Royal Netherlands
Meteorological Institute
Ministry of Infrastructure
and Water Management

Uncertainty analysis of climatological parameters of the Dutch Offshore Wind Atlas (DOWA)

C. de Valk and I.L. Wijnant

De Bilt, 2019 | Technical report; TR-379

Preface

This report presents the KNMI-results of work package 6.1 of the DOWA project (https://www.dutchoffshorewindatlas.nl/about-the-atlas) carried out by ECN part of TNO, KNMI and Whiffle, and funded by the Rijksdienst voor Ondernemend Nederland (RVO).

Abstract

We use statistical methods to assess the uncertainty of the DOWA climatological parameters, and compare it to the uncertainty in the KNW atlas. In particular, the following aspects are studied:

- **The precision of the DOWA and KNW wind data**. Precision of a data source refers to the random (non-systematic, non-predictable) error. Bias can be corrected for, but a lack of precision introduces an uncertainty that cannot be avoided. We use triple collocation and quadruple collocation analysis to assess the precision.

- **The relative bias** (compared to Meetmast IJmuiden mast and lidar measurements) of the DOWA and KNW atlas data. The bias in wind speed is assessed by means of quantile-quantile (QQ) plots that provide the relationship between the sorted values from one data source and the sorted values from another data source. Also, the bias in wind shear is assessed using QQ-plots.

- **The uncertainty in wind resource assessments based on 10 years of wind data**. To assess this uncertainty, we use a bootstrap method, both with and without accounting for multi-year dependence.

Acknowledgement

Ad Stoffelen, Andrew Stepek and Sybren de Haan of KNMI, Jan-Willem Wagenaar and Bernard Bulder of ECN part of TNO and Remco Verzijlbergh of Whiffle provided valuable input and feedback to the uncertainty analysis.

Content

1. Sources of data and information

1.1 Wind atlases

The Dutch Offshore Wind Atlas (DOWA) is a wind reanalysis dataset for the North Sea covering the 10-year period 2008-2017. The regional numerical weather model HARMONIE-AROME was used to downscale the ERA5 global weather re-analysis to hourly weather data on a 2.5 by 2.5 km grid spacing for heights up to 600 m. More information on the models and methods employed can be found under "Innovations in the DOWA project" on the DOWA-website[1].

The KNMI North Sea Wind (KNW) atlas is based on the ERA-Interim reanalysis (the predecessor of ERA5), and an earlier version of HARMONIE. More information on the KNW-atlas can be found on the KNW-website[2].

KNMI North Sea Wind (KNW) atlas	**Dutch Offshore Wind Atlas (DOWA)**
• 1979-2019 (40+ years)	• 2008-2018 (11 years)
• Captures the variability of the North Sea wind climate (40+ years long enough)	• Does not capture the variability of the North Sea wind climate (11 years not long enough)
• Based on re-analysis ERA-Interim and mesoscale weather model HARMONIE Version 37h1.1 (1979-2013) and Version 37h1.2.bugfix (2013-2019), the latter tested and adapted to guarantee a homogeneous dataset (similar results Version 37h1.1 and 37h1.2.bugfix).	• Based on re-analysis ERA5 (follow-up of ERA-Interim with higher spatial and temporal resolution) and mesoscale weather model HARMONIE Version 40h1.2.tg2 (improved wind information because turbulence is better modelled)
• HARMONIE used as downscaling-tool only (data-assimilation of measurements in ERA-Interim only)	• Additional measurements assimilated in HARMONIE (ASCAT-satellite surface wind measurements and MODE-S-EHS aircraft wind profile measurements)
• Climatological information up to and including a height of to 200 m	• Climatological information up to and including a height of 600 m
• Lacks the information required for further LES-downscaling	• Information required for further LES-downscaling included
• Cold starts: limited quality of hourly correlation with measurements (e.g. diurnal cycle)	• No cold starts: better hourly correlation with measurements and representation of the diurnal cycle
• Uniform wind shear correction applied	• No wind shear correction required

[1] https://www.dutchoffshorewindatlas.nl/
[2] http://projects.knmi.nl/knw/index.html

The table above summarizes the differences between the DOWA and the KNW-atlas. Where DOWA covers 10 years, the KNW-atlas covers more than 40 years (1979-2019) and therefore gives a better representation of the inter-annual variability (IAV) of the wind. The DOWA, however, is based on better models: ERA5 is better than ERA-Interim and in the new version of the HARMONIE-model, a better turbulence-scheme and additional assimilation of aircraft and satellite measurements have improved the wind information significantly. Furthermore, the method to produce the DOWA is improved compared to how the KNW-atlas was produced: e.g. no more 6-hourly restarts from the ERA-Interim re-analysis but instead, 3-hourly assimilation of measurements, and ERA5 reanalysis data are only fed at the boundaries.

1.2 Wind measurements

Wind speed and direction of the DOWA re-analysis have been compared to measurements from floating lidars, meteorological masts and scatterometer. The results are described in the validation reports Duncan et al (2019a, 2019b) and Knoop et al (2019). The comparisons cover several sites located offshore and on-shore and focus on the mean and standard-deviation of instantaneous differences between measurements and corresponding values from DOWA. To assess the uncertainty in the DOWA climatology, some other types of statistics are useful, which are discussed in this report.

The assessment of the uncertainty in the DOWA climatology is primarily based on the measurements at Meetmast IJmuiden (MMIJ) 75 km west of IJmuiden. There was no wind farm in the neighbourhood of the Meetmast IJmuiden that could disturb the measurements. Measurement data at the MMIJ-site include:

- Mast-mounted anemometer and wind vane measurements covering the period 2/11/2011 – 11/3/2016
- platform-mounted lidar measurements covering 1/11/2011- 9/3/2016.

The availability of simultaneous mast, lidar and DOWA data offers substantial benefits for quality assessment of the data from all three sources; in particular, it allows for triple collocation analysis (see Section 2.3).

At each of three levels on the MMIJ mast (27, 58 and 87 m), three booms were mounted, each supporting a cup anemometer and a wind vane. In addition, there was an anemometer at the top (92 m), and a LIDAR device was installed for measurements above 90 m. The fact that the wind is measured at three positions at each height makes it possible to select the sensor where the measurements are least disturbed by the mast, or to combine measurements in such a way that mast effects are minimalized. The wind speed that is derived from the raw measurements is often referred to as the "true wind speed" (or "derived wind speed"). The MMIJ measurements used in this study are the true wind speed and directions supplied by ECN part of TNO.

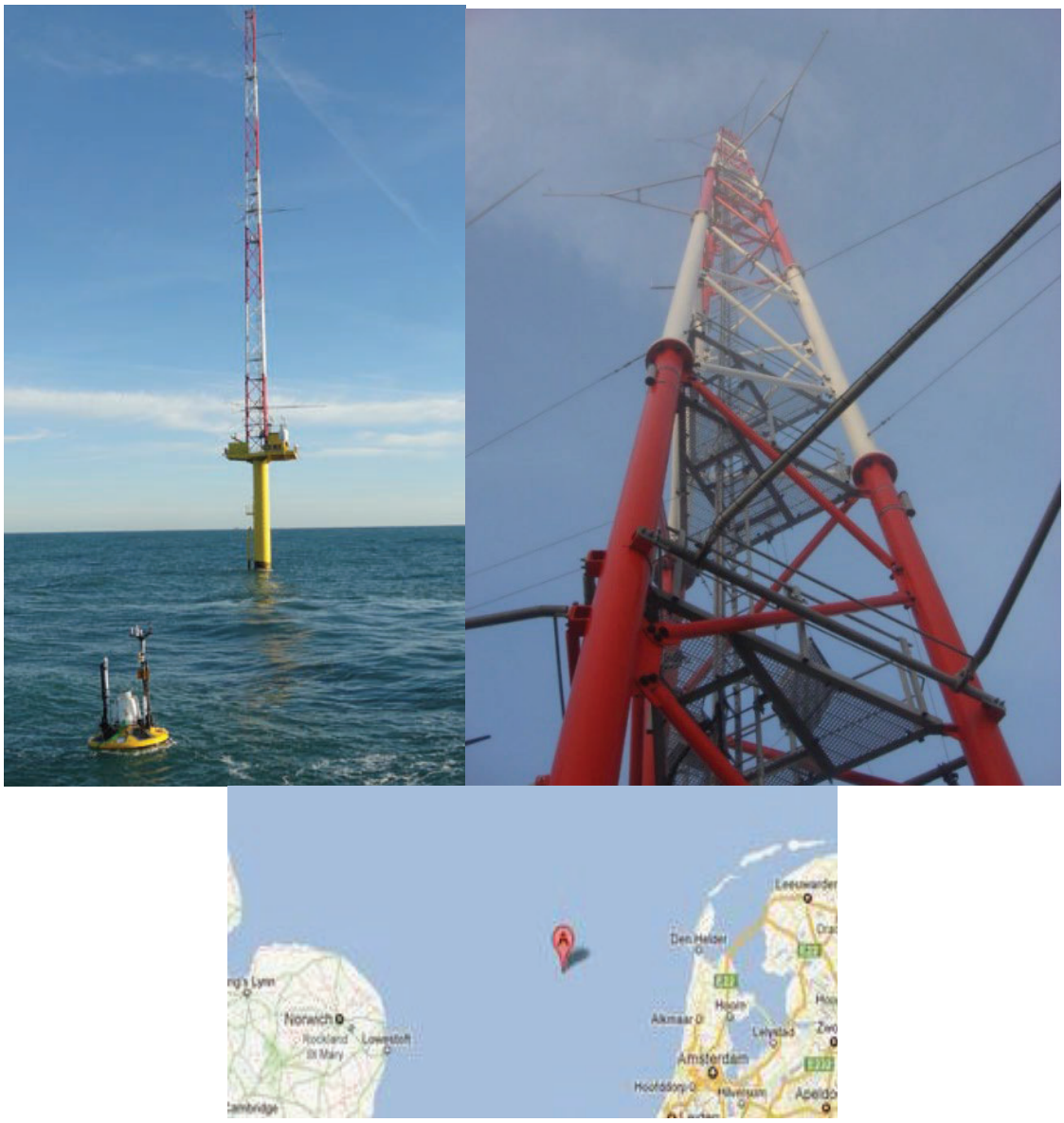

Figure 1: The top left panel shows Meteo Mast IJmuiden and the top right panel shows the mast construction and booms (source: BLIX Consultancy BV). The location of the mast is given in the panel below (source: ECN part of TNO).

In WP6 of the DOWA project, Windpark Egmond aan Zee (OWEZ) is considered as a pilot site for which a wind resource uncertainty assessment is provided for demonstration purposes. For this site, only one year of wind measurements undisturbed by turbine wakes is available: July 2005 - June 2006. After that, the wind from 135°-315° is disturbed and cannot be used for validation of the DOWA data. Therefore, we have focused on the data from Meetmast IJmuiden instead. However, the power curve for the 3MW V90 turbine deployed at OWEZ was used to assess the uncertainty in the reference yield in Chapter 3 of this report.

2. Bias and precision of DOWA wind speed and direction

2.1 Measurement errors and quality control

For the assessment of errors in model data it is important to realize that there are uncertainties in the measurements that are used to validate these model data. These measurement errors include e.g. instrument errors:

- instrument uncertainty assessed in laboratory calibration (e.g. for cup anemometers at Cabauw:1% or 0.1 m/s; see Knoop et al (2019) for further information),
- calibration errors,
- drift/malfunction (instrument errors may change with time!).

In the period 02-11-2011 to 09-03-2016, simultaneous mast and lidar measurements were performed at MMIJ. The lidar measurements cover 90 m and higher, and mast measurements were performed up to a height of 86 m, so the lowest lidar measurement height and the highest mast measurement height almost coincide. DOWA wind data for the nearest grid point were interpolated linearly in the vertical from the model output levels to the measurement heights.

To check the quality of these data, a principal component analysis (also known as empirical orthogonal function analysis; see Chapter 13 of von Storch and Zwiers (2003)) was performed on the matrix formed by the three data records from mast, lidar and DOWA. From the two minor principal components, two periods with anomalous measurements were detected: errors in mast measurement data (apparently a scale error) in the period 02-11-2011 to 25-12-2011, and errors in lidar measurements in the period 04-10-2013 to 11-12-2013. Data of these periods were excluded from further analysis.

2.2 Comparing measurements and model data

In order to be able to compare data sources, representation errors should be taken into account as well. Even an ideal instrument may provide measurements that cannot be used for validation unless the representation errors are known. Representation errors can be due to:

- Obstacles affecting the measurement (e.g. flow disturbance by the wind mast, e.g. JCGM, 2008).
- Location mismatch: even a small location mismatch may make a comparison between measurement and model invalid, especially in a non-uniform environment.
- Scale mismatch: a scale mismatch can show up as an added fluctuation (if sampled too frequently) or smoothing of the signal (if sampled not frequently enough).

In a non-uniform environment, scale and location mismatch are often hard to disentangle. Both instrumental data and model predictions are affected by scale and location mismatches.

Only differences between model and measurement that are larger than the measurement uncertainty can be attributed to model error.

Model data are generally grid box average values at one point in time (e.g. valid at the whole hour), whereas measurements are often recorded and presented as temporal averages (e.g. 10 min or hourly averages). It is not trivial how to properly compare the time-averages at one point in space with space-averages at one instant of time; how measurements are averaged in time affects their statistical relationship to model data (in particular the standard deviation of the differences).

In the following two sections, we will study bias and precision of the DOWA and KNW atlas data as well as of the measurements at MMIJ collocated with the DOWA/KNW data.

Bias is the systematic error in a data source. In a broad sense, it can refer to the deviation of the probability distribution of a prediction or measurement from the true probability distribution, possibly as a function of certain other variables. In a narrow sense, it may refer to an error in the mean value only. What is considered bias therefore depends on how detailed and complete errors in the data are described, i.e., on the "error model" considered (we will see examples in the following sections). It is not possible to assess the absolute bias of a data source by comparing it to another data source (measurements by different instruments and/or model predictions): it is only possible to say how bias in one data source compares to bias in the other data source (relative bias). The absolute bias of a data source needs to be determined in some other way (calibration checks, physical reasoning, etc.).

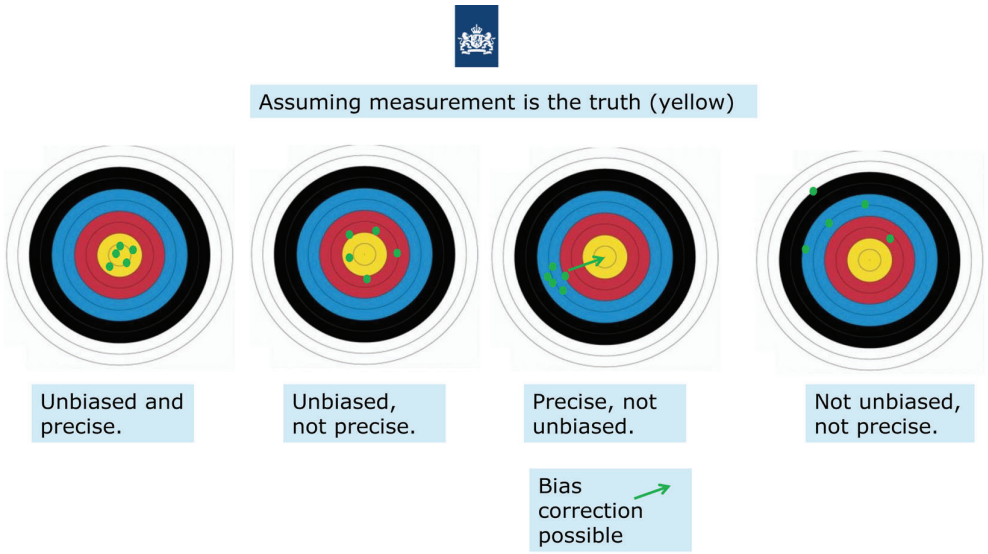

Figure 2: illustration of bias (in the mean) and precision.

Precision is the "smallness" of the non-systematic (non-predictable, random) error. The size of this error is usually given by a standard deviation (or its square, the variance); the smaller the estimated standard deviation of the random error, the more precise the data source.

A bias can be corrected for, but a lack of precision introduces an uncertainty that cannot be avoided. A simple picture of the distinction between precision and bias (in the mean value, in this case) is given in Figure 2.

Unlike the bias, the standard deviation can be determined by comparing datasets: with more than two data sources, it is possible to estimate the standard deviation of the random error of each source; see Section 2.3.

Normally, estimates of bias and precision cannot be easily traced back to particular error sources (representation error, calibration error, disturbances, etc.), but they may be compared to overall error estimates obtained by other means, providing a check on the latter. For cup anemometer and lidar measurements, Section 2.2.1 of Duncan et al (2019a) provides an overview of estimates of errors in measurements by mast-mounted cup anemometer and fixed and floating lidar.

2.3 Assessment of precision

Duncan et al (2019a) and Knoop et al (2019) focus on assessing the relative bias in the mean of the DOWA wind data and the standard deviations of the differences between DOWA and measurement data. In this section, we discuss estimates of the precisions of mast measurements, lidar measurements and DOWA and KNW re-analysis data of wind. These estimates provide information about data quality complementary to the information in these validation reports which compare data from different sources without explicitly assigning observed deviations to individual data sources. Furthermore, when using bias-corrected DOWA or KNW data for wind resource assessment (as in measure-correlate-predict procedures, see e.g. Carta et al (2013)), the accuracy of the bias correction may depend on the precisions of the DOWA or KNW data and of the measurements used for correcting the bias. For bias correction based on quantile-quantile plots, this is explained later in Section 2.4.

For the DOWA wind data at MMIJ, precision was assessed by a triple collocation analysis of wind speed data from mast, lidar and DOWA. The basic form of triple collocation of wind speed data x_1, x_2 and x_3 from three different sources is based on the following error model (e.g. McColl et al, 2014):

$$\begin{aligned} x_1 &= b_1 + a_1 t + \epsilon_1 \\ x_2 &= b_2 + a_2 t + \epsilon_2 \\ x_3 &= b_3 + a_3 t + \epsilon_3, \end{aligned} \qquad (1)$$

with t representing the true wind speed, and ϵ_1, ϵ_2 and ϵ_3 mutually uncorrelated random error terms, uncorrelated with t, and having mean zero and standard deviations σ_1, σ_2 and σ_3, respectively. The bias (systematic error) in x_i in (1) is equal to $b_i + (a_i - 1)t$ for i = 1, 2 or 3.

As we can only measure the bias relative to a particular data source, we may declare that one of the data sources is unbiased, by setting b_i=0 and a_i = 1 in (1) for that data source. From the model thus obtained, the following estimator of the standard deviations σ_1, σ_2 and σ_3 of the random error (i.e. the precision) can be derived (McColl et al, 2014):

$$\begin{aligned} \sigma_1 &= \sqrt{(C_{11} - C_{12}C_{31}/C_{32})} \\ \sigma_2 &= \sqrt{(C_{22} - C_{12}C_{23}/C_{13})} \\ \sigma_3 &= \sqrt{(C_{33} - C_{13}C_{32}/C_{12})}, \end{aligned} \qquad (2)$$

in which $C_{ij} = E\,(x_i - Ex_i)\,(x_j - Ex_j)$ denotes the covariance of x_i and x_j (with E the expectation), which is readily estimated from the data.

Note that in (2), the estimates of the standard deviations depend only on the covariances of the data, so they do not depend on the choice of which data source is unbiased. From the covariances, using the model (1), we can also derive estimates of the bias parameters b_i and a_i which have not been fixed (these estimates represent a relative bias). These bias estimates are not reported here, as we will consider relative bias in more detail in the next section.

To summarize the basic triple collocation method described above:

- you take collocated data from three different sources,
- assume that the relative bias takes the form of a linear term plus a constant,
- assume that the random errors have zero mean values and are mutually uncorrelated, and
- use the covariance matrix of the observations to determine the standard deviations of the random errors (the precisions) of the data sources.

The triple collocation analysis was performed for the full range of wind speed values ("all" in the tables), but also separately for three intervals of the mean wind speed from all three sources: below 7.5 m/s, between 7.5 and 11.8 m/s and above 11.8 m/s. The intervals were chosen in such a way that all intervals contain a third of the data. The precisions (i.e., the estimates of the standard deviations of the random wind speed errors) are shown in Table 1, where the first number in each cell refers to the triple collocation analysis with mast, lidar and DOWA-data and the second number in the cell refers to the triple collocation analysis with mast, lidar and KNW-data. For example: the precision in the DOWA-dataset is 1.24 m/s and of the KNW-dataset 1.48 m/s (all measurements included).

Wind speed range	mast	lidar	DOWA/KNW
All	0.21/0.20	0.36/0.36	1.24/1.48
Below 7.5	0.16/?	0.43/0.48	1.22/1.44
7.5-11.8	0.10/?	0.36/0.38	1.24/1.40
Above 11.8	0.23/0.24	0.31/0.31	1.16/1.41

Table 1: Estimates of precision (standard deviations of the random wind speed errors in formula 1) at MMIJ (in m/s) for mast, lidar and DOWA (left number in each cell) and for mast, lidar and KNW data (right number in each cell). The standard deviations of the wind speed error have been calculated for all wind speeds (1st row) and for three wind speed classes (rows 2-4). "?" indicates that precision cannot be estimated because the covariance matrix is inconsistent with the model (1); this may happen if the data source is much more precise than the other sources.

Table 1 indicates that

1. DOWA wind speed are more precise than KNW wind speed
2. Wind speed measurements from mast-mounted anemometer and lidar are considerably more precise than DOWA or KNW wind speeds
3. Mast measurements are more precise than lidar measurements.

In fact, the random errors in the measurements are almost negligible in comparison to the random errors in DOWA and KNW data. This implies that the DOWA-validation against mast and lidar measurements (Duncan et al, 2019a) provides a reliable picture of the error statistics of DOWA data.

It should be stressed here that even though the DOWA data are less precise than the measurements, their precision may be quite acceptable for many purposes. For example, for wind resource assessment, Weibull fits of the direction-dependent probability distribution of wind speed are estimated. Such estimates are not likely to be affected much by the random error of the data source, as they represent averages over a considerable time interval (say, 10 years). The impact of the random error on such statistics can be assessed by simulation. However, in the analysis of the uncertainty of a wind resource assessment, often the random error in the data source does not need to be addressed explicitly, as is the case when using a bootstrap method as in Ch. 3.

In Table 1, the estimates of precision for the three wind speed classes are likely to be less reliable than the estimates for the full range of wind speed, since the absence of values above or below some threshold violates the error model (1). To some extent, even the estimates for the full wind speed range suffer from this problem, as wind speed cannot be negative. This problem does not occur if wind components are analysed as in Stoffelen (1998), or if the wind vector represented as a complex number is analyzed in this manner (see below).

The estimated standard deviations of random errors in the mast measurements in the low and medium wind speed could not be estimated from the combination

of mast, lidar and KNW data, as the covariance matrices are inconsistent with the model (1) in these cases. This appears to happen when a data source x_i (mast data in this case) is much more precise than the other data sources, making the expression of the form $C_{ii} - C_{ij}C_{ki}/C_{kj}$ in eq. (2) almost zero (so its sign can easily become negative). This means that we should be cautious in interpreting differences between estimates of standard deviations for the mast data for different wind speed ranges in Table 1.

For mast measurements of wind speed above 7.5 m/s, the standard deviations estimated from collocated mast, lidar and DOWA data in Table 1 are consistent with the (supposedly RMS) error cited in Section 2.2.1 of Duncan et al (2019a): 0.1 m/s and 1-2% for higher wind speeds. For wind speeds below 7.5 m/s, the value in Table 1 is higher. Furthermore, the standard deviations in Table 1 do not appear to increase with wind speed (indicated by the wind speed class). This is somewhat curious, as for mast and lidar measurements, instrument error is usually reported in terms of relative error; see e.g. Section 2.2.1 of Duncan et al (2019a). On the other hand, the representation error due to turbulence is quantified by an absolute error size and is normally larger than the instrument error (Stoffelen, personal communication). Also, we should be cautious in interpreting the wind-speed dependence of these numbers, in view of the issues mentioned in the previous paragraphs.

The same triple collocation analysis was performed with the wind vector (represented as a complex number) instead of wind speed. Using the model (1) in this context then presumes that errors in the wind vector are isotropic (errors in the wind components are uncorrelated with identical standard deviations). In this case, the reported standard deviation applies to the magnitude of the random error in the wind vector, which implicitly incorporates the error in wind direction. The precisions of the wind vector error are given in Table 2.

Wind speed range	mast	lidar	DOWA/KNW
All	0.21/0.18	0.89/0.89	1.96/2.16
Below 7.5	0.16/0.23	1.01/1.12	2.02/2.26
7.5-11.8	0.10/0.10	0.87/0.73	1.95/2.12
Above 11.8	0.23/0.22	0.78/0.79	1.86/2.02

Table 2: As Table 1, but for the precision of the wind vector (standard deviation of the magnitude of the random error in the wind vector in m/s).

Comparing these results with those for wind speed in Table 1 indicates that the error in lidar wind direction is rather large, even though the error in wind speed is small. Something similar is seen with DOWA and KNW data, but the relative increase in standard deviation is somewhat smaller in this case. For the mast measurements, the standard deviation is about the same for wind speed and wind vector, indicating that mast measurements of wind direction are precise.

Possibly more reliable estimates of the error in wind direction from the different data sources could be obtained by analyzing the components aligned with and perpendicular to the dominant wind direction separately, but this was not undertaken in the present study.

The average wind speed at 90 m at MMIJ equals 10.0 m/s. This number can be used to normalize the standard deviation of the random error to a coefficient of variation (CV). This results in a CV of 4% for lidar wind speed and a CV of 9% for lidar wind vector difference.

A potential limitation of the basic triple collocation model eq. (1) is that it cannot represent scale mismatch (see Section 2.2). One might suspect that model reanalysis data, lidar and mast measurements do not resolve wind variations at the same spatial and temporal scales. The collocated data at the MMIJ site do not provide information about spatial scale mismatch, but as we have time series, we may obtain information about a possible temporal scale mismatch. For this purpose, the approach of Stoffelen (1998) was adapted to time-series data. It is based on the idea that if a weather prediction model does not resolve the finer scales, then the spectral density of the model output will be lower than the spectral density of measurements at these scales.

But this is not observed at the time scales considered in this study; see Figure A1 in Appendix A. Fluctuations on all time scales from 2 hours to 3 years in the DOWA and KNW data have very similar magnitudes as fluctuations in the measurement data on the same scales. So there is no evidence of a temporal scale mismatch between DOWA, KNW and the measurements.

However, the coherence[3] between DOWA or KNW reanalysis and mast measurements in Figure A2 of Appendix A decays much more rapidly with increasing frequency than the coherence between mast and lidar measurements. In other words: the fluctuations in DOWA and KNW data are realistic, but different from those in the measurements (see also the explanation in appendix A). These fluctuations are well represented by a noise term as in eq. (1). However, the high-frequency fluctuations in the mast and lidar measurements (not present in the DOWA and KNW data) are strongly correlated with each other (see their coherence function in Appendix A), implying that the model in eq. (1) should be replaced by a model of the form

$$\begin{aligned} x_1 &= b_1 + a_1(t+\eta) + \epsilon_1 \\ x_2 &= b_2 + a_2(t+\eta) + \epsilon_2 \\ x_3 &= b_3 + a_3 t + \epsilon_3 \end{aligned} \qquad \text{(3a)}$$

with the labels 1, 2 and 3 corresponding to mast, lidar, and DOWA or KNW, respectively,

t representing the common wind component present in all data sources,
η a wind fluctuation not present in the re-analysis data, and the
ϵ_i are mutually uncorrelated zero-mean noise terms.

[3] Also known as "coherency squared"; it can be regarded as the square of a correlation coefficient between the components of the two signals lying within a narrow frequency band. Phase shifts do not reduce the coherence.

This triple collocation model is equivalent to the model of Stoffelen (1998), containing a term (η in eq. (3c)) representing a wind component not predicted by the models. However, η in eq. (3) is not due to scale mismatch, but to lack of coherence between model reanalysis and measurements.

To estimate the parameters of model (3a) (in particular, the standard deviations of ϵ_1, ϵ_2, ϵ_3 and η), it could be converted to a model similar to eq. (1) by defining $s = t + \eta$:

$$\begin{aligned} x_1 &= b_1 + a_1 s + \epsilon_1 \\ x_2 &= b_2 + a_2 s + \epsilon_2 \\ x_3 &= b_3 + a_3 s + (\epsilon_3 - a_3 \eta). \end{aligned} \qquad (3b)$$

This looks again like model (1), with ϵ_3 replaced by $\epsilon_3 - a_3\eta$ as the noise term for the reanalysis data. However, unlike model (1), this noise is now correlated with s in (3b). Therefore, (3b) cannot be estimated from the covariances of the data in the same way that model (1) was estimated. This shows that we cannot estimate the parameters of the model (3a) as we did with (1).

However, we can estimate a model of the form (3a) if we use both reanalysis datasets simultaneously. A straightforward extension of the model (3a) is the quadruple collocation model

$$\begin{aligned} x_1 &= b_1 + a_1(t + \eta) + \epsilon_1 \\ x_2 &= b_2 + a_2(t + \eta) + \epsilon_2 \\ x_3 &= b_3 + a_3 t + \epsilon_3 \\ x_4 &= b_4 + a_4 t + \epsilon_4, \end{aligned} \qquad (3c)$$

with the labels 1, 2, 3, and 4 corresponding to mast, lidar, DOWA and KNW, respectively. As before, the "noise terms" ϵ_1, ϵ_2, ϵ_3 and ϵ_4 are assumed to be mutually uncorrelated with mean zero. This model is appropriate in view of the coherence functions in Figure A2 in Appendix A: the measurement data are strongly correlated on short time scales less then say 10 hours (represented by the common fluctuation η), but not with the DOWA and the KNW data, and the DOWA and KNW data are not strongly correlated with each other either.

For measurement data ($i = 1,2$), the deviation from the local wind $t + \eta$ equals (see 3c) $b_i + (a_i - 1)(t + \eta) + \epsilon_i$, which has a bias component $b_i + (a_i - 1)(t + \eta)$ and a random component ϵ_i. Therefore, the mean square error MSE, defined as the mean square difference between a measurement and the local wind $t + \eta$, is the sum of the bias MSE $b_i^2 + (a_i - 1)^2(\sigma_t^2 + \sigma_\eta^2)$ and the noise MSE σ_i^2. For mast and lidar measurements, the precision is therefore given by the noise standard deviation σ_i.

For reanalysis data ($i = 3,4$), the deviation from the local wind $t + \eta$ equals (see 3c) $b_i + (a_i - 1)t - \eta + \epsilon_i$, which has a bias component $b_i + (a_i - 1)t$ and a random component $\epsilon_i - \eta$. In this case, the MSE is therefore the sum of the bias MSE $b_i^2 + (a_i - 1)^2\sigma_t^2$ and the random MSE $\sigma_\eta^2 + \sigma_i^2$. For DOWA and KNW data, the precision is therefore given by the non-bias root mean square error (RMSE) $(\sigma_\eta^2 + \sigma_i^2)^{1/2}$.

In the estimation of the parameters of the quadruple collocation model, there is some redundancy (more data covariance values than unknowns); this allows averaging of the estimates of certain parameters. The result of the quadruple collocation analyses is summarized in table 3 (wind speed data) and 4 (wind vector data).

For all data sources, the results of the quadruple collocation analysis agree with the results of the triple collocation analysis in Tables 1 and 2. This is not entirely trivial, because in the triple collocation analysis, we have estimated the model (3b) as if it were (1) (i.e., as if $\epsilon_3 - a_3\eta$ is not correlated with s). Apparently, this does not affect the outcomes much in this case, otherwise the estimates would not agree with the quadruple collocation analysis.

Wind speed range	mast	lidar	DOWA	KNW
All	0.21	0.36	1.23	1.48
0.0-7.5	0.04	0.46	1.26	1.51
7.5-11.8	0.10	0.32	1.27	1.53
11.8-	0.24	0.34	1.13	1.39

Table 3: Estimated precision (non-bias RMSE) in wind speed for mast, lidar, DOWA and KNW data from a quadruple collocation analysis. Rows: for all data, and for data in three classes of wind speed (mean from all three sources). Unit: m/s.

Wind speed range	mast	lidar	DOWA	KNW
All	0.19	0.89	1.95	2.16
0.0-7.5	0.23	0.73	1.91	2.16
7.5-11.8	0.07	0.66	2.03	2.23
11.8-	0.20	1.19	1.89	2.06

Table 4: As Table 3, but for wind vector errors.

2.4 Assessment of relative bias

Relative bias of the three data sources was assessed by means of quantile-quantile (QQ) plots, which give the relationship between the sorted values from one data source and the sorted values from another data source. An example is shown in Figure 1. Such a QQ plot provides much more insight than a simple mean difference: it provides a relationship between two variables, determined through their distribution functions. This relationship is a non-decreasing graph which may be nonlinear.

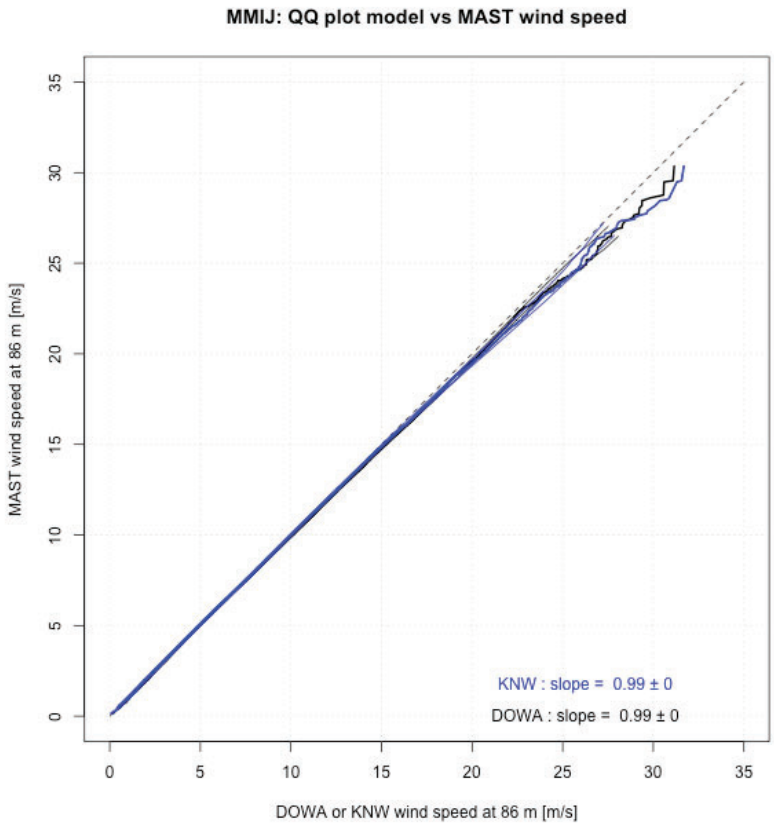

Figure 1: Quantile-quantile plots of wind speed data from two sources: mast vs. DOWA (black thick), and mast versus KNW (blue thick). Thin lines indicate the 90% compatibility interval. Text in plot gives slope estimates with their standard deviations.

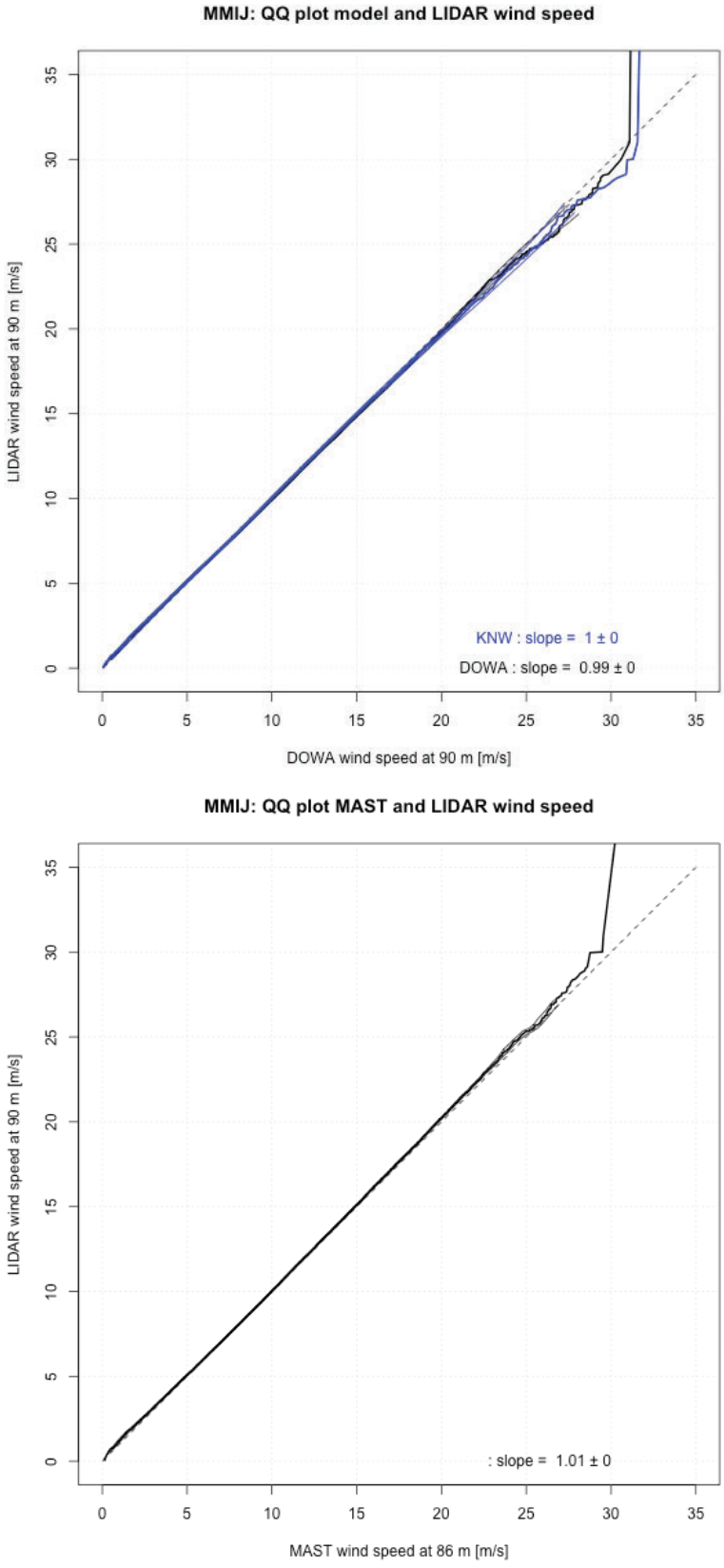

Figure 2: As Figure 1, for lidar vs. DOWA (black), and lidar versus KNW (blue) (top), and for lidar vs. mast (bottom).

In wind resource assessment, regression is often used to establish a relationship between two sets of measurements: usually a long time-series from a weather prediction model or from measurements at a location in the vicinity, and a shorter time-series of measurements at the site of a projected wind farm. This relationship is then used to correct bias in the long time-series, so it can be used to assess the wind resource at the wind farm site.

However, instead of regression, one might use the relationship from a QQ plot for bias correction. This may be a better choice for several reasons: it ensures that the probability distribution of the bias-corrected time-series is correct (which is all that is needed for the purpose), it avoids the estimation bias which tends to plague regression (e.g., different regressions are obtained if "x" and "y" are swapped), and the QQ plot is quite precise, as illustrated below. If appropriate, the raw QQ plot may be further simplified by fitting a line or other function, or by smoothing. In this report, we will only consider linear approximation of QQ plots.

Figure 1 shows the QQ plots for mast measurements versus DOWA reanalysis (black) and for mast measurements versus KNW reanalysis (blue) and Figure 2 shows the QQ plots for lidar measurements versus DOWA/KNW reanalysis and for lidar measurements versus mast measurements. The slopes of the QQ plots in Figures 1 and 2 (computed by total least squares) are all within 1% of the unit slope, which indicates that the distributions of the datasets compared are very similar. All QQ plots are very well approximated by lines through the origin. The anomalies in the tails of the QQ plot in Figure 2 appear to be due to an outlier in the lidar data. Figure 1 and 2 show that the DOWA slightly overestimates high wind speeds.

The QQ plots are very precise, as can be seen in the narrow 90% compatibility intervals[4] shown by thin lines in the plots; these lines are barely visible as they often coincide with the QQ plot. The intervals were estimated by generating 1000 synthetic datasets having similar characteristics as the original dataset, and computing the QQ plot for each of them. From these, we can derive compatibility intervals of the QQ plots, as well as estimates of the standard deviation of the slope, listed in Table 5. The synthetic data were generated by resampling of the time-series by a moving block bootstrap (see Figure 3): time-intervals of 62.5 days are drawn randomly with replacement, and new time-series are created by stringing the data within these blocks together; see e.g. Künsch (1989). Resampling here has the same function as Monte-Carlo simulation, but where the latter requires an explicit probabilistic model, resampling does not; it just needs the data. The resampling represents variability in wind speed values encountered due to using a finite sample of (correlated) data. Random, zero-mean errors in the data may affect the QQ plot in different ways: they may cause bias (which was verified to be small by Monte Carlo simulation), as well as increased variability. However, the increased variability is already represented by the bootstrap ensemble: in the absence of errors in the data, the QQ plots from all bootstrap samples would be on a single graph. Therefore, when using the bootstrap method to assess the uncertainty of

[4] "Compatibility interval" is just a better name for a confidence interval, as it indicates the range of values compatible with the data; see e.g. Amrhein et al (2019).

a QQ plot, there is no need to address the precision of the data explicitly. Seasonality was not preserved in the resampling; this may lead to some overestimation of variability, as the lengths of the seasons in the resampled data are more variable than in the original data. The bootstrap method only works if values in different blocks are weakly dependent, so the block length should be long enough. The chosen block length of 62.5 days is the time over which wind speed effectively de-correlates. A related technique is the jack-knife, based on estimates with part of the sample left out. Jack-knife and bootstrap techniques are recommended for uncertainty assessment in DNV (2011).

The high precision of QQ plots makes them suitable as an alternative to regression-based measure-correlate-predict (MCP) methods (Carta et al., 2013), which are prone to large uncertainty (both bias and variance) if the data are imprecise and/or correlation is low. Indeed, several MCP methods based on fitting relationships to sorted data from two different sources exist; see King and Hurley (2005) (somewhat inappropriately listed under the header "methods based on first-order linear regressions" in Carta et al., (2013))

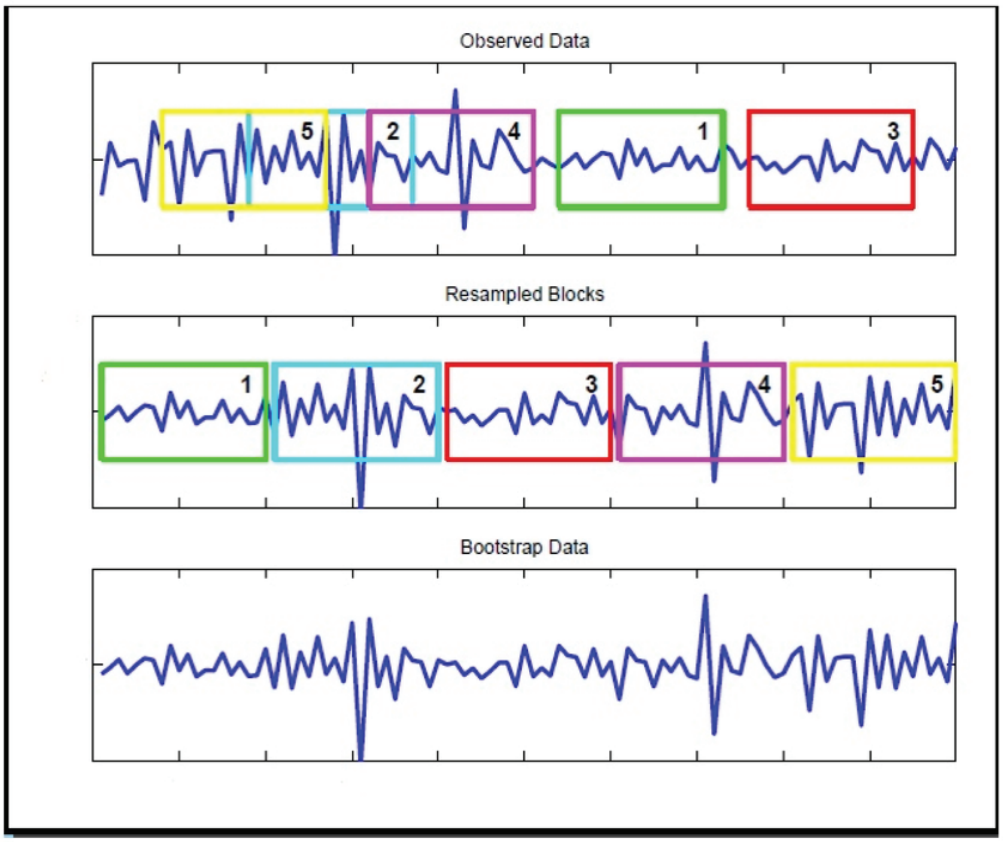

Figure 3: Illustration of the block bootstrap method applied in this study.

QQ plots for 30 degree wind direction bins in DOWA were also computed by (a) selecting all data with direction in a given bin, and (b) applying the block bootstrap to this data sequence, using a block length equal to 62.5 days times the fraction of data having a wind direction in that bin. The direction-dependent slope estimates and their standard deviations are given in Table 5.

Although based on smaller datasets, even the directional QQ plots appear to be very precise, with slope standard deviation within 1%. Bias of lidar wind speed relative to mast wind speed is very small (within 1%). Relative bias in KNW wind speed is also small (only for the bins of 0 and 210 degrees, it reaches 3%). Overall, the KNW data appear to be compatible with a unit slope. For the 90 and 120 degree bins, DOWA wind speed underestimates mast and lidar wind speed by 3-5%, where KNW data are closer to the measurements. For the 270 and 330 degree bins, DOWA is somewhat higher than the measurements, and KNW somewhat lower. Except for this, we can conclude that the bias in DOWA and KNW wind speeds is very similar.

x	DOWA	KNW	DOWA	KNW	Mast
y	Mast	Mast	Lidar	Lidar	Lidar
Omni	**0.99**±0.00	**0.99**±0.00	**0.99**±0.00	**1.00**±0.00	**1.01**±0.00
0	**0.99**±0.01	**1.03**±0.01	**0.98**±0.01	**1.02**±0.01	**0.99**±0.00
30	**1.00**±0.01	**1.01**±0.01	**0.99**±0.01	**1.01**±0.01	**0.99**±0.00
60	**1.01**±0.01	**1.00**±0.01	**1.02**±0.01	**1.01**±0.01	**1.02**±0.00
90	**1.04**±0.01	**0.99**±0.01	**1.05**±0.01	**1.01**±0.01	**1.01**±0.01
120	**1.03**±0.01	**0.99**±0.01	**1.03**±0.01	**0.99**±0.01	**0.99**±0.00
150	**1.02**±0.01	**1.01**±0.01	**1.01**±0.01	**1.00**±0.01	**0.99**±0.00
180	**0.98**±0.00	**0.99**±0.00	**0.99**±0.00	**1.00**±0.00	**1.01**±0.00
210	**0.97**±0.00	**0.97**±0.01	**0.98**±0.00	**0.98**±0.01	**1.01**±0.00
240	**0.98**±0.00	**0.98**±0.00	**0.99**±0.00	**0.98**±0.00	**1.01**±0.00
270	**0.99**±0.00	**1.01**±0.00	**0.99**±0.00	**1.01**±0.00	**1.00**±0.00
300	**0.98**±0.00	**1.01**±0.01	**0.99**±0.00	**1.02**±0.01	**1.00**±0.00
330	**0.97**±0.00	**1.02**±0.01	**0.98**±0.01	**1.02**±0.01	**1.01**±0.00

Table 5: Estimates of slopes (y/x) of QQ plots in bold red ± their standard deviations independent of direction ("omni") and for 30 degree bins of wind direction; the first column lists the centre of each bin.

To assess the bias in the wind shear, it is not enough to rely on bias in mean wind speeds at different heights as presented in the validation reports by Duncan et al (2019a) and Knoop et al (2019). Instead, the QQ plot of DOWA wind speed at two heights was compared to the same plot from either mast or lidar measurements. For the comparison to mast measurements, we took the heights 27 m and 86 m; for the comparison to lidar measurements, we took the heights 90 m and 290 m. Such a QQ plot provides information about the mean wind speed profile as a function of wind speed; differences between the QQ plots

for different data sources indicate differences in the mean wind speed profiles from these sources. This may be relevant information in wind resource assessment if the wind climate at some height above the mean sea level is to be extrapolated to another height.

The results are plotted in Figure 4. The QQ plots for wind at 27 m and 86 m from DOWA and mast data are almost linear. The QQ plots for wind at 90 m and 290 m from DOWA and lidar data are nonlinear: at wind speeds below 10 m/s, the speeds at 90 m and at 290 m are on average equal, but they differ by about 20% at wind speeds around 25 m/s.

The curves for mast and DOWA data agree closely; the same applies to lidar and DOWA data except at wind speeds exceeding 25 m/s, where the difference is somewhat larger. Therefore, overall, bias in DOWA wind shear appears to be quite low, possibly with the exception of high wind speeds above an altitude of 90 m.

Another way to look at differences in wind profiles is to examine the QQ plot of wind (speed) shear derived from different data sources. This provides a relationship between the probability distributions of the shear from these sources, but the connection to wind speed is lost.

These QQ plots are shown in Figure 5 for the wind speed at 86 m minus the wind speed at 27 m from mast and DOWA data, and for the wind speed at 290 m minus the wind speed at 90 m from lidar and DOWA data. For small positive wind speed differences, DOWA matches the measurements well. However, negative and large positive differences are smaller in magnitude for DOWA than for the measurements, so the spread in the wind speed difference is reduced in the DOWA data, compared to the measurements.

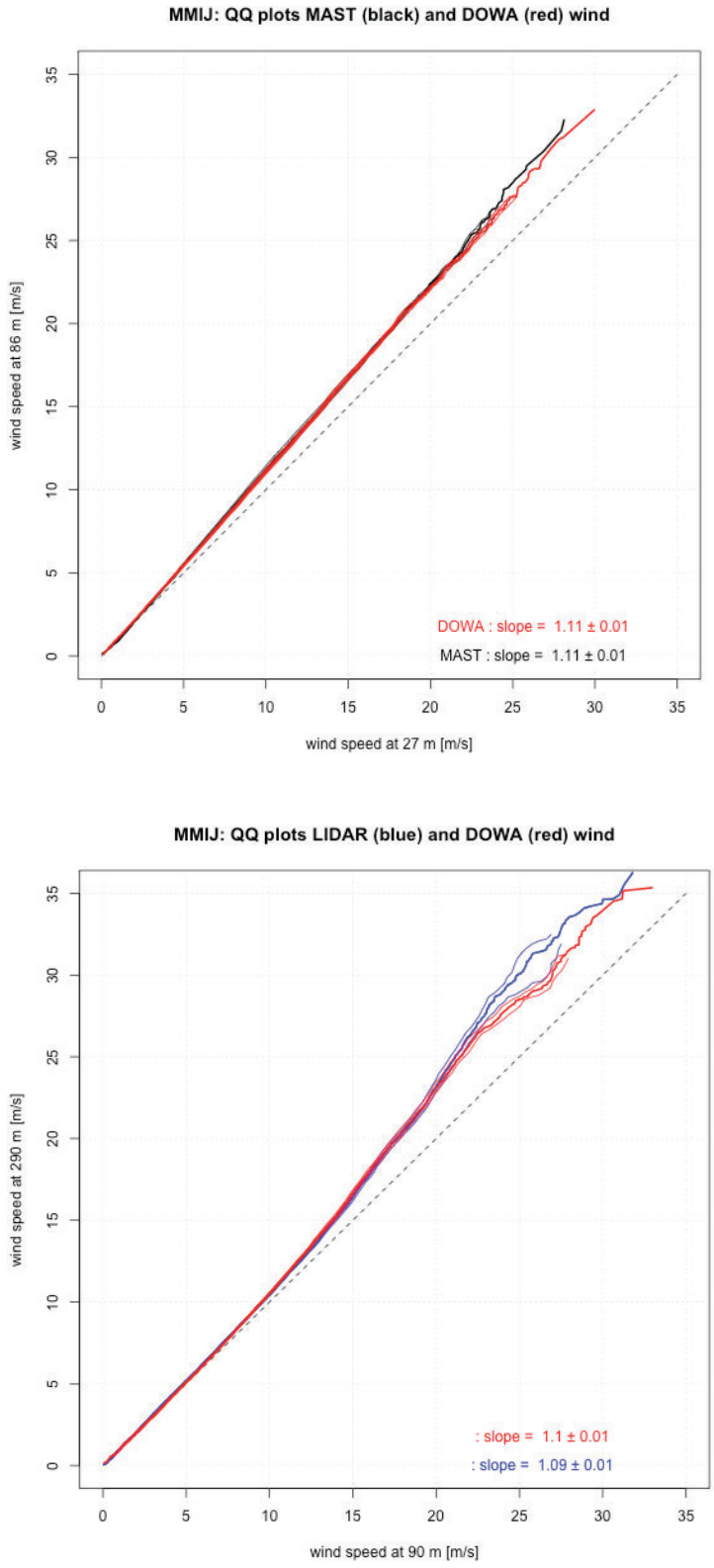

Figure 4: Quantile-quantile plots of wind speed data at two different heights for mast and DOWA data (top) and for lidar and DOWA data (bottom). Unit: m/s.

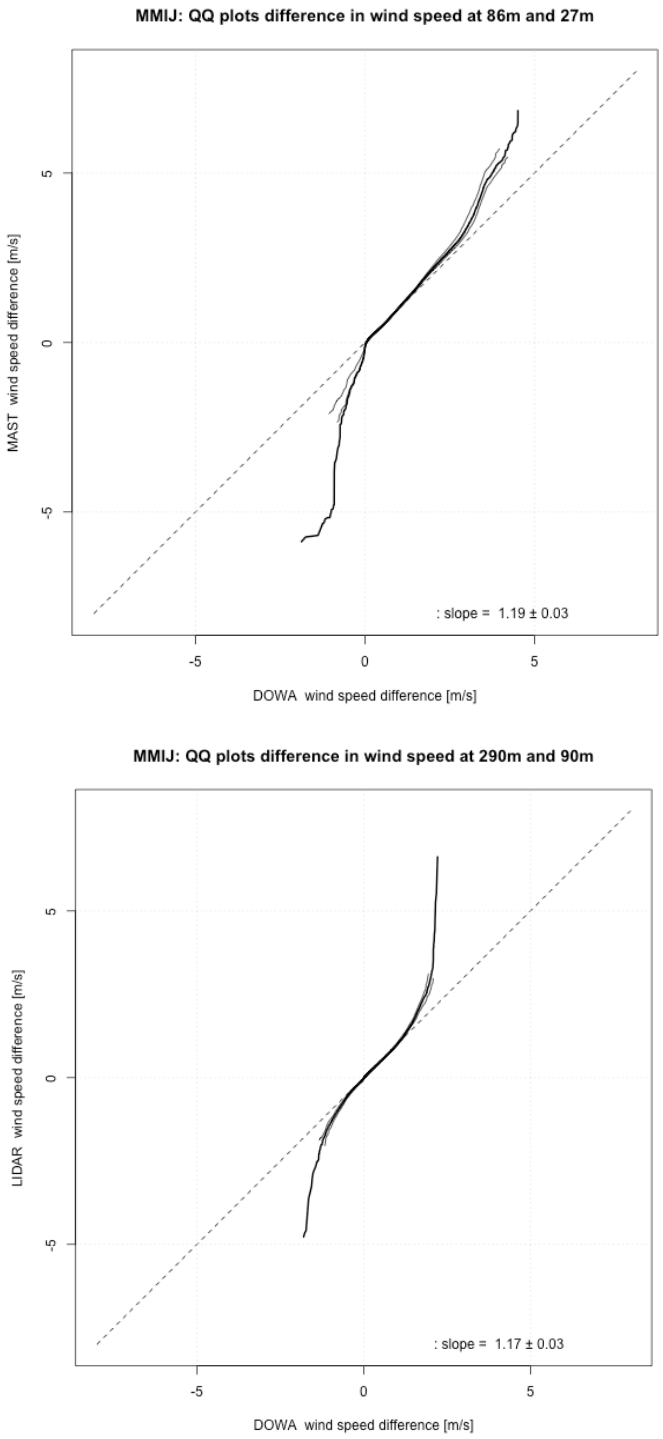

Figure 5: Quantile-quantile plots of the wind speed difference between 86 and 27 m from mast and DOWA data (top) and the wind speed difference between 290 and 90 m from lidar and DOWA data (bottom). Unit: m/s.

3. Uncertainty in wind climatology at hub height

3.1 Introduction

For wind resource assessment, the wind energy sector needs estimates of the directionally dependent distribution functions of wind speed; the directional information is needed for the modelling of wake effects and (on land) spatial variation of wind speed within a wind farm (e.g. due to orography)

To obtain these estimates, typically a 10 year record of wind from a weather prediction model or from measurements at a nearby site is used. The directionally dependent wind speed distributions are commonly approximated by Weibull distributions, estimated from the data.

In the present analysis, we establish the uncertainty of these statistics as derived from 10 years of DOWA data for the case of wind at a height of 90 m at the location of MMIJ. The uncertainty is represented by a an ensemble of such estimates, obtained from the data by a bootstrapping scheme. Bootstrapping means generating synthetic datasets having similar characteristics as the original dataset by (a) drawing random subsets from the original data **with replacement**, and (b) stringing these subsets together to obtain a synthetic sample; see Section 2.4.

We apply two different bootstrap schemes. The first scheme is based on the simplifying assumption that averages of wind statistics in different years are independent. This makes it possible to derive the uncertainty in a 10-year average from the observed variability in 1-year averages. The second scheme does not make this assumption: it derives an estimate of the uncertainty in a 10-year average from the variability of such averages based on a longer dataset; in this case, the 40 year long KNW-dataset.

3.2 Bootstrap without accounting for multi-year dependence

A total of 500 bootstrap samples of 10-year DOWA wind at a height of 90 m at MMIJ were generated. Each sample was obtained by randomly drawing ten 1-year records of DOWA wind from the 10-year record. From each bootstrap sample, Weibull directional wind speed distributions were estimated for each of the twelve 30-degree bins. These estimates were based on the left-censored maximum likelihood method using a censoring threshold of 2 m/s, which is below the typical cut-in speed. Censoring ensures that irrelevant low values of wind speed do not affect the Weibull estimates.

The resulting ensemble of bootstrap samples of

- probability of exceedance of the threshold
- scale parameter
- shape parameter

for each of the twelve directional bins represent the estimated uncertainty in the climatology. The data and code used to generate such data can be obtained from the author for experimentation.

To evaluate the uncertainty in the wind climatology, the median W50 of the 10-year average wind speed was compared to the W90, the 10-year average wind speed exceeded with a probability of 90% (W50 and W90 are comparable to the parameters P50 and P90 well-known in wind energy, where P refers to production and W to wind). The ratio W90/W50 was found to be 0.99, so this bootstrap experiment produces a narrow distribution of 10-year average wind speeds from the 500 bootstrap samples. Adopting a value of 1% for the bias error of the DOWA wind speed as inferred from Section 2.2, the total uncertainty in the 10-year average wind speed would not change much if bias is taken into account as well (as the standard deviation and bias of the 10-year average wind speed are both close to 1%).

To evaluate the uncertainty in reference yield (energy yield without accounting for wake effects and technical losses; see e.g. Mortensen et al (2015)), the ensemble of estimated Weibull parameters was converted to an ensemble of energy production of a single 3MW V90 turbine as deployed at OWEZ over a 10 year period. The power curve for this turbine was linearly interpolated between the tabulated values.

Neglecting other sources of uncertainty, we obtained P90/P50= 0.98, with P50 the median energy production and P90 the energy production exceeded with a probability of 90%.

This value is based on the assumptions that

(1) the variation of annual mean wind speeds over the decade 2008-2017 is representative of the climatology, and

(2) annual averages of wind statistics are independent.

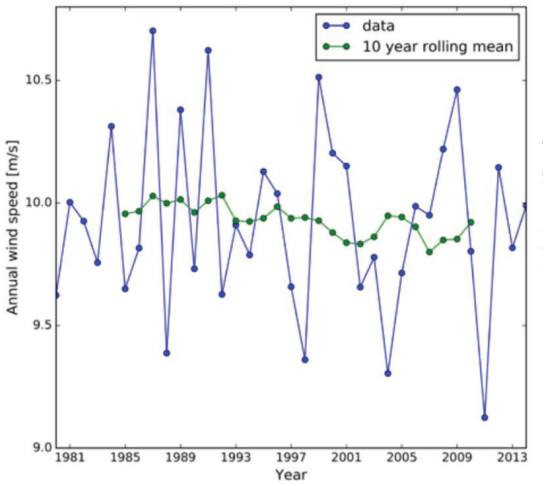

Figure 6: Annual mean wind speed and 1-year rolling mean wind speed at 100 m from KNW data at the location of offshore platform K13 (from http://projects.knmi.nl/knw/IAV_12062017_IW8.pdf).

Both assumptions are highly questionable. For example, Figure 6 shows that at K13 at a height of 100 m, the annual average wind speed varies considerably on multi-year time scales (in fact, more than would be the case if the consecutive annual mean wind speed values were independent). On the other hand, the variation of the 10-year average wind speed in Figure 6 does not seem large. However, this may be a coincidence, as we have only four disjoint 10-year intervals in the 40-years covered by the KNW dataset (and so the sequence of 10-year averages has at most 4 degrees of freedom). In the following section, we try to relax these assumptions.

3.3 Bootstrap accounting for multi-year dependence

From the 40-year KNW data at MMIJ, 500 bootstrap samples of 10 years of wind speed data at a height of 90 m were generated by:

- randomly choosing a starting year of the 10 year period within the first 30 years of KNW data
- resampling the years within this 10 year period as in the previous method.

This procedure is a rough approach to account for interannual variability, in particular because there are only four disjoint 10-year intervals in the 40 years of KNW. We would draw from data over a longer interval if such data were available. Resampling years within 10-year blocks as if annual averages are independent is also a simplification.

The estimate of W90/W50 accounting for multi-year dependence is 0.99, the same value as obtained in Section 3.2 without accounting for multi-year dependence or for the possibility that variability of annual wind speed within the DOWA decade 2008-2017 may not be representative. The estimate of P90/P50 for a single 3MW V90 turbine is 0.98, also the same value as obtained from 10 years of data without accounting for multi-year dependence.

These values of W90/W50 and P90/P50 are highly uncertain, as the block size of 10 years is rather large in comparison to the 40 year length of the data record. However, they indicate that wind variability on multi-year time scales does not contribute a great deal to the uncertainty in reference yield over a 10 year period.

For completeness, we also investigated how the uncertainty would change when using a 4-year dataset instead of a 10-year dataset. So from the 40-year KNW data at MMIJ, 500 bootstrap samples of 4 years of wind data at a height of 90 m were generated. Four years is by approximation the length of the measurement record at Meetmast IJmuiden, so this exercise illustrates the uncertainty due to inter-annual wind variability when using the measurement data from MMIJ

instead of the DOWA data. In this case, W90/W50 reduces slightly to 0.98 and P90/P50 to 0.97.

With a long wind dataset covering a much longer period than 40 years, we might obtain a much more reliable estimate of the effect of multi-annual dependence on the uncertainty in wind climatology from 10 years of data. There are no observational or reanalysis data covering such a long period. Data simulated by a climate model could be used as a surrogate. However, there is no way to verify whether the multi-annual dependence in the climate simulation is realistic, as this too would require a long measurement record.

4. Conclusions

The relative bias of DOWA wind speed is very low (of the order of 1%) in comparison to mast measurements and platform-mounted lidar measurements at Meetmast IJmuiden. For some 30° wind directional bins, this is somewhat larger. For KNW, bias is similar, although slightly lower for some directional bins. Bias in the relation between wind speeds at different altitudes in the DOWA data appears to be very low as well, possibly with the exception of high wind speeds at high altitude (90 - 290 m). The spread in wind shear is lower in DOWA than in the measurements.

These conclusions were derived from QQ (quantile-quantile) plots. The QQ plot has proven to be an accurate method to assess bias, resilient against low precision of the data and low correlation. Therefore, it is recommended as an alternative to regression-based MCP (measure, correlate, predict) methods for correction of bias in reference wind records in wind resource assessment.

The standard deviation (more precisely: the root-mean square non-bias error) of DOWA wind speed is estimated to be 1.23 m/s, as compared to 0.36 m/s for the platform-mounted lidar wind speed and 0.21 m/s for mast-mounted anemometer wind speed. For KNW, the value is 1.48 m/s, so DOWA is more precise.

The uncertainty due to variability of the 10-year average wind climate appears to be small: the ratio P90/P50 for the reference yield over 10 years is estimated to be 0.98, both from 10 years of DOWA data by assuming that annual averages of wind statistics are mutually independent, and from 40 years of KNW data by a method accounting for multi-annual dependence. However, being based on at most 40 years of data, these assessments are highly uncertain.

References

Amrhein, V., Greenland, S., McShane, B. (1019), Scientists rise up against statistical significance (comment). Nature **567**, 305-307.

Carta, J.A., Velàsquez, S., Cabrera, P. (2013), A review of measure-correlate-predict (MCP) methods used to estimate long-term wind characteristics at a target site. Renewable and Sustainable Energy Reviews 27, 362–400.

DNV (2011), Use of remote sensing for wind energy assessments. Recommended Practice DNV-RP-J101, Det Norske Veritas.

Duncan, J.B., Wijnant, I.L. , Knoop, S. (2019a), DOWA validation against offshore mast and lidar measurements (draft). TNO report 2019 R10062.

Duncan, J.B., Marseille, G.J., Wijnant I.L (2019b), DOWA validation against ASCAT satellite winds (draft). TNO report 2018 R11649.

JCGM (Joint Committee for Guides in Metrology) (2008), Evaluation of measurement data — Guide to the expression of uncertainty in measurement. Document 100:2008.

King, C. and Hurley, B. (2005), The SpeedSort, DynaSort and Scatter wind correlation methods. Wind Engineering 29, 217–41.

Knoop, S, Ramakrishan, P., Wijnant, I.L. (2019), Validation of DOWA against Cabauw meteomast wind measurements (draft), KNMI Technical Report.

Knoop, S. (2019), DOWA validation against coastal wind lidar measurements (draft), KNMI Technical Report.

Künsch, H. R. (1989). The Jackknife and the Bootstrap for General Stationary Observations. Ann. Stat. 17 (3): 1217–1241.

McColl, K.A., Vogelzang, J., Konings, A.G., Entekhabi, D., Piles, M., Stoffelen, A. (2014), Extended triple collocation: estimating errors and correlation coefficients with respect to an unknown target. Geophys. Res. Lett. 41, 6229–6236.

Mortensen, N. G., Nielsen, M., & Ejsing Jørgensen, H. (2015). Comparison of Resource and Energy Yield Assessment Procedures 2011-2015: What have we learned and what needs to be done? In Proceedings of the EWEA Annual Event and Exhibition 2015 European Wind Energy Association (EWEA).

Stepek, A., M. Savenije, H. W. van den Brink, and I. L. Wijnant (2015), Validation of KNW-atlas with publicly available mast observations (Phase 3 of KNW project). KNMI Technical Report (TR-352).

Stoffelen, A. (1998), Toward the true near-surface wind speed: error modeling and calibration using triple collocation. J. Geophys. Res. 103, 7755–7766.

Von Storch, H. and Zwiers, F.W. (2003), Statistical analysis in climate research. Cambridge University Press.

Wijnant, I.L., Brink. H. W. van den, Stepek, A. (2014), North Sea wind climatology Part 1: a review of existing wind atlases. KNMI Technical report (TR 342): http://bibliotheek.knmi.nl/knmipubTR/TR342.pdf.

Wouters, D.A.J. and Wagenaar, J.W. (2016), Verification of the ZephIR 300 LiDAR at the ECN LiDAR Calibration Facility for the offshore Europlatform measurement campaign. Report ECN-E—16-029, ECN.

Appendix A: Spectral analysis of wind speed data

Figure A1 shows the frequency spectra of collocated wind speed data from mast (black), lidar (blue), DOWA (red) and KNW (magenta). It shows that DOWA can reproduce the sizes of wind fluctuations at all periods from 3 years (frequency almost 0.0) to 2 hours (frequency is 0.5/hour). The KNW spectrum is only slightly lower than the DOWA spectrum[5]. This means that the wind fluctuations in DOWA and KNW on these time-scales are **realistic**. However, they are not necessarily **real** (that is, matching the measured fluctuations).

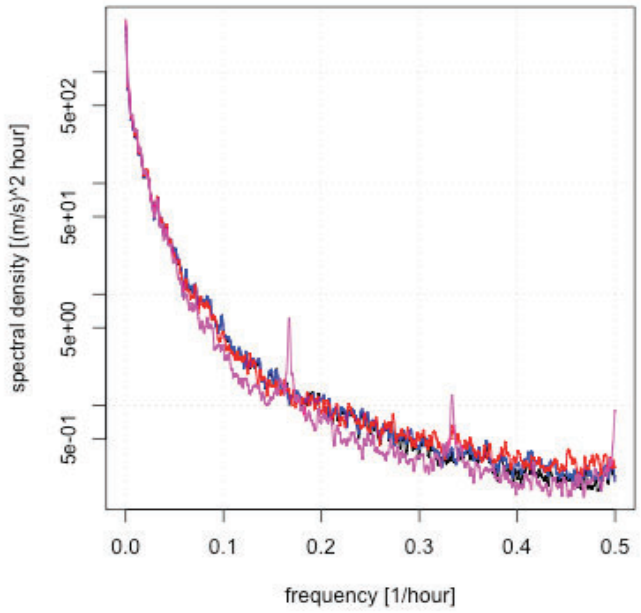

Figure A1: Frequency spectra of collocated wind speed data from mast (black), lidar (blue), DOWA (red) and KNW (magenta). Note that spectra are plotted on a logarithmic scale.

How real the wind fluctuations of DOWA and KNW are is shown by Figure A2. The coherence function[6] between mast and lidar measurements is high over the entire frequency range (black lines). The lowest coherence between mast and lidar measurements is 0.6 for frequency 0.45 (a period of 2.2 hours). The coherence between DOWA and the mast measurements (blue) is much lower. If we assume that a coherence of 0.6 or higher is good (which is approximately the minimal value of the mast/lidar coherence), then we conclude that the coherence between DOWA and mast measurements is good for frequencies below 0.1/hour. So based on this, the DOWA values are real for averaging times

[5] The narrow peaks at about 2, 3 and 6 hours (frequencies of 0.5, 0.33 and 0.17/hour) in the KNW spectrum are most likely related to the 6-hourly restarts.
[6] Also know as "coherency squared"; it can be regarded as the square of a correlation coefficient between the components of the two signals lying within a narrow frequency band. Phase shifts do not reduce the coherence.

above 10 hours. The coherence between the DOWA and the lidar measurements is good for frequencies below 0.07/hour (periods above 14 hours).

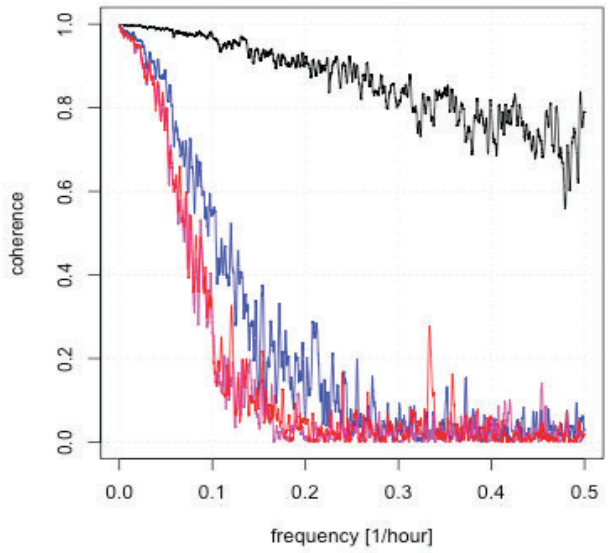

Figure A2: Coherence spectra of collocated wind speed data: mast/lidar (black), mast/DOWA (blue), lidar/DOWA (red) and DOWA/KNW (magenta).

We may conclude from this that the DOWA and KNW reanalyses generate realistic wind speed fluctuations on all time scales, but these agree with the observed wind speed fluctuations only at periods roughly above ten hours.

It is very well possible that the wind fluctuations occur, but not exactly at the location or at the moment that DOWA/KNW projects them (for example, two grid points away or an hour later, and for some other event, three hours earlier). This is one of the reasons why the quality of models often appears to decrease upon increasing the resolution: a model is penalized twice if an event is predicted where it did not occur and no event is predicted where it does occur. This is called the double penalty effect. In low-resolution model output, small-scale features are smeared out (smoothed), which tends to improve the performance statistics. In fact, DOWA provides more real detail that ERA5 (Duncan et al, 2019b), which is remarkable.

For wind resource assessment, it is important that the wind fluctuations of DOWA and KNW have realistic magnitudes at all time scales from 2 hours to 3 years. The similarity between the spectra of DOWA and KNW and the spectra of the measurements shows that this is the case, so the hourly DOWA and KNW data can be safely used as a reference wind data set for, for example, dynamic simulation of power production.

Combining the spectra and coherence functions, we may obtain an estimate of the noise spectrum of DOWA, i.e., the spectrum of wind speed fluctuations in DOWA unrelated to the observed wind. Assuming that the mast measurements have very low noise (as appears from the triple and quadruple collocation analyses in Section 2.3), we may estimate the noise spectrum of DOWA as $S(1-C)$, with S the DOWA wind speed spectrum and C the coherence between DOWA and mast measurement; Figure A3 shows that the wind speed spectrum of DOWA (red) almost coincides with its noise spectrum (cyan) for periods below roughly 7 hours (frequencies above 0.15/hour), i.e. the DOWA wind speed at these time scales is almost entirely noise.

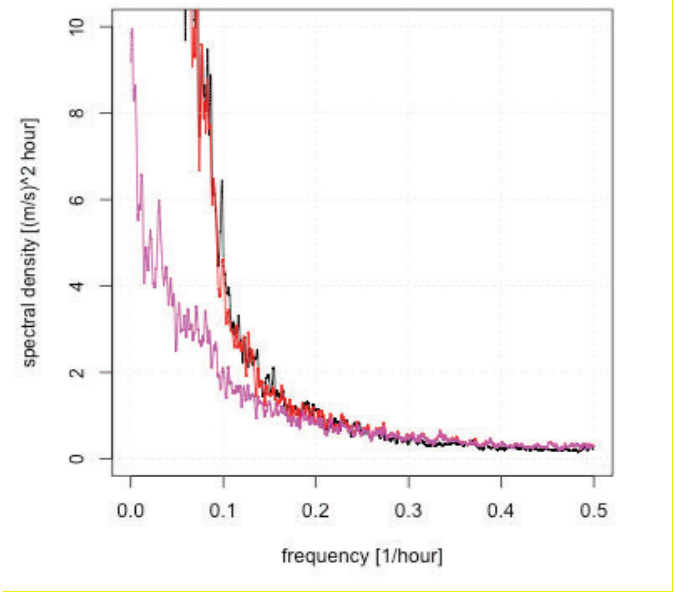

Figure A3: Spectra of mast (black) and DOWA (red) wind speed, and estimated noise spectrum of DOWA (magenta). Note that spectra are plotted on a linear scale.

Royal Netherlands Meteorological Institute

PO Box 201 | NL-3730 AE De Bilt
Netherlands | www.knmi.nl