bias of the P7 model is about 8% and 4% for the 175 and 241-m heights, respectively.

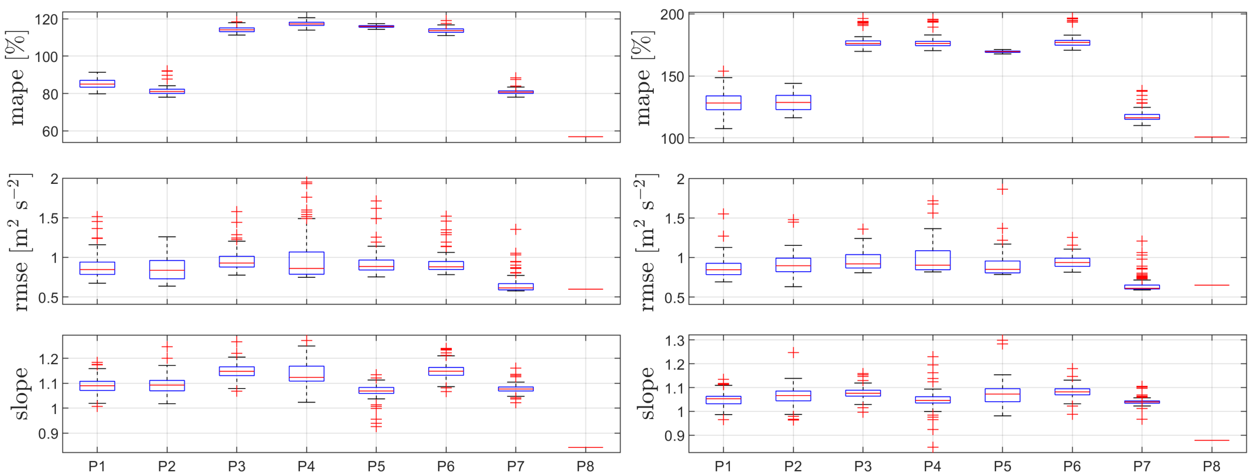

Figure 33: Similar to Fig. 29 but for the 175-m height (left) and 241-m height (right)

The final TI predictions for these two heights are constructed with the P7-based DDNN and results are shown in Figure 34. Unit 542-based TI predictions are clearly outperformed by the unit's own original TI measurements; only under a few mean wind speed bins, unit 542-based TI predictions are closer to the sonic-based TI values. Interestingly, when comparing the TI levels between these two heights within the low wind speed range ($U \lessapprox 5$ m s^{-1}), the lidar-based TI levels are close or higher at 241 m. This might indicate that the 'recovery' of turbulence observed for unit 590 is also observed in the measurements from unit 542.

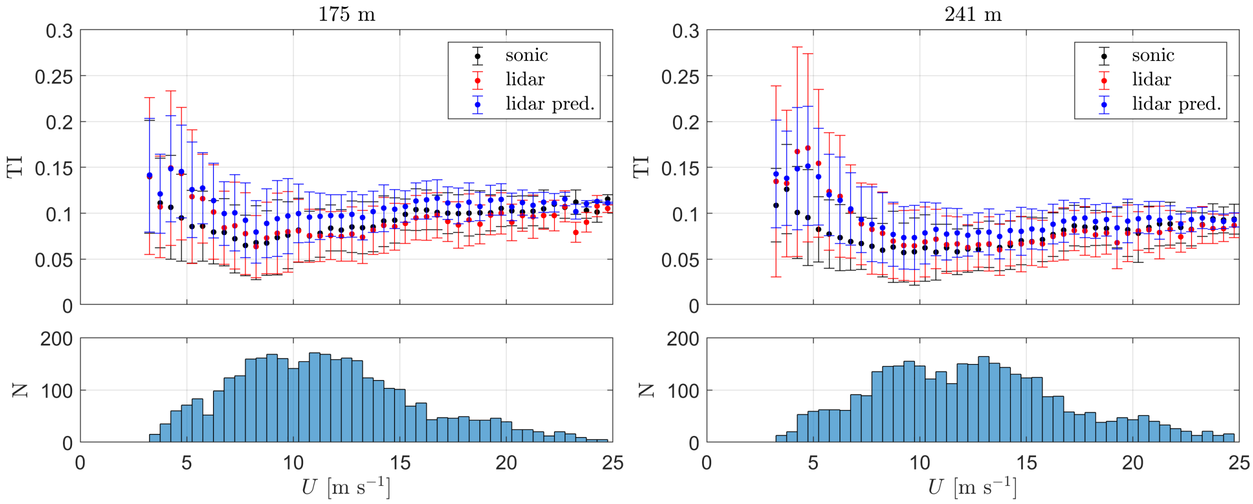

Figure 34: Similar to Fig. 30 but for the 175-m height (left) and 241-m height (right)

7.3.4 Summary

Overall, it seems that P7, namely U_l, σ^2_{ul}, performs the best for the heights 37m and 103m, since it has the lowest mape, the lowest rmse, and slope closer to one compared to the rest NNs. At 175m and 241m, the lidar's own TI measurements are very close to the sonic TI measurements or higher, even though it is expected that the lidar would underestimate TI measurements. Therefore, the lidar predicted TI does not improve the original lidar measurements (does not bring them closer to the TI measurements of the sonics).

8 Discussion

Before discussing the results of the physics-based lidar-turbulence model, PBNNs, and DDNNs evaluated in Sect. 7, it is important to note that directly comparing the PBNNs and DDNNs here proposed is not completely fair. DDNNs are here built using information from unit 590 only and predictions are carried out with a fully independent new lidar dataset (in our case from unit 542). The PBNNs use information from unit 590 to generate the physics-based datasets that serve as input for the predictions, which ultimately are based on the full unit 590 measurements. We can though say that the methodology here proposed, which involves the PBNNs, is a more practical approach than that involving the DDNNs, since we can generate new physics-based datasets to perform predictions at any site and height using the site-specific lidar measurements. DDNNs are on the other hand very local as they are site- and height-dependent; however, note that we could eventually use more than one lidar or more heights to construct the NNs so eventually a lidar measurement network would benefit the DDNNs.

8.1 On the physics-based lidar-turbulence model

The deteriorating behavior of the physics-based lidar-turbulence model at the highest two matching heights can have different explanations. First, the results of the physics-based lidar-turbulence model depend on the goodness and characteristics of the spectral velocity tensor model used. Here, we use the Mann model, which was originally formulated to simulate the spatial structure of stationary homogeneous turbulence under near-neutral atmospheric stability conditions and within vertical levels well inside the surface layer. Here, we are both looking at atmospheric conditions beyond near-neutral stability and, in some or most cases, well above the surface layer. Further, under all atmospheric stability conditions, we are using the proxy for the Mann turbulence length scale by Kelly [2018], which is highly sensitive to estimations of the local vertical wind shear. We also use, for simplicity and to isolate the effect of the turbulent length scale, the same values of the anisotropy parameter ($\Gamma = 3.0$) at all heights, although Γ exhibits height-dependency [Peña, 2019]. Although not shown, the medians of the ratio of the v-to-u variance based on the sonic-anemometer observations tend to be closer to unity the more unstable the atmospheric conditions are and the higher above the ground we observe. This behavior indicates that Γ is probably lower when approaching the above mentioned conditions as turbulence becomes more isotropic. As shown in Peña et al. [2010], the three Mann model parameters can be determined under a broad range of turbulence conditions and heights by fitting measured sonic-anemometer velocity spectra to pre-computed Mann-model spectra; here, we do not attempt this procedure, as we want to use close-to-lidar standard outputs to deduce the Mann parameters.

Note that the physics-based lidar-turbulence model assumes that the flow is homogeneous within the scanning volume, a condition that might not be fully valid, particularly when observing winds at the higher vertical levels at Østerild, since the scanning area of the lidar is enlarged with increasing focus distances. Also, since the lidar-to-sonic turbulence correlations deteriorate with height (see Sect. 5), the lidar Doppler radial velocity spectrum might be more sensitive to noise impacting mostly the lower end of radial velocity bins. Such impact artificially increases the radial and reconstructed velocity variances, which could explain the recovery of turbulence at the highest measurement levels. Finally, 2D turbulence might be present, particularly when measuring above 100 m; the Mann model, which is a 3D turbulence model, could be missing some part of the velocity variability at low-frequencies [Cheynet et al., 2018].

8.2 On the physics-based neural networks

The abilities of the PBNNs depend on both the accuracy of the physics-based lidar-turbulence model and on the accuracy/quality of the lidar turbulence measurements. Here we quality-filter lidar measurements by using the 10-min samples in which the difference in mean wind speed between unit 590 and the cup anemometers is below a threshold. The ratios of lidar-to-sonic u- and v-variance based on measurements show a recovery, particularly above 103 m, which the model cannot predict. The recovery can also be seen clearly in the TI predictions. Within the low mean wind speed range the original lidar TI is consistently lower than the sonic TI at 37 m; the difference in TI between these two estimates reduces with increasing height. The consequence is that the lidar-based prediction overcorrects the sonic-based TI value because the physics-based lidar-turbulence model on the contrary to the measurements cannot predict turbulence recovery at these heights.

8.3 On the data-driven neural networks

The abilities of the DDNNs depend on both the accuracy/quality of the lidar turbulence measurements used to construct the NNs (in this case unit 590) and that of the lidar turbulence measurements used for predictions (in this case unit 542). Note that the differences between the unit 542 and sonic-based TI values (Figs. 30–34) are smaller than those between the unit 590 and sonic-based TI values (Figs. 22–28). These TI differences also decrease with height, which suggests that the turbulence recovery observed in unit 590 measurements might be even stronger for unit 542 measurements; apart from a couple of mean wind speed bins the DDNN-based lidar TI predictions are always higher than the original lidar TI measurements. For unit 542 measurements, we apply the same filtering techniques as for the unit 590 measurements; if we do suffer from noise contamination of the Doppler radial velocity spectrum, this seems to more strongly impact the radial velocity fluctuations of unit 542.

9 Conclusions

The physics-based lidar-turbulence model agrees well with the unit 590 and sonic-anemometer measurements under a number of atmospheric stability conditions and length-scale ranges within the two first lidar-sonic matching heights, although we do not try to adjust the spectral model parameters, which is the basis of the model. Depending on the length-scale range and atmospheric stability condition, the level of agreement (i.e., mean bias) deteriorates at the highest two matching heights (178m and 244m).

The physics-based neural networks, which are trained by combining the unit 590 measurements with the physics-based lidar-turbulence model and applied to correct the unit 590 measurements, accurately predict the TI levels measured by the sonic anemometers. Better results are achieved at the two first matching heights (37m and 103m) within the broad range of mean wind speeds - generally, the lidar underestimates the TI levels and the predictions from the NNs show a slight overestimation compared to the sonics. With increasing vertical level, the method overpredicts the TI for wind speeds up to 11 m/s. These overpredictions are a result of the behavior of the lidar 590 as the measurement height increases particularly above 105 m: within several 10-min periods, the lidar captures high fluctuations in reconstructed horizontal velocity, which are not observed by the matching cup anemometer. It was not possible to correlate these 10-min periods with any of the internal lidar parameters. Therefore, the reliability and accuracy of the results from the PBNNs highly depends on the quality of the input lidar data. Potentially, the results from the PBNNs could be improved if the Doppler radial velocity spectra were available. By analyzing the Doppler spectra of radial velocities, we would be able to identify if background, optical or other types of noise are present; this would allow us to either further filter noise-contaminated data and/or correct the Doppler spectrum to improve the radial velocity estimate.

The data-driven neural networks, trained with unit 590 measurements and applied to correct the unit 542 measurements, accurately predict the TI levels measured by the sonic anemometers at the two first matching heights within most mean wind speed bins. At the two highest matching heights, the uncorrected unit 542 TI measurements are generally closer to the sonic-based values than the unit 542-based predictions, which overestimate the standard turbulence measures.

Although the results at the two highest matching heights are not encouraging for the DDNNs in particular, both presented methodologies appear as possible countermeasures of the lidar-turbulence paradox and can be further 1) improved, e.g., by better matching the local turbulence conditions at which predictions are performed, enlarging the datasets used for training, or improving the neural network architecture itself, and 2) extended, e.g., by using a lidar network for training the neural networks.

Acknowledgements

We would like to thank Kim Jakobsen, Guillaume Mougin and Gry Schachtschabel for the collaboration and discussions throughout this project. We would also like to thank Ørsted for supporting this work by granting us permission to use their lidar data from Østerild collected during 2018 and 2019.

References

E. Cheynet, J. B. Jakobsen, and J. Reuder. Velocity spectra and coherence estimates in the marine atmospheric boundary layer. *Boundary-Layer Meteorol.*, 169:429–460, 2018.

W. Fu, A. Sebastiani, A. Peña, and J. Mann. Dependence of turbulence estimations on nacelle lidar scanning strategies. *Wind Energ. Sci.*, 8:677–690, 2023.

M. Kelly. From standard measurements to spectral characterization: turbulence length scale and distribution. *Wind Energ. Sci.*, 3:533–543, 2018.

J. Mann. The spatial structure of neutral atmospheric surface-layer turbulence. *J. Fluid Mech.*, 273:141–168, 1994.

J. Mann, J.-P. Cariou, M. S. Courtney, R. Parmentier, T. Mikkelsen, R. Wagner, P. Lindelöw, M. Sjöholm, and K. Enevoldsen. Comparison of 3D turbulence measurements using three staring wind lidars and a sonic anemometer. *Meteorol. Z.*, 18:135–140, 2009.

J. Mann, A. Peña, F. Bingöl, R. Wagner, and M. S. Courtney. Lidar scanning of momentum flux in and above the surface layer. *J. Atmos. Ocean. Technol.*, 27:959–976, 2010.

Metek GmbH. Flow distortion correction for 3-d flows as measured by METEK's ultrasonic anemometer USA-1, 2004.

A. Peña. Østerild: a natural laboratory for atmospheric turbulence. *J. Renew. Sustain. Energ.*, 11:063302, 2019.

A. Peña, S-E. Gryning, and J. Mann. On the length-scale of the wind profile. *Q. J. Royal Meteorol. Soc.*, 136:2119–2131, 2010.

A. Sathe, J. Mann, J. Gottschall, and M. S. Courtney. Can wind lidars measure turbulence? *J. Atmos. Ocean. Technol.*, 28:853–868, 2011.

A. Sathe, J. Mann, N. Vasiljevic, and G. Lea. A six-beam method to measure turbulence statistics using ground-based wind lidars. *Atmos. Meas. Tech. Discuss.*, 7:10327–10359, 2014.

C. M. Sonnenschein and F. A. Horrigan. Signal-to-noise relationships for coaxial systems that heterodyne backscatter from the atmosphere. *Appl. Opt.*, 10:1600–1604, 1971.

The MathWorks Inc. Deep learning toolbox: 14.5 (r2022b), 2022. URL https://www.mathworks.com.

DTU Wind Energy is a department of the Technical University of Denmark with a unique integration of research, education, innovation and public/private sector consulting in the field of wind energy. Our activities develop new opportunities and technology for the global and Danish exploitation of wind energy. Research focuses on key technical-scientific fields, which are central for the development, innovation and use of wind energy and provides the basis for advanced education at the university.

We have more than 250 staff members of which approximately 45 are PhD students. Research is conducted within three divisions: Wind Energy Materials & Components, Wind Turbine Design and Wind Energy Systems.

Technical University of Denmark
Department of Wind Energy
Frederiksborgvej 399
Building 118
4000 Roskilde
Denmark
Telephone 46 77 50 85

info@vindenergi.dtu.dk
www.vindenergi.dtu.dk