Article

Considering Uncertainties of Key Performance Indicators in Wind Turbine Operation

Sebastian Pfaffel * , Stefan Faulstich and Kurt Rohrig

Fraunhofer Institute for Energy Economics and Energy System Technology—IEE, Königstor 59, 34119 Kassel, Germany; Stefan.Faulstich@iee.fraunhofer.de (S.F.); Kurt.Rohrig@iee.fraunhofer.de (K.R.)

* Correspondence: Sebastian.Pfaffel@iee.fraunhofer.de; Tel.: +49-561-7294-441

Received: 3 December 2019; Accepted: 21 January 2020; Published: 30 January 2020

Featured Application: The results enable wind turbine operational managers to consider data handling-related uncertainties when using KPIs. The following recommendations allow mitigating uncertainties through simple measures. Furthermore, methods are proposed to correct systematic deviations.

Abstract: Key performance indicators (KPIs) are commonly used in the wind industry to support decision-making and to prioritize the work throughout a wind turbine portfolio. Still, there is little knowledge of the uncertainties of KPIs. This article intends to shed some light on the uncertainty and reliability of KPIs in general and performance KPIs in particular. For this purpose, different uncertainty causes are discussed, and three data handling related uncertainty causes are analyzed in detail for five KPIs. A local sensitivity analysis is followed by a more detailed analysis of the related uncertainties. The work bases on different sets of operational data, which are manipulated in a large number of experiments to carry out an empirical uncertainty analysis. The results show that changes in the data resolution, data availability, as well as missing inputs, can cause considerable uncertainties. These uncertainties can be reduced or even mitigated by simple measures in many cases. This article provides a comprehensive list of statements and recommendations to estimate the relevance of data handling related KPI uncertainties in the day-to-day work as well as approaches to correct KPIs for systematic deviations and simple steps to avoid pitfalls.

Keywords: wind turbines; key performance indicators; performance; operation; uncertainties; sensitivities; KPI

1. Introduction

After the successful installation of a wind farm (WF), the operational performance of its wind turbines (WTs) determines whether a project is economically successful or not. On average, operation and maintenance (O&M) accounts for approximately 25% to 35% [1–4] of the levelized cost of energy (LCOE) and also affects the energetic and monetary revenue. Thus, the performance of WTs is of high importance and closely monitored by operational managers. As part of the digitalization of energy supply, the latest information technology (IT) applications and solutions also penetrate the area of O&M of WTs and WFs. Big data and artificial intelligence (AI) applications in particular support the operational managers with valuable information regarding the WT performance. As more and more tasks and evaluations are automated, reliable and trustworthy results become more important.

KPIs are a common tool throughout all industries to make information easily accessible and comparable [5]. The wind industry is no exception and makes use of KPIs to evaluate the performance and reliability of WTs regularly. A whole set of KPIs is used to evaluate different aspects of WT operation, as shown in [6,7]. Even though further tools like sophisticated machine

learning approaches for condition monitoring systems (CMS) [8] and performance monitoring applications [9] are used more and more often in the industry; KPIs remain the foundation for asset management, reporting, and benchmarking. As KPIs and their trends are part of many decision-making processes, it is necessary to take their reliability into account [5].

The reliability of a KPI can also be referred to as its uncertainty. Even though the uncertainties of KPIs have not been widely discussed in the literature, a few publications are available. Sánchez-Márquez et al. [10] evaluate the uncertainties of KPIs used for balanced scorecards caused by the selected sample size. Perotto et al. [11] discuss the uncertainty of environmental performance indicators on wastewater discharges, and Torregrossa et al. [12] deal with the uncertainty of energy consumption indicators for wastewater treatment plants due to measurement uncertainties and data availability limitations. Another example is the consideration of uncertainties by Feiz et al. [13] for a CO₂ emission indicator used in the cement industry.

Publications on the sensitivity and uncertainty of KPIs related to the wind industry are rare. Martin et al. [14] perform an analysis of the sensitivity of O&M costs on the chosen maintenance strategy; although this approach aims to identify cost reduction factors and not to identify and quantify uncertainties. The work of Dykes et al. [15] follows a similar approach and performs a sensitivity analysis of the wind plant performance to changes in WT design parameters. The only known publication on uncertainties of a performance KPI for WT operation was published by Craig et al. [16], which examines the uncertainty in power curves due to different methods for data filtering and power curve modeling. This work is continued in another publication [17] to quantify the uncertainties of wind plant energy estimation resulting from power curve uncertainties and different sampling periods.

A broader assessment of the uncertainties of O&M-KPIs used in the monitoring of WT is lacking. Thus, this paper discusses sensitivities and uncertainties of the most important performance KPIs and tries to quantify uncertainties to assess their relevance for the daily work. Another scope of this paper is to provide recommendations to reduce or mitigate KPI uncertainties by simple measures. The results are also meant to raise awareness of uncertainties when dealing with KPIs in general. The authors try to view uncertainties from the perspective of an operational manager and consider uncertainties that are relevant to and can be influenced by this specific role. By quantifying KPI uncertainties and providing mitigation strategies, this paper aims at more reliable and trustworthy KPIs as part of decision support tools.

2. Key Performance Indicators

In operational management of WTs, KPIs are used to evaluate and categorize the performance of WTs with regard to (1.) Predetermined target values, which are in many cases part of contractual agreements, (2.) comparable WTs (e.g., of the same WF), or (3.) the long-term trend of the WT itself. In any case, KPIs are always compared to a reference, and significant deviations usually trigger actions [18,19].

A high deviation to the reference reflects a high need for intervention and leads to a higher prioritization in the work of an operational manager. This usually means a more detailed review of the performance of the WT and the operational data to identify the underlying causes. At this point, one has to ask what deviation to the reference value can be considered to be healthy or can be explained due to the uncertainty of the KPI itself, and thus either should not trigger any action or lower its priority. To answer this question is one of the main scopes of the present work.

As a single KPI is usually not sufficient to rate the performance of a WT, in most cases, a set of different metrics is used. The selected KPIs are calculated on a regular base. Monthly, quarterly, half-yearly, and yearly periods are common resolutions for KPIs and are normally synchronized with reporting periods [18,19]. According to Gonzalez et al. [7] and Pfaffel et al. [6], KPIs for O&M of WTs belong to one of the following five different groups.

- Performance KPIs
- Maintenance KPIs

- Reliability KPIs
- Health, Safety and the Environment (HSE) KPIs
- Finance KPIs

Financial and HSE KPIs are not wind-specific, and are thus excluded from further evaluation. Furthermore, the calculation of HSE, reliability, and some maintenance KPIs is solely based on distinct events instead of measurements. As the idea to provide uncertainties for these KPIs is misleading, they are excluded as well. According to Hahn et al. [20], standardization of the collected data and a review of documentation processes provides much more utility in this case. To select a total of five KPIs for further assessment out of the remaining KPIs, the prioritization of KPIs, as provided by [6], is used. The selected KPIs are briefly described and defined in Sections 2.1–2.4 based on [6].

2.1. Wind Conditions

The wind conditions define the quality of a specific WT site and are not a KPI itself but described by several KPIs. The most known KPI is the average wind speed, see Equation (1). As it reflects the usable energy inadequately, the authors complemented it by the average wind power density for this publication, see Equation (2). The wind power density includes the cubic relation between wind speed and the usable energy, and also considers varying air densities. Both KPIs are calculated based on n observations in the observation period (e.g., day, month, year).

$$\bar{V} = \frac{1}{n} \cdot \sum_{i=1}^n V_i \quad (1)$$

where

$$\begin{aligned} \bar{V} &= \text{Average wind speed,} \\ V_i &= \text{Measured wind speed for the } i\text{th observation} \end{aligned}$$

$$\bar{P}_{\text{Wind}} = \frac{1}{n} \cdot \sum_{i=1}^n \frac{1}{2} \rho_i V_i^3 \quad (2)$$

where

$$\begin{aligned} \bar{P}_{\text{Wind}} &= \text{Average wind power density,} \\ \rho_i &= \text{Air density for the } i\text{th observation,} \\ V_i &= \text{Measured wind speed for the } i\text{th observation} \end{aligned}$$

2.2. Capacity Factor

The capacity factor normalizes the energy yield to the rated power of a WT, see Equation (3). It is a common metric to compare WTs of the same WF or review the long term trend of a specific WT.

$$CF = \frac{W_{\text{Actual}}}{W_{\text{Nominal}}} \cdot 100\% \quad (3)$$

where

$$\begin{aligned} CF &= \text{Capacity factor} \\ W_{\text{Actual}} &= \text{Actual energy yield of the observation period,} \\ W_{\text{Nominal}} &= \text{Nominal possible energy yield of the observation period} \end{aligned}$$

2.3. Power Curve

The power curve is not a KPI itself but an essential tool in WT monitoring, and is also necessary to calculate further KPIs. It represents the relation between wind speed and power output of a WT. Exemplary use cases for power curves during the operational period of a WT are a basic monitoring

of the WT performance through a visualization of the operating data or the calculation of energy losses during downtimes.

2.4. Production-Based Availability

The production-based availability is the ratio between the actual energy yield and the potential energy yield of a WT if downtimes would not occur (actual yield plus energy losses), see Equation (4). Compared to the time-based availability, downtimes are weighted by their energy losses. The corresponding energy losses can be determined using several methods, inter alia by using the power curve and wind speed measurements.

$$A_W = \left(1 - \frac{W_{\text{Potential}} - W_{\text{Actual}}}{W_{\text{Potential}}} \right) \cdot 100\% \quad (4)$$

where

$$\begin{aligned} A_W &= \text{Production-based availability,} \\ W_{\text{Actual}} &= \text{Actual energy yield of the observation period,} \\ W_{\text{Potential}} &= \text{Potential energy yield of the observation period} \end{aligned}$$

3. Uncertainties

“Uncertainty” is a term and concept used in various ways. Lindley [21] describes it generically to be a situation where one does not know whether a statement or information is true or false. According to the Guide to the Expression of Uncertainty in Measurement (GUM) [22], it “means doubt about the validity of the result of a measurement”. This definition, in the context of measurements and metrics that also includes KPIs, corresponds to the idea of this work.

Uncertainties are closely related to sensitivities. Sensitivity analyses evaluate the effect of single parameters or a group of parameters on a result and are, for example, also used when assessing different scenarios. In the context of uncertainties, a sensitivity analysis helps to understand whether and how substantial deviations of one parameter affect the measurement or KPI and is performed as the first step of an uncertainty analysis. Sensitivity analyses can be distinguished into local and global analyses. A local sensitivity analysis evaluates the impact of deviations of an isolated parameter while keeping the other inputs stable, whereas a global sensitivity analysis varies all input parameters at the same time so that interdependencies can be discovered. Screening methods combine both attempts [22,23].

Uncertainties are either aleatoric or epistemic. Aleatoric uncertainties are caused by the natural randomness of processes and are thus not predictable. Epistemic uncertainties are the result of a lack of information or knowledge and can be avoided or corrected in theory. [24] The standard deviation (σ) is a common metric to quantify uncertainties and assumes that the related results follow a normal distribution [22]. Density functions of single uncertainties can be combined to overall uncertainties through analytical or numerical methods. For this purpose, it is vital to know whether single uncertainties are independent or not [25–28].

Utilization of uncertainties in the wind industry includes the determination of wind potential and energy yields [29], the determination of power curves [30,31], or wind power forecasts [32,33]. The results in corresponding expert reports, measurements, or predictions are displayed with uncertainty information to enable the consideration of individual safety requirements in decision-making processes. The only known publication related to uncertainties of KPIs in the wind industry is the already mentioned work of Craig et al. [16] on the uncertainty of power curves.

Figure 1 presents a fishbone diagram of possible causes of uncertainties collected for this work. The single causes haven been assigned to one of the four groups: data handling, measurement, method, and definition. As data handling-related uncertainties are the only uncertainties that are, to a certain

extent, in control of the operational manager, this work focuses on this group. The other groups are briefly discussed below.

Measurement uncertainties are the apparent causes of uncertainties for KPIs and well researched in general. Relevant references for uncertainties of inputs like power, wind speed, wind direction, and temperature are the International Electrotechnical Commission (IEC) standards 61400-12-1/2 [30,31] (tables G1–G8) and 61400-15 [34], the technical guideline six [29] of the Fördergesellschaft Windenergie und andere Dezentrale Energien (FGW e.V.), and the recommendations of the Measuring Network of Wind Energy Institutes (MEASNET) [35]. Default or standard values for individual uncertainties of the measurements are also given in the sources mentioned.

Figure 1. Exemplary causes for uncertainties of performance KPIs grouped by their categories in a causal diagram.

Uncertainties resulting from corrections to the measured values are assigned to the uncertainty group “method”. Among other factors, the wind speed measurement of the nacelle anemometer is corrected by a transfer function to account for the slowed down and disturbed wind speed at the measuring position behind the rotor. When determining a power curve, the wind speed is also corrected to reference values for turbulence and air density, for example. Further details on the methods and the associated uncertainties can be found in the IEC standard 61400-12-1/2 [30,31].

Further uncertainties can arise when KPI definitions are changed, different definitions are used or from an inaccurate formulation of the definition. As KPIs are usually calculated and compared in a closed operational management software or benchmark platform, consistent definitions can be assumed for the entire portfolio and all periods. However, this uncertainty group should be considered whenever precalculated KPIs from different sources or systems are compared. The IEC standard 61400-26-1/2 [36,37] already provides four exemplary definitions of the time-based availability and three definitions of the production-based availability, let alone many more company-specific versions.

Data handling related uncertainties originate from the way how data are collected, processed, and stored. Some of the listed uncertainty causes can be easily avoided by simple measures or proper documentation of the data, and are thus not further investigated. Examples are missing information on the time zone of timestamps, a missing synchronization, or missing information on the time allocation (start, center, end) of an aggregation period. Uncertainties due to different data filtering approaches have been investigated by Craig et al. [16] for power curves. This work focuses on the remaining three uncertainty causes.

- **Data resolution :** Operational data of WT are mostly stored in resolutions of 5, 10, or 15 min. Older projects also use resolutions as low as 30 min, whereas recent projects sometimes store Hz data.

When data is aggregated, information like extreme values gets lost in the averaging process. This behavior is not to be confused with the Nyquist–Shannon sampling theorem, which determines the necessary sampling rate in a measurement [38,39]. The IEC 61400-12 [30,31] requires for example, a minimum sampling rate of 1 Hz for most measurements—ambient temperature or air pressure can also be sampled once per minute—whereas it defines aggregated 10 min data to be the foundation for power curve calculations. This work expects the underlying measurement to be done correctly. An additional uncertainty cause is related to the sample length of an aggregation period. Whenever samples are missing in the aggregation period, the remaining samples are weighted more strongly than in other periods. Although missing samples might show up in Hz data, they might not be detectable in 5 min data anymore. Schmiedel [40] investigated the information loss when switching from 1 min data to 15 min data in a different context.

- **Data gaps/data availability:** Data gaps in a long history of operational data are normal. As shown in [41], the data availability is, on average, at 95%. In most cases, the data availability is higher (median at 99%), but in other cases also much lower. Current data availability requirements seem somewhat arbitrary. The German guideline for site assessment [29] requires, for example, a data availability of 80% without any further notices or restrictions. The technical guideline 10 of the FGW e.V. [42] allocates data gaps to be downtime in its availability calculation. Reasons or data gaps are manifold and include failed sensors, downtime of the whole turbine, including its controller, as well as telecommunication interruptions [43]. Many WT still use the GSM standard of the cellular network. If data are missing, the remaining data points are unlikely to be representative of the period considered. Thus, uncertainties in the calculated KPIs are to be expected. Existing literature mainly deals with different approaches to fill data gaps [38,44,45] and not with their effect on KPIs.
- **Data completeness:** Especially for older WT, the number of available measurements is often very limited, and supporting inputs for the calculation of KPIs are missing. Similar situations occur when sensors fail for a longer time. If the missing measurements are not vital for the calculation but increase the accuracy, they can be neglected or sometimes replaced by approximations. This approach leads in any case to uncertainties. The present paper discusses this issue for the example of air density since it is a vital factor in obtaining comparable results in power performance assessments [30,31]. Especially data sets of older WTs do not comprise all required measurements (air pressure, temperature, and humidity) for an accurate density calculation, which makes our example a common issue in the industry.

4. Method and Data

This work follows the recommendations of the GUM [28] to evaluate different uncertainty causes for various KPIs. In the first step, a local sensitivity analysis determines whether a KPI is affected by a potential uncertainty cause or not. A more detailed analysis is carried out for sensitive relationships to quantify the uncertainties. Herefore, the GUM method “Type A” for the evaluation of standard uncertainties is applied. This approach determines measurement uncertainties through many repeated measurements. The empirical distribution function of all measurements allows calculating the expected value of the measurement as well as the deviations of the single measurements to the expected value. If the distribution function follows a normal distribution, which is commonly assumed for uncertainties, the uncertainty can be quantified by the standard deviation, see Equation (5) [28].

Transferring this method to the present work means calculating KPIs in many experiments either through multiple manipulations of a selected dataset or by investigating many different datasets to determine the expected uncertainty for different scenarios. The described method is intended to quantify the uncertainties due to random (aleatoric) deviations. As the considered uncertainty causes also lead to systematic (epistemic) deviations, as shown later on in the paper, this work also tries to correct or mitigate those systematic deviations.

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2}$$

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i \quad c_v = \frac{\sigma}{\mu} \quad (5)$$

where

σ = standard deviation over n measurements/experiments,

x_i = result of the i th measurement/experiment,

μ = average over n measurements/experiments,

c_v = coefficient of variation

Figure 2 illustrates the overall method of the present paper to determine sensitivities and uncertainties for selected KPIs (Section 2) and uncertainty causes, see Sections 4.2–4.4 for details. Different data sets (Section 4.1) were used depending on the evaluated uncertainty cause. The approaches for the different uncertainty causes were similar. The original data was used to calculate reference values for the single KPIs. To obtain sufficient data for uncertainty analyses, the original data was manipulated in numerous experiments. In the last step, uncertainties were derived by comparing the results of the experiments to the reference values of the single KPIs.

Figure 2. Approach of this paper to analyze sensitivities and uncertainties for selected KPIs and uncertainty causes.

4.1. Data and Toolset

The present work makes use of several different data sets. To be able to work with a complete and high-quality dataset, comprising all relevant measurements (pressure, temperature, humidity, etc.), the data of a 200 m high wind met mast was analyzed. The met mast is located at a complex, forested site at “Rödeser Berg” close to Kassel in central Germany. The availability of comprehensive meteorological data allows investigating the relevance of complete measurement data for the accuracy of the different KPIs. A detailed description of the location, as well as the instrumentation of the met mast, can be found in a publication by Klaas et al. [46]. Furthermore, data of the Wind Energy Information Data Pool (WInD-Pool) [41] was available where needed. The WInD-Pool includes data in the Hz resolution

for a German offshore WF (5 MW WT) in the North Sea, which was needed for this work to evaluate the effect of different resolutions. Further details cannot be stated due to confidentiality reasons. Wherever possible, however, the data from ENGIEs open data WF “La Haute Borne” [47] was used to ensure the replicability of the results. The WF comprises four Senvion MM82 WT with a rated power of 2 MW each and is located in the department Meuse in northeast France. For each WT, operational data over approximately six years was available. All datasets were filtered for implausible values and stuck signals [48], for example, the plausible wind speed range is 0 m/s to 50 m/s.

For the handling, filtering, and manipulation of the data as well as the calculation of KPIs, the Python Open-Source Tool for Wind Farm Operational Performance Analysis (OpenOA) [49,50] was used. OpenOA is developed by the National Renewable Energy Laboratory (NREL) and extended according to the requirements of this work. OpenOA is available under an open source license as Python source code on Github [51]. Due to the public availability of the ENGIE Open-Data-Dataset and OpenOA, the results can be replicated.

4.2. Data Resolution

The effect of different data resolutions was assessed based on 1 Hz data from the WInD-Pool for a 5 MW offshore WT. The dataset comprises of 4,727,951 entries, or 55 days worth, of high-resolution operating data. Data of additional WT was available and used to validate the results. After filtering and merging, 98.5% of the data proved to be good quality, and data gaps were filled by linear interpolation to obtain a continuous and complete time series. This was necessary to investigate the effects of resolution changes without being affected by data gaps and resulting differently weighted data points.

The 1 Hz data was then resampled to different resolutions up to hourly values. Special requirements were also considered, such as for vector functions to average angular values (e.g., wind direction) [30,52]. This averaging process used the 1 Hz data for each resolution so that sound statistics (min, max, std.) were available. All considered KPIs were then calculated for each resolution to assess sensitivities and uncertainties.

4.3. Data Availability

The relevance of varying data availabilities on different KPIs was assessed based on 10 min data of the WT “R80711” of the ENGIE Open Data WF for the year 2017. Approximately 2% of the data was categorized as implausible values or stuck signals and thus removed. The remaining data serve as the 100% data availability reference for this analysis. In this instance, filling the data gaps was not necessary and would rather falsify the results. In the next step, data gaps were added to the time series to reduce the data availability. This was done in one percent increments for the data availabilities ranging from 100% down to 5%. Two different methods were used to evaluate not only the relevance of the data gaps but also the effect of their distribution in the time series. In the first method, single data gaps were randomly added until the target data availability was reached to achieve a uniform distribution; longer gaps occurred only by chance. The second method added one long subsequent data gap of the required length to reach the targeted data availability by using random starting points.

As every experiment was expected to provide different results, 1000 experiments were carried out per targeted data availability and method to get reliable results. The dataset was split into single months as most KPIs are calculated monthly. Thus, a total of about 2.2 million experiments were carried out. As a single month of data is not sufficient to calculate power curves, they were calculated based on the entire year using the IEC and spline approaches implemented in OpenOA, which totaled in approximately 285,000 power curves.

4.4. Data Completeness

To analyze the effect of missing air density measurements as well as their approximation on the calculation of KPIs, one year of 10 min data of the 200 m met mast was used. This dataset

was selected due to its comprehensive, reliable, and well-calibrated measurements. After filtering the data availability of the selected year was 99%. In addition, three years (99.9% data availability) of comprehensive data of a North Sea offshore WF was used, as the met mast does not include power measurements as well as to validate the results.

In the first step, the air density and wind power density sensitivities were evaluated to the availability of temperature, pressure, and humidity measurements using Equation (6) [31], as well as to the site altitude above sea level. Furthermore, the air density was approximated by adding additional inputs step-by-step to assess their relevance starting with the standard atmosphere. All results were compared to a reference case comprised of all required inputs.

$$\rho = \frac{1}{T} \left(\frac{B}{R_0} - \phi P_w \left(\frac{1}{R_0} - \frac{1}{R_w} \right) \right) \quad (6)$$

where

- ρ = derived air density,
- T = measured absolute air temperature,
- B = measured air pressure,
- R_0 = gas constant of dry air 287.05 J/kgK,
- ϕ = relative humidity (range 0 to 1),
- R_w = gas constant of water vapor 461.5 J/kgK,
- P_w = vapor pressure for T

5. Results

This section reviews selected results of the sensitivity and uncertainty analysis in detail. For each uncertainty cause, one KPI is presented based on the relevance of the results. Other results are only briefly mentioned for the sake of brevity. Table 1 provides a summary of all reviewed uncertainty causes and KPIs as part of the discussion (Section 6).

5.1. Data Resolution

For KPIs with a linear dependence, changes of the data resolution showed hardly any impact. The smoothing effect of the averaging process when resampling the data did not affect KPIs, such as the average wind speed or capacity factor. Minor deviations were caused by missing data points, which led to a different weighting of the remaining data points. Different aggregation time periods led to changing allocations and weightings of the data and thus to slight deviations. The observed deviations can be considered negligible given the number of samples per aggregation period do not deviate too much.

Equations of the second- or a higher-degree were affected by the smoothing effect. Missing information on extreme values led to lower KPI results since losing higher values has a greater impact than losing lower values. For the wind power density, one can correct this systematic deviation if the standard deviation of the wind speed is available, see Figure 3a. The correction assumes a normally distributed wind speed within the aggregation period as used to describe wind speed turbulences [31]. Of note, this is only valid for short periods. Long term wind speed distributions are generally characterized by Weibull or Rayleigh distributions. The correction algorithm showed an uncertainty of ~0.1% for the investigated examples since this assumption is not always met. This is still fifteen times lower than the observed systematic deviation when switching from 1 Hz data to hourly data.

The smoothing effect also affected the calculation of empirical power curves. Systematic deviations were witnessed in the transition area to the rated power, where high resolutions show a kink while lower resolutions smooth the transition and give the power curve its expected look. The IEC standard 61400-12-1/2 [30,31] requires 10 min data for empirical power curves. Different data resolutions led to deviations in the power curves and thus uncertainties when calculating energy yields. A relative

standard deviation of approximately 0.4% was observed when calculating potential annual energy productions using power curves based on different resolutions over various wind speed distributions. Although this result should not be considered to be generally valid, its order of magnitude is expected to be correct.

Figure 3. Effect of different data resolutions on (a) (wind power density) the calculation of an average wind power density and (b) (availability) availability metrics in the case that downtimes are solely identified based on SCADA-Data.

This uncertainty also propagates into the calculation of the potential energy yield when assessing the production-based availability, but can be neglected due to its low impact. Much more important is the effect shown in Figure 3b. Initiatives such as the WInD-Pool, derive downtimes from wind speed and power measurements because distinct event information is often times missing. Although high-resolution data allows a clear distinction between operating times and standstill, this distinction becomes blurred at lower resolutions. The high-resolution area of Figure 3b shows a dropping in availability. This behavior can be explained by the assimilation of time steps adjacent to downtime because the negative power of several non-operational time steps can negate the positive power of single time steps. As the resolution further decreases, this behavior reverses. Short downtimes, such a five min-long reset of a WT, disappear in 10-min data. Of course, the behavior depicted in Figure 3b strongly depends on the distribution and length of downtimes, but the overall message is clear: Deriving downtime events from operational data is afflicted with significant uncertainties. Instead, distinct event information should be used whenever possible.

5.2. Data Availability

All considered KPIs are affected by decreasing data availabilities since the data is less representative of the considered period. Figure 4 shows the coefficient of variation in the results for (a) the wind speed, (b) the wind power density, and (c) the capacity factor, depending on the data availability and compared to a reference value based on a 100% data availability. The coefficient of variation was chosen over the standard deviation to make different datasets comparable.

The results illustrate that data availability is not the decisive factor, but rather the representativeness of the remaining data. Although the data availability is equivalent, short and random data gaps show only a minuscule effect, but long and subsequent data gaps lead to high uncertainties. Hereby, the results also emphasize the need of an entire year of measurements if representative results are to be achieved. The two depicted data gap distributions represent the best- and worst-case scenarios. For mixed experiments of shorter and longer data gaps, the coefficient of variation is expected to be between both cases.

A metric for the quality or representativeness of the remaining time series is needed to estimate case-specific uncertainties for data gap distributions witnessed in the field. Different attempts of the authors as well as existing metrics, such as the work of Bertino [53], did not lead to satisfying results or were not applicable for the described case. Thus, there is need for further research.

As for the data resolution (Section 5.1), KPIs with a nonlinear relation to the wind speed measurement show a higher sensitivity to decreasing data availability. When comparing the results for the wind speed (Figure 4a) to the wind power density (Figure 4b) and the capacity factor (Figure 4c), the coefficient of variation is more than twice as high.

Figure 4. Coefficients of variation for different KPIs caused by a reduction in data availability. Randomly distributed data gaps are compared to continuous data gaps. (a) wind speed, (b) wind power density, (c) capacity factor.

Even though the power curve of a WT represents the relation between two measurements (wind speed and power) and should not be affected by data gaps, the experiments carried out in this work showed increasing uncertainties with decreasing data availabilities, see Figure 5. Still, the uncertainties are much lower compared to the examples in Figure 4. If the data gaps are randomly distributed, the uncertainty of the resulting annual energy production is even for a data availability of just 20% below 0.2% and thus negligible. Subsequent data gaps cause not only considerably higher aleatoric uncertainties but also systematic deviations. These uncertainties can be mainly related to seasonal effects, such as varying air densities and turbulence intensities. By correcting the power curves for both effects [31,54], the uncertainties can be significantly reduced. At a data availability of 80%, the uncertainty for subsequent data gaps can be reduced from approximately 0.6% to 0.3%. However, some systematic uncertainties remain, and long data gaps should be avoided whenever possible. Even though the uncertainties caused by random data gaps could be only slightly decreased, it is highly recommended to perform a density and turbulence correction of power curves in any case. Figure 4 also highlights the need for a comprehensive data set for power curve assessments. Less than three months of data (here 25% data availability) cause an increasing uncertainty as the wind speed range is not sufficiently covered in many cases. Again, the order of magnitude of the above values is considered valid, while the concrete uncertainties can vary for other datasets. If an empirical power curve is used to calculate the production-based availability, its uncertainty propagates into the yield loss calculation.

Figure 5. Effect of decreasing data availability on the uncertainty of power curves. The uncertainty is represented by deviations in the calculation of a generic annual energy yield using a rayleigh wind speed distribution (6.45 m/s) as reference.

5.3. Data Completeness

The importance of complete inputs for the calculation of KPIs is discussed in this section using the example of an air density measurement. Figure 6 shows the impact of the missing air density measurement and the quality of different approximations on the calculation of wind power density. If no air density measurement is available, it is common to use the standard atmosphere of 1.225 kg/m^3 . If the considered site or nacelle is at a higher altitude, this assumption leads to systematic high results. A correction for altitude likely results in an underestimation since the average temperature is usually below the standard temperature of 15°C . If the air density is also corrected for the average temperature, a yearly wind power density KPI is already close to the reference as lower and higher temperatures balance out. Good results and drastically reduced deviations are achieved when temperature measurements are available.

Figure 6. Monthly sensitivity of the mean wind power density to missing information.

If the small and systematic underestimation of the air density is ignored, the remaining deviations can be explained by weather-related pressure fluctuations. Thus, the uncertainty of an approximated air density can be derived from a known and site-specific distribution of the air pressure. As the sensitivity of the air density to changes of the humidity is low, the recommendation of the IEC standard 61400-12-1/2 [30,31] to use a default humidity of 50% is reasonable.

The presented results are validated against the North Sea WF data. The higher weather-related standard deviation of air pressure at the offshore location led to a higher standard deviation of the air density and thus of the wind power density. The discussed uncertainty propagates through the power curve to the production-based availability.

6. Discussion

This paper evaluates the sensitivity and uncertainty of five different KPIs to three different uncertainty causes related to data handling. Section 5 presents the most relevant findings of this work. A summary of the effect of each uncertainty cause on each KPI is provided in Table 1.

Table 1. The effect of the evaluated uncertainty causes on different KPIs.

Uncertainty Causes	Key Performance Indicators				
	Wind Speed	Capacity Factor	Wind Power Density	Power Curve	Production-Based Availability
Data Resolution	-	-	!	!	!
Data Availability	!	!	!	!	↗
Data Completeness	-	-	!	↗	↗

- not affected; ! affected; ↗ uncertainty propagation.

Furthermore, the following list provides a summary of key statements and recommendations for the discussed uncertainty causes and KPIs. It is intended to make the reader aware of potential pitfalls and sources of error and uncertainty in the day-to-day work with KPIs.

Data Resolution

- The mean wind speed and capacity factor are not affected by the data resolution.
- The systematic influence on the mean wind power density is caused by its cubic equation. By using the standard deviation of the wind speed, it can be corrected to an uncertainty of about 0.1%.
- By filtering the data thoroughly, effects on the power curve can be reduced to an acceptable level. The standard deviation of the resulting energy yields can be estimated at 0.4%.
- Whenever downtime events are derived from operational data, the number and length of downtime events heavily depend on the chosen data resolution. Availability metrics are affected accordingly. Thus, discrete event information should be used whenever possible. The production-based availability is furthermore affected through the power curve.

Data Availability

- The effect of decreasing data availability mainly depends on the distribution of data gaps in the whole dataset. While keeping the data availability constant, continuous data gaps cause much higher deviations than short and randomly distributed gaps. This emphasizes the need to use a full year of measurements if representative results are to be achieved. If data gaps are entirely random, the deviations can be almost neglected, and low data availabilities provide valid results.
- All discussed KPIs are affected by decreasing data availability. KPIs with a cubic equation, like the capacity factor or wind power density, show a higher sensitivity than the mean wind speed, which has a linear dependence. Site-specific simulations are necessary to quantify uncertainties.
- Short and random data gaps are of low importance when calculating the power curve. In contrast, long and continuous data gaps cause significant and systematic deviations and uncertainties, which can only be partially corrected.

Data Completeness

- Missing measurements of the air density or air pressure result in considerable deviation in the mean wind power density and power curve.
- The air density can be well approximated by a correction to the location altitude and the measured ambient temperature. The weather-related fluctuation of the air pressure leads to an uncertainty of about 0.013 kg/m³.

- Measuring the ambient humidity is only of little additional use. It can be assumed to be 50%.
- When approximating the air density, the uncertainty of the mean wind power density is approximately 1 W/m^2 . At high wind speeds, this value may be higher.
- If a density correction of the wind speed is carried out based on an approximated air density, the corresponding uncertainty of the calculated power curve leads to an uncertainty of about 0.25% when calculating annual energy yields.

The results show that KPIs are affected by differing approaches and carefulness in data handling. In many cases, the resulting KPI uncertainties can be significantly reduced by simple measures, and the KPIs remain reliable. These are mostly systematic or epistemic uncertainties that can be largely corrected by additional effort. However, there are also substantial aleatoric uncertainties that depend, among other things, on the location and weather conditions but also on technical specifications and the operational behavior of WTs.

The present work provides a quantification of the uncertainties wherever possible. However, uncertainties vary with site- and technology-specific conditions in many cases. These results cannot be generalized and require specific simulations for valid uncertainty quantification. Regardless, the presented results can be used to estimate the magnitude and relevance of the uncertainties.

As shown in Section 3, there are additional potential causes for KPI uncertainties. Some of them are already part of previous publications and standards, while others require further research. Although this work covers the most important KPIs, there are many additional KPIs that could be of interest in the future. The mechanisms to simulate and test data handling-related uncertainties remain unchanged, and can be easily applied to additional KPIs. Furthermore, there is a need for additional research to quantify uncertainties due to decreasing data availability. A data quality index is required to describe the randomness and length of data gaps used to estimate the corresponding uncertainties more precisely. The current results provide a best- and worst-case scenario, which represent a starting point for future work.

To consider KPI uncertainties in daily work, there is a need to incorporate the corresponding calculations and visualizations into operational management and business intelligence systems. As shown by the IEC standard 61400-12-1/2 [30,31] as a best practice example, guidelines to combine various uncertainty causes and default values need to be agreed upon.

Author Contributions: S.P., S.F., and K.R. designed the research question; S.P. performed the experiments, analyzed the data and wrote the paper; and S.F. and K.R. reviewed the paper. All authors have read and agreed to the published version of the manuscript.

Funding: The research work of Fraunhofer IEE for this paper was funded by the German Federal Ministry for Economic Affairs through the WInD-Pool (grant No. 0324031A) project. Costs to publish in open access were covered by the Fraunhofer-Gesellschaft.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	artificial intelligence
CMS	condition monitoring system
FGW e.V.	Fördergesellschaft Windenergie und andere Dezentrale Energien
GUM	Guide to the Expression of Uncertainty in Measurement
HSE	Health, Safety and the Environment
IEC	International Electrotechnical Commission
IT	information technology
KPI	key performance indicator

LCOE	levelized cost of energy
MEASNET	Measuring Network of Wind Energy Institutes
NREL	National Renewable Energy Laboratory
O&M	operation and maintenance
OpenOA	Open-Source Tool for Wind Farm Operational Performance Analysis
SCADA	supervisory control and data acquisition system
WT	wind turbine
WF	wind farm
WInD-Pool	Wind Energy Information Data Pool

References

1. Stehly, T.J.; Beiter, P.C.; Heimiller, D.M.; Scott, G.N. *2017 Cost of Wind Energy Review*; National Renewable Energy Lab. (NREL): Golden, CO, USA, 2017; doi:10.2172/1475534. [\[CrossRef\]](#)
2. Wiser, R.; Bolinger, M.; Lantz, E. Assessing wind power operating costs in the United States: Results from a survey of wind industry experts. *Renew. Energy Focus* **2019**, *30*, 46–57. [\[CrossRef\]](#)
3. IEA Wind. *IEA Wind TCP Task 26—Wind Technology, Cost, and Performance Trends in Denmark, the European Union, Germany, Ireland, Norway, Sweden, and the United States: 2008–2016*; National Renewable Energy Laboratory: Golden, CO, USA, 2018.
4. Arwas, P.; Charlesworth, D.; Clark, D.; Clay, R.; Craft, G.; Donaldson, I.; Dunlop, A.; Fox, A.; Howard, R.; Lloyd, C.; et al. *Offshore Wind Cost Reduction: Pathways Study*; The Crown Estate: London, UK, 2012.
5. PriceWaterhouseCoopers. *Guide to Key Performance Indicators: Communicating the Measures that Matter*; PriceWaterhouseCoopers: London, UK, 2007.
6. Pfaffel, S.; Faulstich, S.; Sheng, S. Recommended key performance indicators for operational management of wind turbines. *J. Phys. Conf. Ser.* **2019**, *1356*, 012040. [\[CrossRef\]](#)
7. Gonzalez, E.; Nanos, E.M.; Seyr, H.; Valdecabres, L.; Yürüşen, N.Y.; Smolka, U.; Muskulus, M.; Melero, J.J. Key Performance Indicators for Wind Farm Operation and Maintenance. *Energy Procedia* **2017**, *137*, 559–570. [\[CrossRef\]](#)
8. Stetco, A.; Dinmohammadi, F.; Zhao, X.; Robu, V.; Flynn, D.; Barnes, M.; Keane, J.; Nenadic, G. Machine learning methods for wind turbine condition monitoring: A review. *Renew. Energy* **2019**, *133*, 620–635. [\[CrossRef\]](#)
9. Traiger, E. Machine Learning for Automated Detection of Wind Farm Underperformance. In Proceedings of the the WindEurope Analysis of Operating Wind Farms Workshop, Vilnius, Lithuania, 16 May 2018.
10. Sanchez-Marquez, R.; Albarracin Guillem, J.M.; Vicens-Salort, E.; Jabaloyes Vivas, J. A statistical system management method to tackle data uncertainty when using key performance indicators of the balanced scorecard. *J. Manuf. Syst.* **2018**, *48*, 166–179. [\[CrossRef\]](#)
11. Perotto, E.; Canziani, R.; Marchesi, R.; Butelli, P. Environmental performance, indicators and measurement uncertainty in EMS context: A case study. *J. Clean. Prod.* **2008**, *16*, 517–530. [\[CrossRef\]](#)
12. Torregrossa, D.; Schutz, G.; Cornelissen, A.; Hernández-Sancho, F.; Hansen, J. Energy saving in WWTP: Daily benchmarking under uncertainty and data availability limitations. *Environ. Res.* **2016**, *148*, 330–337. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Feiz, R.; Ammenberg, J.; Baas, L.; Eklund, M.; Helgstrand, A.; Marshall, R. Improving the CO₂ performance of cement, part I: Utilizing life-cycle assessment and key performance indicators to assess development within the cement industry. *J. Clean. Prod.* **2015**, *98*, 272–281. [\[CrossRef\]](#)
14. Martin, R.; Lazakis, I.; Barbouchi, S.; Johanning, L. Sensitivity analysis of offshore wind farm operation and maintenance cost and availability. *Renew. Energy* **2016**, *85*, 1226–1236. [\[CrossRef\]](#)
15. Dykes, K.; Ning, A.; King, R.; Graf, P.; Scott, G.; Veers, P. Sensitivity Analysis of Wind Plant Performance to Key Turbine Design Parameters: A Systems Engineering Approach. In Proceedings of the AIAA SciTech 2014, National Harbor, MD, USA, 13–17 January 2014; doi:10.2514/6.2014-1087. [\[CrossRef\]](#)
16. Craig, A.; Optis, M.; Fields, M.J.; Moriarty, P. Uncertainty quantification in the analyses of operational wind power plant performance. *J. Phys. Conf. Ser.* **2018**, *1037*, 052021. [\[CrossRef\]](#)
17. Craig, A. Uncertainty Quantification in Wind Plant Energy Estimation. In *AIAA Scitech 2019 Forum*; American Institute of Aeronautics and Astronautics: Reston, VA, USA, 2019; p. 1. [\[CrossRef\]](#)

18. Fitz-Gibbon, C.T. *Performance Indicators*; Multilingual Matters: Bristol, UK, 1990.
19. Shahin, A.; Mahbod, M.A. Prioritization of key performance indicators. *Int. J. Product. Perform. Manag.* **2007**, *56*, 226–240. [\[CrossRef\]](#)
20. Hahn, B. 17. *Wind Farm Data Collection and Reliability Assessment for O&M Optimization: Expert Group Report on Recommended Practices*, 1st ed.; IEA Wind TCP: Olympia, WA, USA, 2017.
21. Lindley, D.V. *Understanding Uncertainty*; Wiley: Hoboken, NJ, USA, 2013. [\[CrossRef\]](#)
22. JCGM. *Evaluation of Measurement Data—Guide to the Expression of Uncertainty in Measurement (GUM)*: JCGM 100:2008; JCGM—Joint Committee for Guides in Metrology: Sevres, France, 2008.
23. Dangendorf, S.; Burzel, A.; Wahl, T.; Mudersbach, C.; Jensen, J.; Oumeraci, H. *Unsicherheits- und Sensitivitätsanalyse im Rahmen einer integrierten Risikoanalyse: Zwischenbericht Aktivität 4.5 im Forschungsprojekt XtremRisK (03F0483A)*; TU Braunschweig: Brunswick, Germany, 2012.
24. Daneshkhah, A.R. Uncertainty in Probabilistic Risk Assessment: A Review. *Nucl. Eng. Des.* **2004**, *115*, 173–179.
25. Mudersbach, C. Untersuchungen zur Ermittlung von hydrologischen Bemessungsgrößen mit Verfahren der instationären Extremwertstatistik. Ph.D. Thesis, Universität Siegen, Siegen, Germany, 2009.
26. Ang, A.H.S.; Tang, W.H. *Probability Concepts in Engineering: Emphasis on Applications in Civil & Environmental Engineering*, 2nd ed.; Wiley: Hoboken, NJ, USA, 2007.
27. Merz, B. *Hochwasserrisiken: Grenzen und Möglichkeiten der Risikoabschätzung*; Schweizerbart: Stuttgart, Germany, 2006.
28. JCGM. *Auswertung von Messdaten—Eine Einführung zum “Leitfaden zur Angabe der Unsicherheit beim Messen” und zu den dazugehörigen Dokumenten*: JCGM 104:2009; JCGM: Sevres, France, 2009.
29. FGW. *Technical Guideline for Power Plants Part 6: Determination of Wind Potential and Energy Yields*; FGW: Berlin, Germany, 2017.
30. International Electrotechnical Commission. *Power Performance Measurements of Electricity Producing Wind Turbines (IEC 61400-12-1)*; International Electrotechnical Commission: Geneva, Switzerland, 2017.
31. International Electrotechnical Commission. *Power Performance of Electricity-Producing Wind Turbines Based on Nacelle Anemometry (IEC 61400-12-2)*; International Electrotechnical Commission: Geneva, Switzerland, 2013.
32. Zhang, J.; Hodge, B.M.; Gomez-Lazaro, E.; Lovholm, A.L.; Berge, E.; Miettinen, J.; Holttinen, H.; Cutululis, N.; Litong-Palima, M.; Sorensen, P.; et al. *Analysis of Variability and Uncertainty in Wind Power Forecasting: An International Comparison: Preprint*; U.S. Department of Energy Office of Scientific and Technical Information: Oak Ridge, TN, USA, 2013.
33. Lange, B.; Rohrig, K.; Ernst, B.; Schlögl, F.; Cali, Ü.; Jursa, R.; Moradi, J. Wind power prediction in Germany—Recent advances and future challenges. In Proceedings of the Poster at European Wind Energy Conference, Athen, Greece, 27 February–2 March 2006.
34. Sherwin, B.; Fields, J. *IEC 61400-15 Working Group: Progress Update #2 – Meeting 13*; National Renewable Energy Laboratory: Golden, CO, USA, 2018.
35. MEASNET. *Calibration Uncertainty Parameters in MEASNET Wind Tunnels used for Anemometer Calibration*; MEASNET: Madrin, Spain, 2012.
36. International Electrotechnical Commission. *Production Based Availability for Wind Turbines (IEC 61400-26-2)*; International Electrotechnical Commission: Geneva, Switzerland, 2014.
37. International Electrotechnical Commission. *Time Based Availability for Wind Turbines (IEC 61400-26-1), 2010-12.*; International Electrotechnical Commission: Geneva, Switzerland, 2012.
38. Madsen, H. *Time Series Analysis*; Texts in Statistical Science Series; Chapman & Hall/CRC: Boca Raton, FL, USA, 2008; Volume 72.
39. Tavner, P.; Edwards, C.; Brinkman, A.; Spinato, F. Influence of Wind Speed on Wind Turbine Reliability. *Wind Eng.* **2006**, *30*, 55–72. [\[CrossRef\]](#)
40. Schmiedel, A. Untersuchung des Informationsverlustes von Zeitreihen beim Übergang von Minuten- zu Viertelstundendurchschnittswerten. Bachelor’s Thesis, Technische Hochschule Chemnitz, Chemnitz, Germany, 2011.
41. Faulstich, S.; Pfaffel, S.; Hahn, B. Performance and reliability benchmarking using the cross-company initiative WInD-Pool. In Proceedings of the RAVE Offshore Wind R & D Conference, Bremerhaven, Germany, 14 October 2015.

42. FGW. *Technical Guideline for Power Plants Part 10: Determination of Site Quality Following Commissioning*; FGW: Berlin, Germany, 2018.
43. Hirsch, J.; Faulstich, S.; Fraunhofer Institute for Energy Economics and Energy System Technology. *HERA-VPP—High Efficiency, Reliability, Availability of Virtual Power Plants: Abschlussbericht: Laufzeit des Vorhabens: 01.08.2014–31.01.2016*; Fraunhofer Institute for Energy Economics and Energy System Technology: Kassel, Germany, 2016. [\[CrossRef\]](#)
44. Little, R.J.; Rubin, D.B. *Statistical Analysis with Missing Data*; Wiley Series in Probability and Mathematical Statistics; Wiley: New York, NY, USA, 1987.
45. Aubinet, M.; Vesala, T.; Papale, D. (Eds.) *Eddy Covariance: A Practical Guide to Measurement and Data Analysis*; Springer Atmospheric Sciences; Springer: Dordrecht, The Netherlands, 2012.
46. Klaas, T.; Pauscher, L.; Callies, D. LiDAR-mast deviations in complex terrain and their simulation using CFD. *Meteorol. Z.* **2015**, *24*, 13. [\[CrossRef\]](#)
47. ENGIE. Welcome to ENGIE's First Open Data Windfarm. Available online: <https://opendata-renewables.engie.com> (accessed on 24 April 2019).
48. MEASNET. *Evaluation of Site-Specific Wind Conditions: Version 2*; MEASNET: Madrid, Spain, 2016.
49. Optis, M.; Perr-Sauer, J.; Philips, C.; Craig, A.E.; Lee, J.C.Y.; Kemper, T.; Sheng, S.; Simley, E.; Williams, L.; Lunacek, M.; et al. OpenOA: An Open-Source Code Base for Operational Analysis of Wind Power Plants. *Wind Energy Sci. Discuss.* **2019**, 1–14, doi:10.5194/wes-2019-12. [\[CrossRef\]](#)
50. Optis, M. OpenOA: Open-Source Tool for Wind Farm Operational Performance Analysis. In Proceedings of the Drivetrain Reliability Collaborative Meeting 2019, Golden, CO, USA, 19–20 February 2019.
51. NREL/OpenOA: *GitHub Repository*; National Renewable Energy Laboratory: Golden, CO, USA, 2019. Available online: <https://github.com/NREL/OpenOA> (accessed on 18 June 2019).
52. Grange, S.K. *Technical Note: Averaging Wind Speeds And Directions*; University of Auckland: Auckland, New Zealand, 2014; doi:10.13140/RG.2.1.3349.2006. [\[CrossRef\]](#)
53. Bertino, S. A Measure of Representativeness of a Sample for Inferential Purposes. *Int. Stat. Rev.* **2006**, *74*, 149–159. [\[CrossRef\]](#)
54. PCWG. Power Curve Working Group—Part of CFARs. Available online: <https://pcwg.org> (accessed on 10 October 2019).

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).