Processing of sonic anemometer measurements for offshore wind turbine applications

IOP Conf. Series: Journal of Physics: Conf. Series **1356** (2019) 012006 doi:10.1088/1742-6596/1356/1/012006

Processing of sonic anemometer measurements for offshore wind turbine applications

A Nybø, F G Nielsen and J Reuder

Geophysical Institute, and Bergen Offshore Wind Centre (BOW), University of Bergen, Norway

E-mail: astrid.nybo@uib.no

Abstract. Quality assured measurements from offshore masts may provide valuable information of the characteristics of the offshore wind field, which is of high relevance for simulations of offshore wind turbines' dynamic response. In order to obtain these high quality data sets, a processing procedure tailored to offshore wind turbine applications must be followed. In this study, existing quality control routines applied in literature are evaluated, and a complete procedure is developed for sonic anemometer measurements. This processing procedure is applied to measurements at three heights from 16 months of measurements at FINO1. The processing procedure results in a data set of more than 6 000 30-minute periods of high quality time series showing a large variety in terms of wind speed and turbulence intensity. Together with an assessment of the stationarity, this processed data set is ready for use in offshore wind turbine research.

1. Introduction

Wind turbines in operation are exposed to and have to react on a complex flow environment, where vertical wind shear, turbulence intensity, coherence and atmospheric stability are closely interconnected [1–3]. The analysis of their dynamic response requires the understanding of the offshore wind characteristics and thus the availability of high quality wind measurements offshore.

Observations are rather sparse, as the infrastructural requirements connected to offshore measurements are considerable. In-situ measurements with cup or sonic anemometers require a mast or tower structure that is expensive, limited to one location and that causes additional disturbances in the flow measurements from certain directions. Lidar-based remote sensing can be performed from offshore platforms, buoys or ships and provides a more flexible measurement setup with respect to localization. The inherent volume averaging in the lidar measurements limits, however, the use of those data for the characterization of turbulence [4].

Sonic anemometers are robust instruments without moving parts that can provide high frequency measurements (typically 10 to 100 Hz) of the 3D wind field and the sonic temperature even in harsh offshore environments. The preparation of the raw data for further analysis requires a thorough and extensive quality control procedure, that removes unphysical values (e.g. spikes) or situations where the measurements are disturbed by flow-distortion or precipitation, assesses stationarity and organizes the data in appropriate time windows for the relevant applications. Quality control routines for sonic anemometer measurements have been described by e.g. Foken and Wichura [5], Vickers and Mahrt [6] and Foken et al. [7, 8]. They mainly focus on data

processing for the use in boundary layer meteorology, where a correct representation of the turbulent fluxes of heat, moisture and momentum is of highest importance. All the above publications mention spike detection procedures and stationarity tests, which are relevant for wind speed analysis as well. Time series designated for the use in load simulations for turbines might, however, require a more specific handling and quality control. The time resolution has to be high enough to allow for the simulations of high frequency phenomena, as, e.g., blade tip vibrations. The data set has, on the other hand, to be long enough to cover the relevant low frequency motions as well. Natural periods of a floating wind turbine are in the range of minutes. This requires time histories of wind in the order of at least 30 minutes. State of art turbines have rotor diameter of approximately 200 m, i.e. the height range from 20 m to 250 m above sea level is of interest in wind turbine applications.

The wind measurements from the FINO1-platform, assessed in this study, have earlier been analyzed from an offshore wind turbine perspective (e.g. [3, 9–13]). The quality control routines of the sonic data in the mentioned literature include the removal of periods where the wind origins from certain directions in order to exclude disturbances due to mast shadow, and the application of a stationarity test or removal of trends. The studies by Cheynet et al. [3, 9] also include the organization of the data in longer periods, the removal of situations where the turbulence intensity is not within reasonable limits and the correction of tilt angles of the sonic anemometers.

This study describes a thorough processing procedure of sonic wind measurements, starting from available raw data and ending in a complete set of processed time series. These measurements may further be used in analyzing characteristics of the offshore wind field relevant to the wind turbine, or even as input in wind turbine simulations in the design phase of offshore wind turbines. The main objective is to achieve a wide variety of undisturbed measurement situations for such further analysis.

2. Data

The measurements used in this study are performed at the German wind energy research platform FINO1 [14], located in the German Bight of the North Sea, approximately 45 km North of the German Island of Borkum, as shown in Figure 1a. FINO1 was deployed in July 2003 and consists of a jacket foundation carrying a 16x16 m working platform (20 m above mean sea level), an elevated helicopter deck (25 m) and an 81.5 m high meteorological mast with its highest measurement level at 101.5 m above mean sea level [15].

The mast is densely equipped with standard meteorological sensors for wind, temperature and humidity at levels between 30 and 100 m [18]. High frequency wind measurement by sonic anemometers (Figure 1b) are operationally performed at 40, 60 and 80 m. The 40 and 80 m data are sampled at 20 Hz, while the 60 m data are sampled at 10 Hz. The sonic raw data are stored in blocks of approximately ten minutes duration, as commonly used in wind research and engineering. The present study uses wind speed in three directions and temperature from the mentioned sonic anemometers from June 2015 to September 2016. This period coincides with an extensive offshore field campaign at FINO1 (OBLEX-F1, [19]) and was chosen due to the availability of a wide range of complementary met-ocean measurements that might be beneficial for future investigations. In this study, wind direction information from wind vanes is used to calibrate the sonic anemometers, and precipitation information, in the form of a flag indicating "rain" or "no rain", is used to remove disturbances.

For reference, data from cup anemometer measurements[1] are also presented. Measurements made at 100 m are used for the wind speed in order to avoid flow distortion due to the mast. These cup anemometer data are used together with wind vane measurements of wind direction at

[1] http://fino.bsh.de/

(a) (b)

Figure 1. Map of FINO1 and surrounding wind parks at the time of measurements (dark blue) in the North Sea (a) [16], and layout of the FINO1 mast (b) - the red dots indicate the locations of the sonic anemometers (modified from [17] © Forschungs- und Entwicklungszentrum Fachhochschule Kiel GmbH)

90 m. The availability of the 30-minutes blocks of cup anemometer and wind vane measurements for our investigation period is 95 and 96 % respectively.

As shown by Figure 1a, the wind speed measured at FINO1 may be influenced by nearby land to the East and South, in addition to surrounding wind farms. However, the wind coming from North-West, where the sonic anemometers are pointed towards, represents close to undisturbed offshore conditions.

3. Method

The measurements undergo a processing procedure that is followed by a stationarity assessment, illustrated in Figure 2.

3.1. Missing and formatted measurements

Data are missing for some shorter periods, probably due to defects and maintenance on the measurement equipment. In cases where more than ten seconds are missing in a ten-minute period, the complete ten-minute period is disregarded.

Some additional ten-minute periods are disregarded due to issues with the data logger causing corrupted data files. These files are characterized by symbols at random positions in the data files, causing erroneous numbers.

The ten-minute blocks are not evenly distributed with time series of exactly ten minutes, but consists of e.g. a period of ten minutes and two seconds followed by a period of nine minutes and 58 seconds. In order to even these time series out, they are first organized in daily vectors and thereafter split into intervals of exactly ten minutes duration for further processing. Values are mirrored/deleted at the end of the day or before a missing ten-minute period when only a few seconds are missing/too many data points exist.

3.2. Spike detection and removal

The next step of the processing procedure of the measured values is a spike detection and removal routine. Spikes are considered as unphysical outliers in the data, which may origin from single

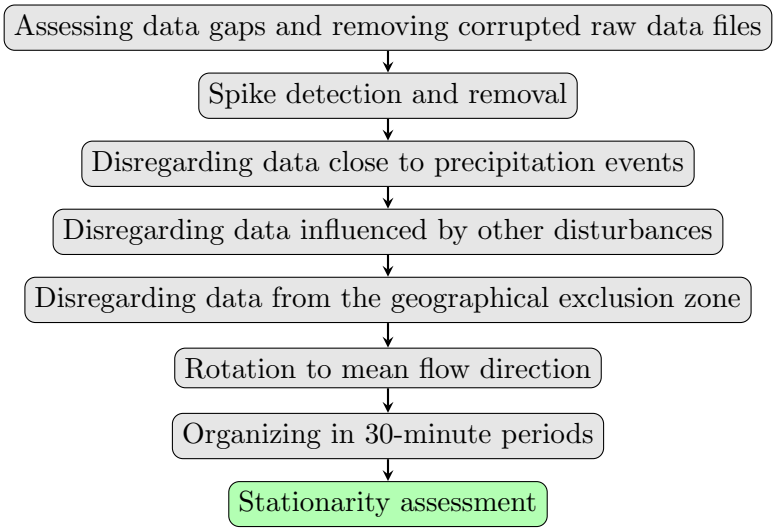

Figure 2. Scheme of the processing procedure followed by a stationarity assessment

water droplets or other sources of measurement errors. In the literature, the spike detection method by Højstrup is widely used [20]. Vickers and Mahrt [6] and Schmid et al. [21] describe methods based on his work, where spikes, x'_{spike}, are defined as extreme deviations from the mean, such that $|x'_{spike}| > a * \sigma$, where a is a scaling factor (3-4 is normally used) and σ is the standard deviation of the process. The methods are iterative, meaning that the mean and the standard deviation are computed again after the first spike removal, as the spikes would affect the statistics of the first iteration. The methods proposed by Vickers and Mahrt and Schmid et al. vary slightly, but they are both dependent on the iteration process, which is a disadvantage when handling very large data sets. Papale et al. [22] (based on Sachs [23]) proposes a similar method which is rather based on the median, avoiding the iteration process.

In this study, a method based on Rinker et al.'s [24] spike detection and removal procedure is applied. This method is chosen as it avoids iteration, it detects large gradients instead of data points, and it easily finds spikes extending over several data points. The method is based on calculation of differences between neighboring samples, Δx. If any of these differences are larger than a certain limit, they are defined as "extremes", Δx_{ext}. The limit is determined by a scaling factor, a, times the standard deviation of the differences between neighboring samples. When calculating this standard deviation, the spikes are excluded in the statistics by only using a given threshold, b, of the differences which have the lowest values, $\sigma_{\Delta x,b}$. By following this approach, the iteration procedure is avoided. In this study, an additional parameter, c, is introduced in the calculation of the limit in order to avoid detecting peaks that are large compared to surrounding values, but very low in absolute value. The spike detection and removal procedure is shown in Figure 3. A spike is detected when "extremes" are found at the beginning or end of each ten-minute interval, or two "extremes" with opposite directions are found within a maximum number of samples, also referred to as the spike width, N_{spike}. The spikes are replaced by the value before/after if found at the end/beginning of the ten-minute interval, and otherwise replaced with interpolation of the values before and after the spike.

The parameters required by the method of [24] are adjusted to detect the intended spikes of the sonic measurements of this study, corresponding to a spike width of six data points ($N_{spike} = 6$), a scaling factor of five ($a = 5$) and a threshold of 99 % ($b = 0.99$), in addition to the mentioned introduced parameter of 0.3 ($c = 0.3$). It is found that there is in general very

few spikes in the time series. However, some of the spikes may have very large values, thus a proper identification and removal routine is important to the quality of the time series. In the processed data set at all heights, there is in average 0.00002 % spikes, with the maximum in one processed 30-minute period being 0.2 %. All periods have far less spikes than the recommended maximum level by Vickers and Mahrt, 1 %, hence no ten-minute period is removed based on this criteria.

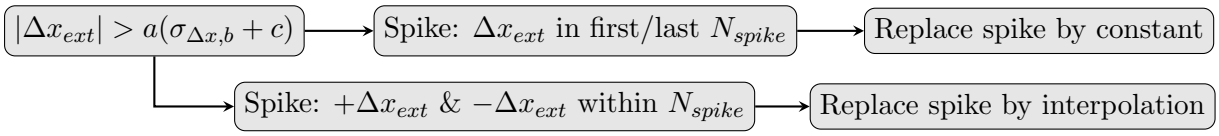

Figure 3. Scheme of the spike detection and removal procedure

3.3. Precipitation

It is common practice to exclude periods during and following precipitation when processing sonic anemometer measurements, as the precipitation may cause erroneous measurements [6, 8, 25, 26]. A consequence of this approach is, however, that the resulting time series are biased towards situations without precipitation. The precipitation information in the present study is limited to "rain" or "no rain" as mentioned in Section 2. As water droplets may stick to the transducers even after precipitation ends, disregarding the period of precipitation alone might not be sufficient. In this study, we chose to discard every ten-minute period when precipitation is registered, in addition to the preceding ten and succeeding 50 minutes.

3.4. Other disturbances

Particles located between the pair of transducers can potentially disturb the measurements. In addition to precipitation, the measurements may be affected by fog, frost, sea spray or larger aerosol particles. As we do not have any measurements to indicate these disturbances, and the study includes a few months where no precipitation information is available at all, we need to rely on characteristics observed during rainy periods in order to exclude these disturbances. In agreement with [25], we see a significant increase in standard deviation of the sonic temperature in rainy situations. We chose to exclude ten-minute periods where this standard deviation is higher than 0.3 K. A stricter threshold of 0.15 K is chosen for the months where no rain information is available. This approach may also remove some periods with significant trend in temperature, but these situations are probably not stationary and therefore anyway undesirable for further use.

3.5. Geographical exclusion zone

The sonic anemometers are pointed towards Northwest, being in the mast shadow when the wind blows from Southerly and Easterly directions, as shown by Figure 1b [27]. The orientation of the sonic anemometers and the potentially disturbed sector, also referred to as the "geographical exclusion zone", is shown in Figure 4. Wind blowing from North corresponds to 0 degrees, while wind from East corresponds to 90 degrees. We chose, in accordance with previous studies based on FINO1 sonic anemometer data [3, 9–11], to exclude the periods of potential disturbances. In this study, a very conservative approach is chosen, where all ten-minute periods with a mean wind direction between 45 and 225 degrees are excluded. The excluded wind directions, with winds coming from South and East, are also influenced by the proximity of land as they have already blown over the Netherlands and Germany (Figure 1a). The remaining data set is mostly undisturbed by such effects and thus represents realistic offshore wind conditions.

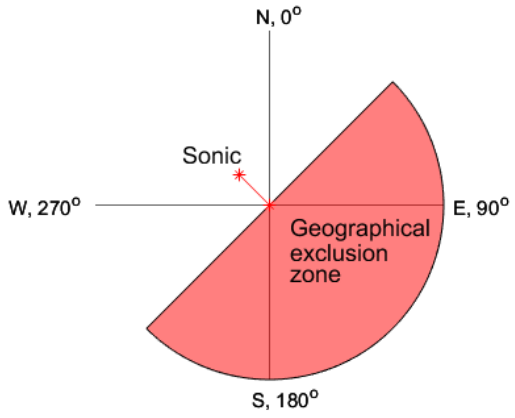

Figure 4. Illustration of the location of the sonic anemometers relative to the mast and the excluded wind directions (the geographical exclusion zone)

3.6. Rotation

For further analysis, e.g. related to the inflow of a wind turbine, we are interested in the mean wind speed in the mean flow direction, rather than the three components given by the sonic anemometers. There are mainly two ways of rotating the wind into the mean flow direction, the planar fit method and the double rotation method [28]. The two methods are quite similar, but treat the vertical velocity differently. The planar fit method only depends on a fixed tilt angle of the anemometer, allowing for a non-zero mean vertical velocity over the averaging periods. This is advantageous for a correct calculation of the fluxes, and the planar fit method is therefore widely used in boundary layer meteorology. The double rotation method depends on both a fixed tilt angle and the wind direction, leaving a zero mean vertical wind speed in all periods.

The double rotation method is chosen in this study mainly because wind turbine simulations are commonly performed with zero mean vertical velocity, which is also recommended or required by certain simulators. This may, however, cause some uncertainty in flux calculations, e.g. for determining stability. The double rotation method is used for all heights for 30-minute periods, but the 40 and 60 m data are rotated back to the mean flow direction at 80 m. In the processed data set at all heights, the average wind veer between 40 and 80 m corresponds to only one degree.

3.7. Organizing in 30-minute periods

The data are finally grouped in 30-minute continuous periods. During 30 minutes, we expect that all relevant frequency ranges for the wind turbine simulations are included, ranging from the very high frequencies relevant for blade tip vibrations to the slow floater motions requiring far longer periods. If ten minutes are missing in one 30-minute period, the whole 30-minute period is disregarded. The data may further be grouped into one hour periods, often used for wind turbine simulations, but the same 30-minute period may also be used twice with a smooth transition, as all relevant frequency ranges are assumed resolved within 30 minutes.

3.8. Stationarity

Natural wind is not a stationary process. However, most analysis of wind turbines assumes stationarity. In using measured wind data, a decision of what is "sufficient stationarity" must be made. As this criterion is strongly case dependent, we have chosen not to include stationarity tests as part of the processing procedure. A quantitative evaluation of the stationarity of a clearly non-stationary process, such as the wind speed, is also not straightforward. Including such tests could therefore remove close to stationary periods or even include too many non-stationary periods.

As for the previously introduced processing procedures, a qualitative stationarity analysis would have been an easier and more precise approach, but the amount of data excludes this

alternative. A common quantitative approach is developed by Foken and Wichura [5]. This method evaluates the variability of the mean of shorter intervals, e.g. five to ten minutes against the total mean of e.g. 30 to 60 minutes. Mahrt [29] has developed a similar test evaluating the standard deviation, which is combined with the Foken test by Cava et al [30]. These methods are widely used in meteorology, but their relevance to wind turbine applications may be questioned. The methods are developed for fluxes, not wind speeds. In addition, a time series with sudden small gusts will be considered stationary according to these methods. Several other methods exist in meteorology, but similar to the mentioned methods, most of these have not been developed with wind speed applications in mind. Ohbrai et al. [31] applied a stationarity test on FINO3 wind data, based on maximum variation between consecutive ten-minute mean values. This method could be used in our study with a maximum variation customized to the frequency of the measurements, but is disregarded as it would not capture non-stationarity due to a slow, but significant trend over the time series. Cheynet et al. [3, 9] has applied a two-step stationarity test in his studies. The first step is a linear trend test, which is adapted in this study and explained in the next paragraph. The second step is a reverse arrangement test described in Bendat et al. [32] and previously used for wind measurements by Chen et al. [33].

The processed time series of this study are already expected to be relatively stationary as rainy periods and periods with high standard deviation of temperature are removed. However, it is anyway recommended to test stationarity, e.g. by the two-step process explained in this paragraph. We may avoid de-trending the time series by rather applying the first step of the stationarity test of Cheynet [3, 9], in order to minimize the altering of the raw data. A maximum slope of 20 % is accepted over one hour ($\Delta t = 60min$), as shown by equation 1.

$$\frac{|\Delta U_{\Delta t}|}{\overline{U_{\Delta t}}} \leq \frac{0.2 * \Delta t(min)}{60min} \tag{1}$$

The second step requires that the maximum moving mean and moving standard deviation of each ten-minute block is less than 40 % off from the hourly mean and mean of the moving standard deviation.

$$max\left(\frac{|\overline{U_{60min}} - \overline{U_{10min}}|}{\overline{U_{60min}}}\right) \leq 0.4 \text{ and } max\left(\frac{|\overline{\sigma_{U_{10min}}} - \sigma_{U_{10min}}|}{\overline{\sigma_{U_{10min}}}}\right) \leq 0.4 \tag{2}$$

This simple two-step method gives a rough picture of the stationarity. For further precision, the frequency spectrum of the velocity variations in different parts of each time series should be evaluated, with special focus on the frequency range close to resonances of the wind turbine to be considered.

4. Results

After the processing, a complete data set ready for analysis is available. In the following paragraphs, we will visualize parts of the processing procedure for an example period, show the fractions of periods removed due to the different processes and at last present an overview of the complete processed data set.

The first four steps of the processing, assessing data gaps, spike detection and removal, removal of rainy periods and removal of periods of other disturbances, are visualized in Figure 5 for one day in September 2015. It is noticeable that the time series are discontinuous, especially towards midnight where we see many gaps. These gaps correspond to ten-minute intervals of missing data or corrupted raw data files. We can also observe the spike detection and removal process close to 18:00. The spike is probably detected in the temperature time series, but an interpolation is performed on all measured quantities. However, we observe another spike close to the mentioned one at 18:00 in both horizontal velocities, which is not detected by the spike

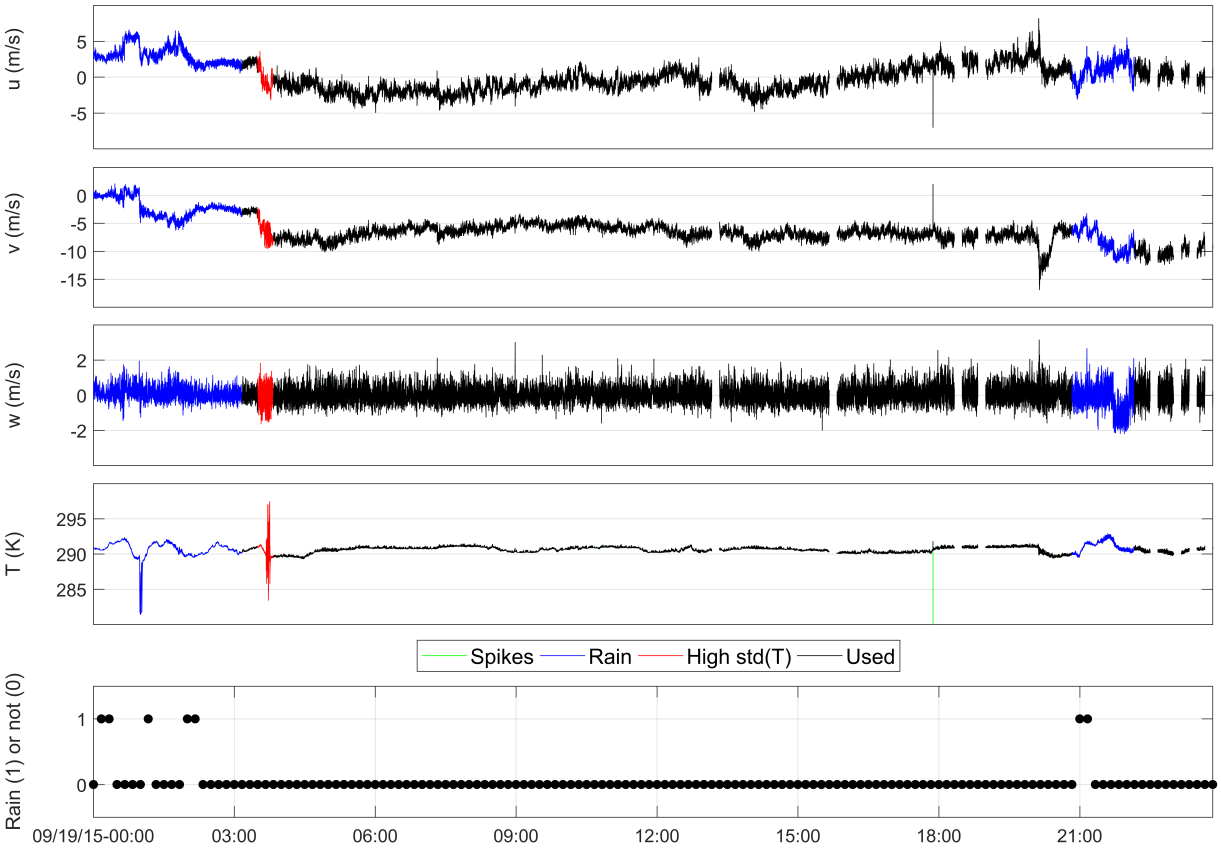

Figure 5. Visualization of the first four steps of the processing procedure with example time series of the three wind components and temperature, in addition to rain information for one day

routine. This is due to the length of the spike being longer than the chosen spike width of six data points used in the procedure. However, there is a sudden increase, and a sudden decrease with a plateau in between, typical for an unrealistic spike. This is therefore a good example that the processing procedure is not perfect, and that a manual inspection of certain time series may be necessary. The last subplot shows if there is rain or no rain in each ten-minute interval. There is a clear relation between precipitation and increased sonic temperature fluctuations. The figure shows that periods are removed both ten minutes before and 50 minutes after the precipitation is registered. In both rainy periods, it is obvious that this approach is necessary as the high fluctuations continues for a long time after the precipitation is registered. At about 04:00, it may even seem that a longer interval after precipitation should have been removed due to persistent large variations in both temperature and velocities. However, this ten-minute interval is removed in the next processing procedure that disregards periods of high temperature standard deviations. As shown in the figure, when unrealistic behavior is noticed at either one of the parameters, all measurements are removed. This is both due to the measurement principle being the same, so unnatural behavior is probably present in all parameters even though it is only obvious in one, and because all four parameters are necessary for further analysis. As mentioned earlier, it is also important to highlight that the processing procedure is not perfect, which is clearly shown by the non-detected spike close to 18:00 in this plot.

The data availability is distinctly reduced by the processing procedure, leaving only 27 %

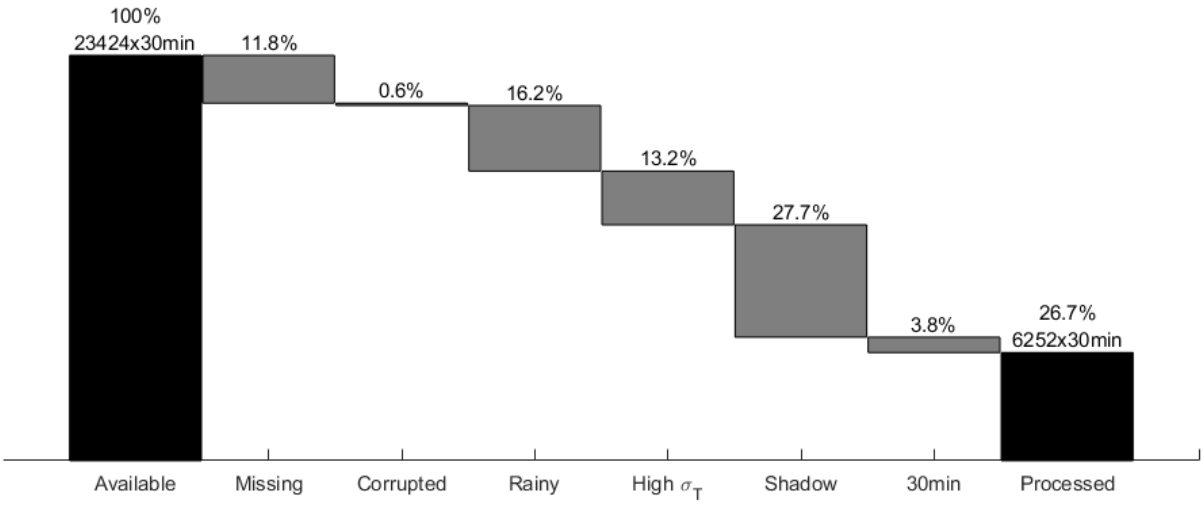

Figure 6. Total number of 30-minute periods from June 2015 to September 2016, fractions of removed data at all heights relevant to the total number of periods, and net remaining availability after processing of the sonic anemometer data at all heights

Table 1. Fractions (%) of removed data at all heights relevant to the total number of periods per season, and remaining availability (%) after processing of the sonic anemometer data at all heights in the last column

Season	Missing	Corrupted	Rainy	High σ_T	Shadow	30min	Processed
`Summer 15`	53	1	0	20	11	4	12
`Autumn 15`	5	2	25	7	34	4	24
`Winter 15/16`	0	0	28	9	36	3	23
`Spring 16`	3	0	14	22	28	5	28
`Summer+Sept. 16`	1	0	15	9	30	4	42

of the entire period available for further analysis. However, this still corresponds to 6252 30-minute periods. The applicability of the processed data set depends on the aim of the further analysis of the data, but 6252 30-minute periods of high frequency data representing a large variety of offshore wind characteristics are more than enough to have significant relevance in offshore wind turbine research. Figure 6 shows the amount of data removed in each step of the processing procedure. The largest contributors are originally missing data, periods removed due to precipitation or high temperature standard deviation and periods where the wind origins from the geographical exclusion zone. Even though the latter excludes only 28 % of the complete period, it excludes close to 50 % of the available data after previous processing steps. The overview presented in Figure 6 is representative for measurements at all heights, meaning that if only the wind speed at 40 m is in the exclusion zone, periods of all heights are removed. The same procedure is followed for Table 1 and all following figures as well, but separate data sets are saved for each height in case only specific heights are interesting for further analysis. As mentioned earlier, we do not consider the stationarity assessment as part of the processing procedure, and the available time series presented in the figures of this section are therefore not

assessed in terms of stationarity yet.

Despite the reduced availability in the processed data set, it covers a wide range of relevant environmental conditions. Table 1 shows that all seasons are well represented, with the summer of 2016 compensating for the summer of 2015 having large amounts of missing data. Figures 7-9 present an overview of the processed data set in terms of wind speed, turbulence intensity and direction at 80 m height. The turbulence intensity, defined by the ratio of the standard deviation to the mean wind speed, is based on 20 Hz data over 30 minutes. It is clear from these figures that a wide range of offshore conditions is present in the processed data set, with wind speeds from zero to 24 m/s and turbulence intensities from a few percent to extreme values of more than 40 %. By comparing to the cup anemometer data with far better coverage (indicated by the red line in Figure 7 left and small wind rose in Figure 8 left), it is shown that the conditions covered by the sonic data set are representative for the selected period. There is a clear trend in the relation between the wind speed and the turbulence intensity, but a wide variety of turbulence intensity may still be found for most wind speeds. Figure 8 clearly shows that data with directions between 45 and 225 degrees are removed, but all other directions are represented in the processed time series.

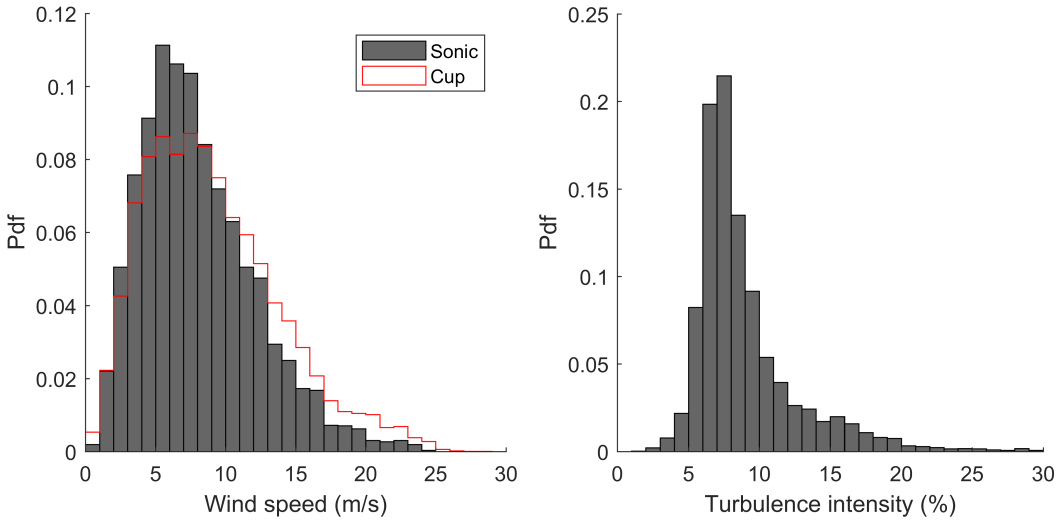

Figure 7. Probability density functions of 30-minutes mean wind speed (left) and turbulence intensity based on 20 Hz data over 30 minutes (right) at 80 m elevation after the processing procedure

Figure 9 shows the relation between wind speed and turbulence intensity and compares the 90th percentile of the FINO-1 data of this study, the IEC standards[2] [34, 35] and the findings of Ernst and Seume [13] and Türk and Emeis [12]. As expected, the measurements of this study fits better with the IEC offshore standard than the onshore, with an especially good fit above ten m/s. They are also well correlated with other FINO1 findings, but some deviations are present at intermediate wind speeds. The higher turbulence intensity of this study may be explained by the lower position of the sonic anemometer, the longer averaging interval, the higher frequency of measurements, and the difference in measurement device.

[2] Onshore: $\sigma = I_{ref}(0.75V_{hub} + 5.6m/s)$. Offshore: $\sigma = \frac{V_{hub}}{ln(z_{hub}/z_0)} + 1.28 * 1.44 * I_{ref}$ where the roughness length is given by $z_0 = \frac{A_c}{g}[\frac{\kappa * V_{hub}}{ln(z_{hub}}]^2$. g is the gravitational acceleration, κ is von Karman's constant (0.4) and A_c is Charnock's constant (0.011). I_{ref} is the typical turbulence intensity at 15 m/s, where 0.12, corresponding to turbulence class C, is used.

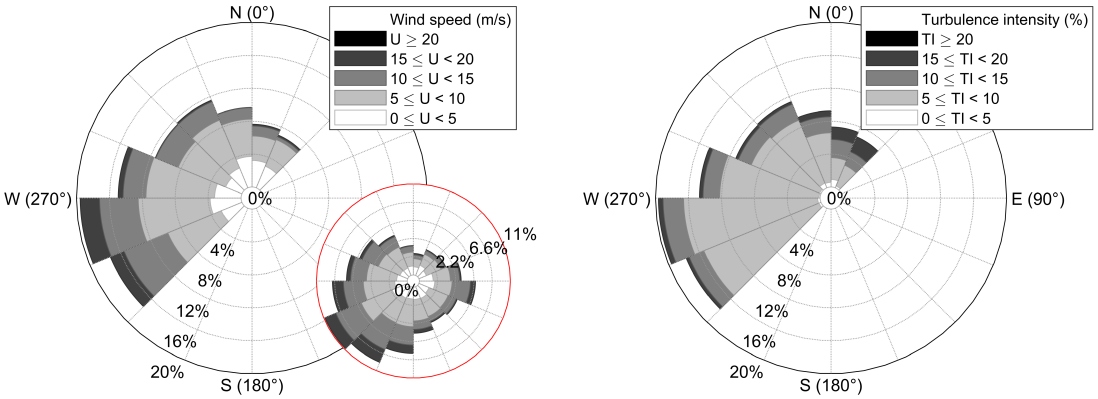

Figure 8. Wind direction dependent probability density functions of 30-minutes mean incoming wind speed (left (cup anemometer data in small circle)) and turbulence intensity based on 20 Hz data over 30 minutes (right) at 80 m elevation after the processing procedure

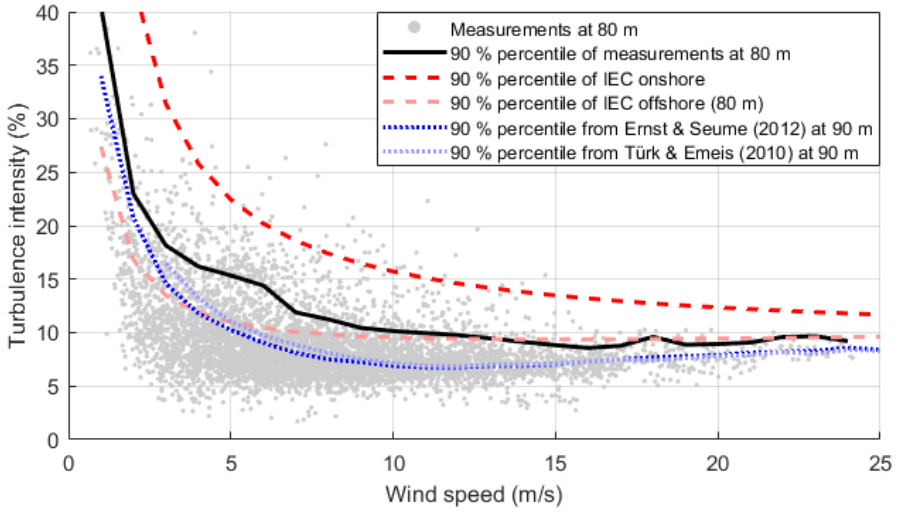

Figure 9. Turbulence intensity (based on 20 Hz data over 30 minutes) as function of wind speed of the processed FINO1-data set of this study, in addition to the 90th percentile of the the same data set, the IEC standards and literature findings of FINO1

5. Discussion

As mentioned in the results, a large data set of great variety in offshore wind conditions is ready for wind turbine research. However, there are also some limitations to the processed data set, which are already briefly mentioned. Even though the data set contains more than 6 000 30-minute periods of different situations, the data availability of the period is below 30 %. As shown by comparing to the cup anemometer data in Figure 7, the sonic anemometer data are more skewed towards lower wind speeds, leaving very few 30-minute periods of wind speeds higher than 20 m/s. This can be partly explained by the higher position of the cup anemometer, but it may also originate from the processing procedure resulting in the sonic data set with bias towards situations without precipitation. Last, the measurements are gathered over one year and three months, with the summer season represented twice. The statistics presented in Figures 7-9 may thus not be considered as statistics for one year at FINO1, but rather as an overview

of available situations for further use.

The processing procedure is thorough and quite conservative, with only 27 % availability after processing. Therefore, we expect that the risk of excluding data with valid measurements is larger than including data with errors in the processed data set. However, it was shown in Figure 5 that something very similar to an unrealistic spike was not detected. If only a small share of the data set, e.g. a few typical atmospheric conditions are analyzed further, and the quality requirements of these time series are very high, we therefore recommend a manual inspection prior to use, as illustrated in Figure 5.

Accepting the mentioned limitations, there are still a lot of applications for the processed data set. Figure 7-9 show the mean wind speed, the mean turbulence intensity and the mean direction, which are easily extracted. One may furthermore use the data set to calculate the roughness length of the ocean, the temperature fluxes and the Monin Obukhov length or Richardson number in order to classify the atmospheric stability. Taken into account the limited availability, it is probably more interesting to study some specific situations than trying to provide statistics of the characteristics at FINO1. These situations may be used to evaluate the co-variance between the mentioned parameters. In the frequency domain, further work could evaluate the energy spectra of the wind and temperature time series and study the coherence between the different heights in certain situations. The data set may also be used more directly towards the dynamic response of offshore wind turbines, e.g. by comparing to the standard turbulence models, wind profiles or turbulence intensities with height [34, 35].

6. Conclusions

In this study, existing quality control routines for sonic anemometer data are discussed and evaluated in terms of their relevance for offshore wind turbines. It is found that standard procedures, typically used in boundary layer meteorology, are not sufficient, and a tailored procedure for offshore wind turbine applications is necessary. This study presents a thorough processing procedure, applicable to sonic anemometer measurements for corresponding applications. The processing includes an assessment of data gaps due to instrument downtime, removal of corrupted raw data files, spikes, periods with precipitation or other disturbances and periods with wind speeds originating from the geographical exclusion zone, in addition to a coordinate transformation of the wind speeds to the mean flow direction and finally, organizing the data in 30-minute periods.

The processing procedure is applied to sonic anemometer measurements, routinely collected at the German offshore research platform FINO1, during the years 2015 and 2016. The applied conservative filtering reduces the raw data set of more than 20 000 30-minute periods to around 6 000. The remaining data set covers, however, a great variety of offshore conditions, despite the considerable reduction to 27 % of the original period.

The study is limited to the processing procedure and the presentation of the processed data set. For further use of the data, it is recommended to evaluate if additional stationarity tests, introduced in Section 3.8, should be applied. The proper choice of a corresponding method will always depend on the scientific question to address.

The data set will allow a wide range of applications related to offshore wind turbine design and analysis. Further studies will compare the characteristics of selected situations to standard turbulence models, and use these situations for studies of the dynamic response of offshore wind turbines.

Acknowledgments

The authors would like to thank DEWI (Deutsches Windenergi Institut), and especially Richard Fruehmann, for providing the FINO1 high-resolution sonic anemometer data, and BSH (The Federal Maritime and Hydrographic Agency of Germany) for providing slow meteorological

reference data. We would also like to thank Martin Flügge and Etienne Cheynet for input and support.

References

[1] Barthelmie R J, Crippa P, Wang H, Smith C M, Krishnamurthy R, Choukulkar A, Calhoun R, Valyou D, Marzocca P, Matthiesen D, Brown G and Pryor S C 2014 *Bull. Amer. Meteor. Soc.* **95** 743–756
[2] Emeis S 2014 *Meteorol. Appl.* **21** 803–819
[3] Cheynet E, Jakobsen J B and Reuder J 2018 *Bound.-Layer Meteorol.* **169** 429–60
[4] Sathe A, Mann J, Gottschall J and Courtney M S 2011 *J. Atmos. Ocean. Tech.* **28** 853–868
[5] Foken T and Wichura B 1996 *Agric. For. Meteorol.* **78** 83–105
[6] Vickers D and Mahrt L 1997 *J. Atmos. Ocean. Tech.* **14** 512–26
[7] Foken T, Leuning R, Oncley S R, Mauder M and Aubinet M 2012 Corrections and data quality control *Eddy Covariance* (Dordrecht: Springer) pp 85–131
[8] Foken T, Göckede M, Mauder M, Mahrt L, Amiro B and Munger W 2005 Post-field data quality control *Handboook of Micrometeorology* (Dordrecht: Springer) pp 181–208
[9] Cheynet E 2019 *Manuscript submitted to Lecture Notes in Civil Engineering*
[10] Mücke T, Harkness C and Argyriadis K 2012 *EWEA Copenhagen* **2102** 1–10
[11] Eliassen L and Obhrai C 2016 *Energy Procedia* **94** 388–98
[12] Türk M and Emeis S 2010 *J. Wind. Eng. Ind. Aerodyn.* **98** 466–71
[13] Ernst B and Seume J R 2012 *Energies* **5** 3835–55
[14] FuE-Zentrum FH Kiel GmbH 2019 FINO1: Forschungsplattformen in Nord- und Ostsee Nr. 1 URL `https://www.fino1.de/`
[15] Neumann T, Kolopp K, Strack M, Mellinghoff H, Söker H, Mittelstaedt E, Gerasch W J and Fischer G 2003 *DEWI Mag* 32–46
[16] 4C Offshore Ltd 2019 Global Offshore Renewable Map - 4C Offshore URL `https://www.4coffshore.com/offshorewind/`
[17] Forschungs- und Entwicklungszentrum Fachhochschule Kiel GmbH 2019 Fotos von FINO1 URL `https://www.fino1.de/de/medien/fotos.html`
[18] Neumann T and Nolopp K 2007 *DEWI Mag* **30** 42–46
[19] Bakhoday-Paskyabi M, Fer I and Reuder J 2018 *Ocean Dynam.* **68** 109–130
[20] Højstrup J 1993 *Meas. Sci. Technol.* **4** 153–57
[21] Schmid H P, Grimmond C S B, Cropley F, Offerle B and Su H B 2000 *Agric. For. Meteorol.* **103** 357–74
[22] Papale D, Reichstein M, Aubinet M, Canfora E, Bernhofer C, Kutsch W, Longdoz B, Rambal S, Valentini R, Vesala T and Yakir D 2006 *Biogeosciences* **3** 571–83
[23] Sachs L 1996 *Angewandte Statistik: Anwendung Statistischer Methoden* (Berlin: Springer)
[24] Rinker J M 2016 *An Empirically Based Stochastic Turbulence Simulator with Temporal Coherence for Wind Energy Applications* Ph.D. thesis Duke University
[25] Zhang R, Huang J, Wang X, Zhang J A and Huang F 2016 *J. Ocean Univ. China* **15** 389–98
[26] Emeis S 2010 *Measurement Methods in Atmospheric Sciences - In Situ and Remote* (Stuttgart: Borntraeger Science Publishers)
[27] Westerhellweg A, Neumann T and Riedel V 2012 *DEWI Mag* **40** 60–66
[28] Wilczak J M, Oncley S P and Stage S A 2001 *Bound.-Layer Meteorol.* **99** 127–50
[29] Mahrt L 1998 *J. Atmos. Ocean. Tech.* **15** 416–29
[30] Cava D, Donateo A and Contini D 2014 *Agric. For. Meteorol.* **194** 88–103
[31] Obhrai C, Kalvig S and Gudmestad O T 2012 *The 22nd Int. Offshore Polar Eng. Conf.* 440–50
[32] Bendat J S and Piersol A G 2011 *Random Data: Analysis and Measurement Procedures* Wiley Series in Probabillity and Statistics (Hoboken: John Wiley & Sons)
[33] Chen J, Hui M C and Xu Y L 2007 *Bound.-Layer Meteorol.* **122** 105–21
[34] International Electrotechnical Commision 2005 *IEC 61400-1 Wind Turbines - Part 1: Design requirements*
[35] International Electrotechnical Commision 2009 *IEC 61400-3 Wind turbines - Part 3: Design requirements for offshore wind turbines*