DTU Library

Uncovering wind power forecasting uncertainty sources and their propagation through the whole modelling chain

Yan, Jie; Möhrlen, Corinna; Göçmen, Tuhfe; Kelly, Mark; Wessel, Arne; Giebel, Gregor

Published in:
Renewable and Sustainable Energy Reviews

Link to article, DOI:
10.1016/j.rser.2022.112519

Publication date:
2022

Document Version
Publisher's PDF, also known as Version of record

Renewable and Sustainable Energy Reviews 165 (2022) 112519

Contents lists available at ScienceDirect

Renewable and Sustainable Energy Reviews

journal homepage: www.elsevier.com/locate/rser

Uncovering wind power forecasting uncertainty sources and their propagation through the whole modelling chain ☆

Jie Yan [a], Corinna Möhrlen [b], Tuhfe Göçmen [c], Mark Kelly [c], Arne Wessel [d], Gregor Giebel [c,*]

[a] North China Electric Power University, State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, Beijing, PR China
[b] WEPROG, Dreijervaenget 8, 5610 Assens, Denmark
[c] Technical University of Denmark, Department of Wind and Energy Systems, Frederiksborgvej 399, 4000 Roskilde, Denmark
[d] Fraunhofer Institute for Energy Economics and Energy System Technology IEE, Kassel, Germany

ARTICLE INFO

Keywords:
Wind power
Forecast uncertainty
Modelling chain

ABSTRACT

Wind power forecasting has supported operational decision-making for power system and electricity markets for 30 years. Efforts of improving the accuracy and/or certainty of deterministic or probabilistic wind power forecasts are continuously exerted by academics and industries. Forecast errors and associated uncertainties propagating through the whole forecasting chain, from weather provider to end user, cannot be eliminated completely. Therefore, understanding the uncertainty sources and how these uncertainties propagate throughout the modelling chain is significant to implement more rational and targeted uncertainty mitigation strategies and standardise the forecast and uncertainty validation. This paper presents a qualitative review on wind power forecasting uncertainty. First, the definition of uncertainty sources throughout the forecast modelling chain acts as a guiding line for checking and evaluating the uncertainty of a wind power forecast system/model. For each of the types of uncertainty sources, uncertainty mitigation strategies are provided, starting from the planning phase of wind farms, the establishment of a forecasting system through the operational phase and market phase. Our review finalises with a discussion on uncertainty validation with an example on ramp forecast validation. Highlights are a qualitative review and discussion including: (1) forecasting uncertainty exists and propagates everywhere throughout the entire modelling chain, from the planning phase to the market phase; (2) the mitigation efforts should be exerted in every modelling step; (3) standardised uncertainty validation practice, including why global data samples are required for forecasters to improve model performance and for forecast users to select and evaluate forecast model outputs.

1. Introduction

High penetration of wind power has been recognised globally as one of the most important features of current and future sustainable power systems. The natural randomness and variability of the wind itself can aggravate negative impacts of wind power on power system operation and market trading, which strengthens the significance of forecasting technology. Wind power forecasting (WPF) started more than three decades ago [1], with the first operational forecasting tools arriving at system operation level some 10 years later at the Danish transmission system operators ELSAM and Elkraft System [2]. Since then, researchers have been making continuous efforts to improve the forecasting accuracy and reliability.

It is impossible to achieve perfect predictions of wind power at any given time or location, due to many reasons; for instance, one may consider the endogenetic randomness of weather systems, and varying wind turbine performance. Many researchers have established that chaotic atmospheric motions have temporal and spatial scales that typically span more than six orders of magnitude [3–5]. Along with the complex wind field, wind turbine performance creates nonlinear and time-varying uncertainties in wind power forecasting. To improve the value of forecasts and their usage, we qualitatively review in this work three questions, when considering the inherent uncertainty in each forecast: why, when, and to what extent the forecasting uncertainty will happen. There is plenty of literature in this area [6], which can be grouped into the following three categories.

☆ 'This paper was coordinated under the auspices of IEA Wind Task 36 'Forecasting for Wind Energy'. Corinna Mörlen, Tuhfe Göçmen, Mark Kelly and Gregor Giebel were funded by the Danish EUDP project "IEA Wind Task 36 Phase II Danish Consortium", Grant Number 64018-0515. Jie Yan was funded by the "Young Elite Scientists Sponsorship Program by CAST", Grant Number 2019QNRC001.

* Corresponding author.
E-mail address: grgi@dtu.dk (G. Giebel).
URL: http://vindenergi.dtu.dk (G. Giebel).

https://doi.org/10.1016/j.rser.2022.112519
Received 12 October 2021; Received in revised form 22 April 2022; Accepted 26 April 2022
Available online 27 May 2022

Nomenclature

AEP	Annual Energy Production
AGCRN	Adaptive Graph Convolutional Recurrent Network
AI	Artificial Intelligence
ARIMA	Auto-regressive Integrated Moving Average
ARMA	Auto-regressive Moving Average
BRP	Balance Responsible Party
Catboost	Categorical Boosting
CFD	Computational Fluid Dynamics
COSMO LEPS	The Limited-area Ensemble Prediction System developed in the framework of the Consortium for Small-scale Modelling
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
DCRNN	Diffusion Convolutional Recurrent Neural Network
DWD	The German Weather Service
EC	European Commission
ECMWF	European Centre for Medium-Range Weather Forecasts
ENTSO-E	European Network of Transmission System Operators for Electricity
EPEX	European Power Exchange
EPS	Ensemble Prediction System
ERCOT	Electric Reliability Council of Texas
EU	European Union
FLOPs	Floating Point Operations
FN	False Negatives
FP	False Positives
GAN	Generative Adversarial Network
GBM	Gradient Boosting Machine
GCN	Graph Convolution Network
GMAN	Graph Multi-Attention Network
GMM	Gaussian Mixture Model
GRU	Gated Recurrent Unit
GUI	Graphical User Interface
GUM	Guide to the expression of uncertainty in measurement
HSSD	High-speed Shutdown
IEA	International Energy Agency
IEE	(Fraunhofer) Institute for Energy Economics and Energy System Technology
IWES	(Fraunhofer) Institute for Wind Energy Systems
KNN	K- Nearest Neighbor
LightGBM	Light Gradient Boosting Machine
LSTM	Long Short Term Memory
MAE	Mean Absolute Error
MEASNET	Measuring Network of Wind Energy Institutes
ML	Machine Learning
MPPT	Maximum Power Point Tracking
MSEPS	Multi-Scheme Ensemble Prediction System
NCEP	National Centers for Environmental Prediction
NOAA	National Oceanic and Atmospheric Administration
NWP	Numerical Weather Prediction
RANS	Reynolds Averaged Navier–Stokes
RMSE	Root Mean Square Error
RNNs	Recurrent Neural Networks
SCADA	Supervisory Control and Data Acquisition
StDev	Standard Deviation
STGCN	Spatial–Temporal Graph Convolutional Network
TN	True Negatives
TP	True Positives
TSO	Transmission System Operator
UQ	Uncertainty Quantification
WFIP	Wind Forecasting Improvement Project
WPF	Wind Power Forecasting
WRA	Wind Resource Assessment
Xgboost	eXtreme Gradient Boosting

- **Qualitative Uncertainty Source Identification:** In the guideline of the World Meteorological Organization on communicating uncertainty in forecasts [7], the main sources of uncertainty in weather forecasts are identified as atmospheric unpredictability, (observational) data interpretation, the process of composing a forecast, and forecast interpretation. Other researchers identified the uncertainty sources in numerical weather prediction (NWP) model parameterisations [8], power curves, and prediction algorithms [9–11]. Uncertainty sources have also been found in typical models, case studies and other reviews [6,12,13].
- **Qualitative Uncertainty Description:** In order to take the uncertainty of the weather prediction into account and to be able to define periods where forecasts have high or low predictability [8,14], probabilistic forecasting and the use of ensemble forecasting arose in the early 2000's [2,15,16]. More than a decade later, complex modelling uncertainty was addressed by the validation & verification framework of [17], though it is not (yet) widely applied by WPF end-users. Since 2016, the International Energy Agency (IEA) Wind Task 36[1] on Forecasting has invited researchers worldwide to discuss and refine our understanding on the state-of-the-art of error and uncertainty quantification in NWP and wind power forecasting models [18]. We analyse the extended existing literature on probabilistic forecasting that determine, estimate, represent and communicate the uncertainty in weather and wind power forecasts, decision-making, validation & verification [8,10,17,19–22].
- **Uncertainty Use Cases:** Many end users today (in the 2020's) are still discussing how to employ probabilistic forecast types in their daily decision processes, scheduling, trading, balancing, etc. Even though there are quite a number of use cases of how to employ probabilistic forecasts in the decision making processes in power system management and trading (see e.g. [20,22–24]), a common question posed is how to approach uncertainty in *binary* or discrete decision processes [21,25–27].

These categories have underlined the significance of forecast uncertainty and its facets for forecasters and decision-makers alike. Consequently, they have been increasing their efforts in uncertainty quantification (UQ) for the planning of wind farms (e.g. [28]), wind farm

[1] Task 36 (2016–2021) initiatives were in 2022 transferred into a broader perspective and relaunched as Task 51 "Forecasting for the Weather Driven Energy System"

performance (e.g. [29,30]) and its application to operational forecasting and marketing practices [31–35]). However, awareness and understanding of probabilistic forecasts – and application of UQ to such – are not (yet) widespread enough to support uncertainty mitigation and improved use of uncertainty in WPF by many end-users (see e.g. [36,37]). To shed light into this research gap, one aim of this review is to contribute with a novel and systematic logic to summarise existing work relating to WPF uncertainty. It analyses the quality of knowledge regarding the uncertainties involved, and representation of the latter in the WPF modelling chain throughout the existing literature.

This review article is designed to be used as a kind of "uncertainty dictionary" for the community, including clear definitions and comprehensive description of all uncertainty sources relevant to wind power forecasting (WPF). Our primary objective is to guide forecast users and wind farm developers through the typical model chain: to point out where and how uncertainty arises and propagates, in order to increase their awareness of potential issues and pitfalls — before they examine and improve their WPF models and forecasting systems. Fig. 1 shows the uncertainty chain through the three main phases of a wind power project: the planning phase, the operational phase, and the marketing phase. A second objective of this review is to describe the mitigation of uncertainty in those three phases, to help overcome the perception of WPF uncertainty sources as potential barriers for the integration of uncertainty forecasts into energy-related decision-making problems [20,24].

The logic behind our objectives led to this review being organised as follows: Section 2 reviews the sources of uncertainty in wind power forecasting induced from the data, model, wind-to-power conversion, etc. From Sections 3 to 5, the authors review the uncertainty evaluation and mitigation strategies in the planning, operation, and market phases corresponding to each uncertainty source mentioned in Section 2. Section 6 describes advanced algorithms and methods to mitigate uncertainty in WPF. Section 7 discusses the uncertainty in validating forecast methods and models. Section 8 summarises our review and look into current trends and future work associated with uncertainty aspects of wind power forecasting.

2. Uncertainty sources in wind power forecasting

The concept of uncertainty can be defined as not knowing the exact value of a quantity, and can be divided into two different parts.

The first is *aleatoric* uncertainty, which encompasses random variability due to the stochastic behaviour of a system (e.g., the atmosphere); it leads to limited ability in repeating identical conditions, such as measuring or modelling wind speed under some presumed 'state' of the atmosphere.

The second type of uncertainty is *epistemic*, which is related to a lack of knowledge; this can involve measurement error, model deficiencies, limited number of observations (statistical sampling), and representativeness of models or observations. It is worth noting, as Kiureghian & Ditlevsen [38] wrote, that "the nature of uncertainties and how one deals with them depends on the context and application".

In the following subsection we describe the sources of uncertainty in the modelling chain, as well as the input/output data to/from the models, and the methods that describe the model uncertainties; this is shown in Fig. 2. It can be said that these imply an initial focus upon epistemic uncertainties.

2.1. Data quality and availability

Data is the foundation of wind power forecasting, particularly because WPF often employs data-driven models. Several factors related to the data introduce uncertainties into WPF:

- data variety
- data size and representativeness
- data contamination
- measurement/instrumental uncertainty

In general, data from a larger number of applicable categories, more sufficient and representative data samples, and less data contamination are likely to lead to smaller forecast uncertainty. In practice, a certain level of data quality and amount of data types and sample sizes are needed, to keep forecast uncertainty at an acceptable level — for both forecast model training and actual operational forecasting.

2.1.1. Data variety

The basic data requirement for wind power forecasting and training/tuning is wind speed and direction data at a wind farm along with power data, typically from a wind turbine's SCADA system. Besides wind, other weather data from NWP models as well as meteorological measurements, produced in the planning phase of a wind farm, are often available for initial forecast model setup and/or training. Dependent on the wind power forecasting methodology more or less of the collected and measured types of data might be required. Forecast uncertainty can be reduced as well as enhanced by the amount of available input data. Generally, the higher the number of degrees of freedom in a model, the higher the uncertainty of its predictions. Another aspect of the data variety with respect to uncertainty is the quality of the data. With low-quality data, the uncertainty may not be reduced, even if using the most sophisticated algorithms.

2.1.2. Size & representativeness of training sample

The data used for WPF model training should ideally cover the weather types, wind scenarios, time stamps and wind turbine operational conditions for a given location, to ensure generalisation and use of data-driven modelling. One-year data sets are generally regarded as a safe size for WPF training samples. However, fewer samples can provide adequate performance under certain scenarios; e.g. newly-built wind farms or locations with complex weather systems, if a refined technical route (see subsection Refined Route 2.2.2) is applied.

To train WPF models, historical performance measurements of a wind park and associated weather forecasts are needed. These should ideally be available over a period of at least one year to cover all seasons. If shorter training periods are used, strong wind periods might not be covered, and the model will not be able to reproduce the upper part of the turbine power curves. It is also possible that individual wind directions can be underrepresented, so that the model cannot sufficiently capture the wake effects there. Weather models are continuously being developed. With longer wind to power training periods, the weather data may come from several generations of a numerical weather prediction model and can introduce additional uncertainties into the power conversion training. The effect on the power training is hence difficult to estimate, as it depends on the extent of the changes made in the NWP model over the training period.

2.1.3. Data contamination

Good training samples should be able to represent the actual wind conditions and wind-to-power conversion process for a wind turbine at a given site. This can be identified using a scatter-plot of the actual (measured) wind turbine power curve, as shown by the dotted circles in Fig. 3. Data contamination typically arises in situations like:

- gaps and errors of wind observations and power measurements, due to loss of data, communication failure, offset or drifting of instruments, etc. (see Fig. 3);
- the data record is complete but does not reflect the actual and representative wind energy conversion process, due to e.g. wind turbine fault and maintenance or curtailment.

In order to ensure and improve the data quality, it is necessary to 'clean' the data samples before model training. We note that the data cleaning process will inevitably give data that deviates from the real measurements, to varying degrees; excessive cleaning increases the uncertainty of the forecast (thus the cleaning or selection process can be seen as an optimisation problem).

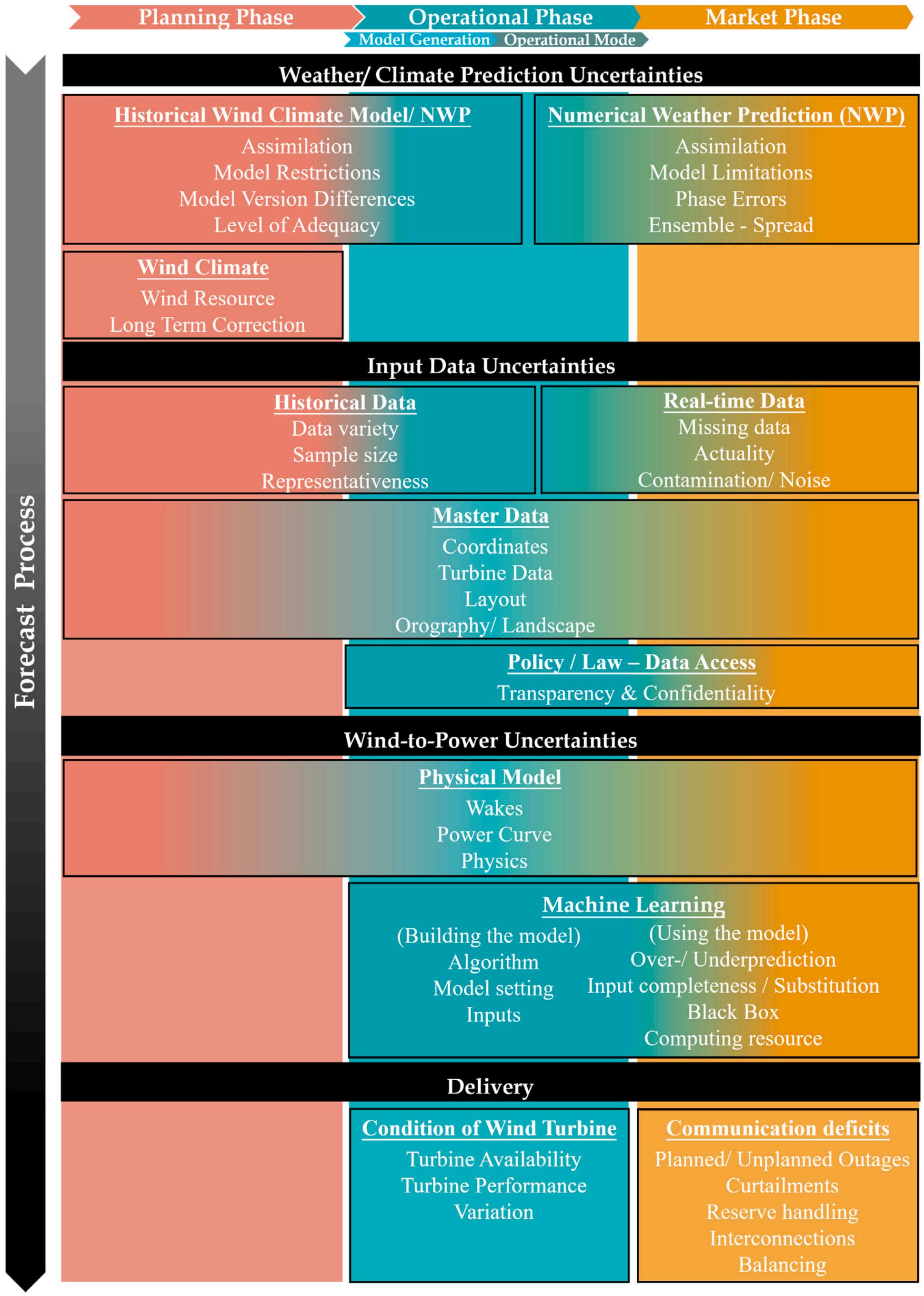


Fig. 1. Overview of the Uncertainty Chain through the 3 phases of a wind project.

2.1.4. Measurement/instrumental uncertainty

Some emphasis may be put on measurement uncertainty, because it: [a] is propagated through the modelling that is involved in the other uncertainty (sub-)components; [b] is involved in uncertainty mitigation; and [c] it also has impact on all phases.[2]

The criteria, pre-defined acceptance range and type of test required to be conducted in the planning phase of a wind project for both wind and power measurements is defined in the IEC 61400-12-1 (2017) and 61400-12-2 standards [39]. The calibration for such industrial wind instrumentation is further detailed in an Annex F of the 61400-12-1 ("Cup anemometer calibration procedure"). Alternatively, the calibration rules of cup anemometers used in wind energy projects can be found in the so-called 'Round Robin rules' recommended by the international Measuring Network of Wind Energy Institutes (MEASNET)[3].

[2] The data contamination described in the previous subsection (Section 2.1.3) could also be considered uncertainty in measurement of the power, though it is typically considered separately, because the data selection/cleaning process extends beyond a simple measurement.

[3] Earlier guidelines on instrument calibration and measurement campaigns for the wind industry can be found in [e.g. 40,41]

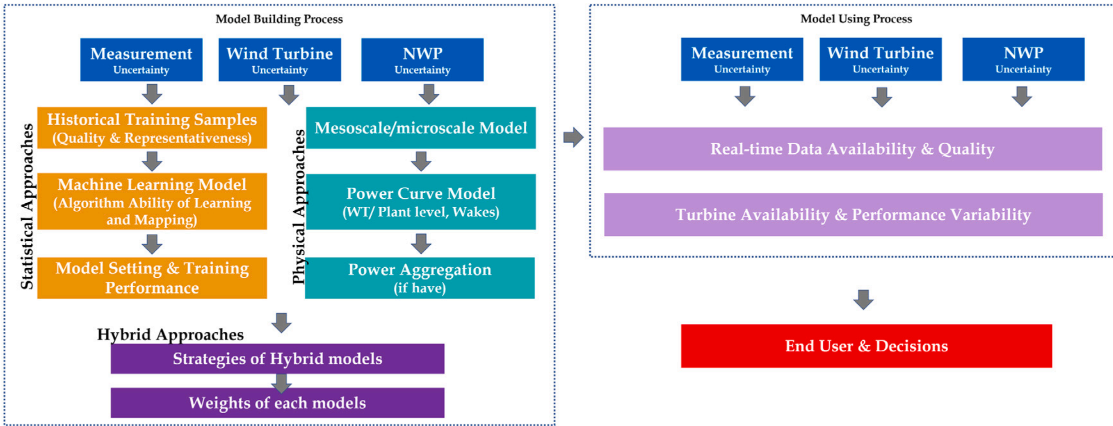


Fig. 2. Overview of the Uncertainty Propagation through the WPF modelling chain.

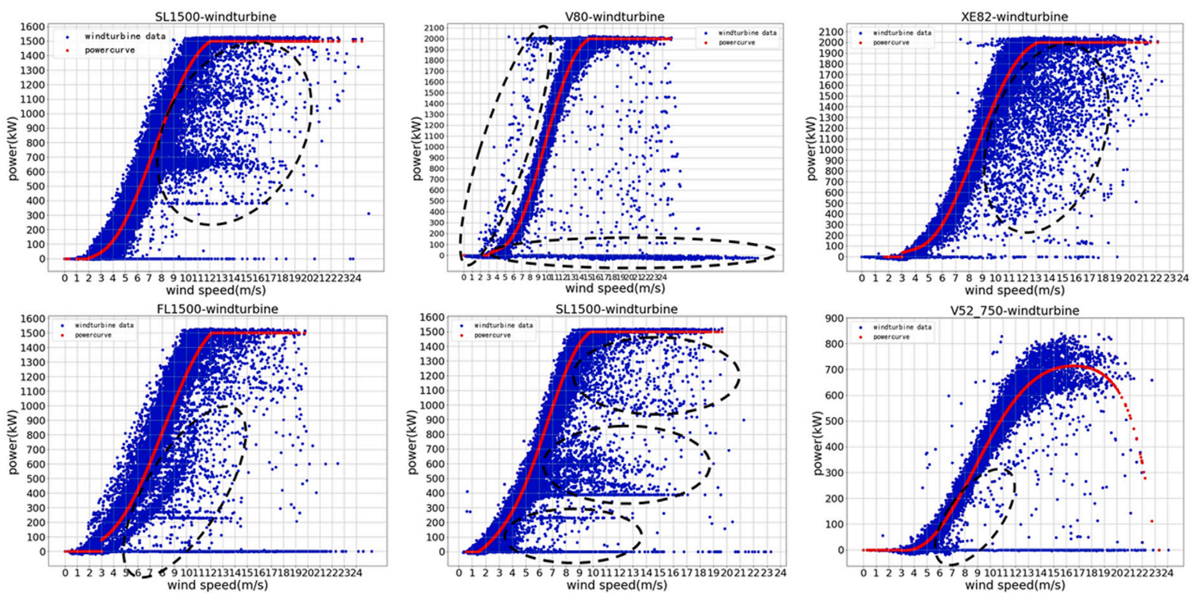


Fig. 3. Data quality problem 2: abnormal power data.

The associated measurement uncertainty evaluation principles are described in detail in an Annex D of the IEC 61400-12-1, which also refers to the so-called GUM.

The 'GUM', which is a commonly used abbreviation for the ISO Guide to the expression of Uncertainty in Measurements [42,43] and its supplement [44]. It defines that the value of a measurand as only complete, when it is accompanied by a statement of the uncertainty of the measurand and otherwise must be considered an approximation or estimate of that measurand. Another way of stating this definition (3.1.2 of the GUM), is that any measurement has an uncertainty associated with it.

Following the GUM, there are two types of measurement uncertainty within any standardised measurements:

- "type A" random errors, i.e. the measuring results of the same quantity with two different instruments are never the same.
- "type B" systematic errors or bias, i.e. offsets in the instruments.

These roughly correspond to the more general concepts of aleatoric and epistemic uncertainties, but are not exactly congruent with them. The 'GUM' [43] is less general, although intends to be applicable to a broad range of measurements such as:

- quality control and quality assurance;
- law and regulation enforcement and compliance;
- calibrating standards, instruments and performing of tests allowing tractability of national standards;
- development, maintenance, and comparison of physical reference standards and materials.

In summary, all of these standards or guidelines (IEC 61400-12, MEASNET, 'GUM' of JCGM) act to hold measurements with sufficient quality to be input for forecasting tools; this facilitates production of high quality forecasts.

2.2. Modelling approaches and technical routes for WPF

Overall, the forecasting techniques for wind power can be divided into deterministic and probabilistic approaches. The latter includes the uncertainty analysis as well as associated risk indices and comprehensive reviews can be found in [6,10,19,20] for wind power forecasting.

For both deterministic and probabilistic forecasting, the methodology can further be divided into three categories: physical, statistical and hybrid. It should be noted that increasingly popular machine learning applications for wind power forecasting (also referred as intelligent methods) are typically considered as an extension of the statistical methods e.g. [45]. However, they can also be considered in the hybrid category, depending on the application [46]. Here in the first part of this section, the main characteristics of the forecasting approaches are described and the corresponding uncertainty factors are listed.

The application of these modelling approaches in terms of the target variable may vary. The technical routes of the WPF in terms of the temporal and spatial scales are grouped in the second part of this section and their potential implications on the associated uncertainties are discussed.

2.2.1. Modelling approaches – statistical & physical & hybrid

Statistical methods generally refer to the application of mathematical statistics, probability theory, and stochastic processes to forecasting problems. They typically use a large amount of historical data for model training or error fitting, establish a mapping relationship between input variables and output variables, and predict the future wind power value (or interval of such) based on the trained model. Established techniques include exponential smoothing, auto-regressive moving average (ARMA) models [47], and auto-regressive integrated moving average (ARIMA) models [48,49]. These can be applied to both deterministic and probabilistic forecasting, as seen in e.g. [50,51]. One of the first operationally applied methods for wind to power conversion was artificial neural networks (ANN) [52]. With the increased use of big data analytics, statistical methods for wind power forecasting today also include artificial intelligence techniques such as machine learning (ML) approaches; common methodologies include support-vector machines, regression tree models, random forest models, gradient boosting, and Gaussian processes. Further ML approaches follow from deep learning and its sub-categories, *i.e.* feed-forward, convolutional and recurrent neural networks, long-short-term memory and gated recurrent unit. If the historical samples are sufficient in volume and variety, most statistical methods can obtain usable prediction accuracy and generalisation ability. Therefore, statistical methods are currently the most widely used forecasting methods, and are a research hotspot in the field of wind power forecasting.

Machine Learning models try to reproduce the data set they have been trained on, as accurately as possible, i.e. to reproduce every detail later in the forecast. If this gets out of hand – the ML model following every fine detail of only the training data set – then the model is considered to be over-trained. To prevent the model from adapting all the noise from the input data and any artefacts, restrictions are placed on it so that it generalises better. This results in a balancing act between generalisation and over-training. On the other hand, an overly generalised model would have the consequence that the output signal would be strongly damped. This trade-off leads to the following uncertainties in the forecast, depending on whether the model is over-generalised or over-trained:

- The generalisation tends to lead to the model searching for its optimal curve in the noisy data. This leads to an underestimation of the forecast, which can result in the model failing to forecast the nominal power of the wind farm.
- The over-training can result in a model trying to reproduce learned artefacts in the later forecast mode, e.g. a curtailment at a specific wind speed/wind direction combination or storm cut-offs that was generated by a single gust.

The balance between over- and under-training leads to the so-called ‘bias–variance dilemma’ [53]: it is usually impossible to fulfil a minimisation of both bias and variance without losing predictive skill.

The (mean-square) model error can be decomposed into three components: bias, variance and a noise term. Bias here means the deviation of the real value from the best possible model (i.e. with training data that represent the whole complexity of the problem) and variance denotes deviation of training from the best possible model. With a simple model, the variance can be kept low because the trained model tends to be closer to the best possible. On the other hand, the error compared to the measured values increases, i.e. the bias. If a more complex

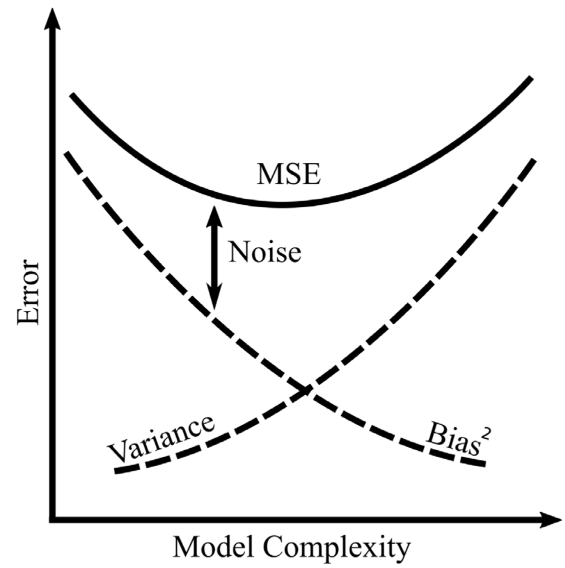


Fig. 4. Generic depiction of model error as a function of the model complexity. The decomposition of the mean-squared error (MSE) in terms of squared bias, variance, and a residual noise term can be seen.

model is now assumed, the error of the optimal model compared to reality is reduced. On the other hand, the deviation of the individual training from the best possible model increases. Through these two dependencies, the total error including noise describes a parabola as a function of model complexity (Fig. 4). The goal is to adjust the model complexity in an optimal way: the model error of the training compared to the best possible model should be balanced with the error of the best possible training compared to the measured values (variance versus bias). Then the minimum of the parabola in Fig. 4 is reached, and the error of the prediction (in terms of model complexity) is minimised.

The uncertainties of statistical models come primarily from the data/training samples, training performance (as a trade-off between generalisation and under-training), and the algorithm matching with the data type, data volume and data quality.

Physical methods typically account for effects of the surface that are on a “micro-scale”, such as terrain differences, surface roughness, and increasingly other factors such as atmospheric stability, to describe the wind field in and around a wind farm. Generally, NWP data are used as boundary conditions to calculate the wind speed and direction at the hub height of each wind turbine using Computational Fluid Dynamics (CFD) simulations as the physical method (flow model). Accordingly, based on the theoretical or fitted power curve of the wind turbine, the predicted wind speed or wind speed error can be converted into a single-point predicted wind power or power interval. The advantage of a physical model is that it does not require historical data and is suitable for building forecasting model(s) of newly built wind farms. However, properly using advanced physical CFD models (typically RANS[4]) can have a high technical threshold and require substantial computational resources, particularly in complex terrain. In order to improve the computational efficiency and accuracy of a physical model with WPF, Li et al. [54] proposed a pre-calculation strategy for the physical forecasting model.

[4] The term CFD is commonly used in some sectors and regions to mean Reynolds-Averaged Navier–Stokes modelling. RANS solves the mean equations of motion, and is often used to model mean flow over complex terrain. CFD also includes large-eddy simulation (LES), which is generally too demanding of computational resources to use outside of research. Technically, CFD also includes simpler models based on reduced flow equations (e.g. WAsP), but the term is not typically applied to denote such.

The uncertainties of a physical model arise from approximations made (in the NWP and flow models), their limited resolution of orography and flow, and time stepping. Much of this can be described as model representativeness. A microscale flow (e.g. CFD) simulation process can improve forecasts, but also has uncertainties, including its turbulence parameterisation and associated parameter choices, resolution dependence, limited or nonexistent treatment of buoyancy, and domain-size limitations.

Hybrid methods are a combination of physical and statistical method. Every wind power forecasting model has errors — i.e. uncertainty in its results, and in some cases, even large biases cannot be avoided. The essential idea of the hybrid forecasting method is to integrate the forecasting results from a variety of algorithms or technical routes (multiple forecasting models), so that large deviations from the target observation due to a single forecasting model can be reduced, along with the overall uncertainty. In other words the predictions from various forecast models are combined, to give better performance on average than any single model would have in terms of uncorrelated error of the target observations. Uncertainty in hybrid models is generated mainly from the choice of sub-models and model weights. A good model portfolio can take advantage of different sub-models; however, poor design may also exacerbate the shortcomings of sub-models.

2.2.2. Technical routes: Decentralised, centralised, & refined

In this subsection we discuss the technical routes of building a WPF model, along with how the associated uncertainties are involved, and can even be exacerbated. Three main routes will be defined and discussed: decentralised, centralised and refined methods. Note that the term decentralised/centralised technical routes is to be understood as an overall design difference when providing independent WPFs and the data sharing aspects of it. This could be for individual wind farms, or for a combined forecast model for a region, control zone or country-wide.

A *decentralised route* predicts the aggregated power output of a wind farm by only using the operational and meteorological data from the wind farm itself. It can also be simply interpreted as different forecast providers independently establish a WPF model, for each wind farm. The advantage of this technical route is that data can be obtained without considering data sharing issues around other wind farms or weather stations. The model scales will not be very large and therefore easier to train for, especially in terms of computational resources. Further, the optimisation goals in the model construction process are clear, that is, to improve the power forecasting accuracy of the target wind farm. However, its disadvantage is that the independent modelling process of any single wind farm ignores the meteorological and spatio-temporal correlations between wind farms, which limits the improvement of forecast accuracy. Moreover, delays in model upgrade and maintenance of a WPF system installed in a relatively remote wind farm will inevitably appear, because of the long journey for the predictor/prediction users to the site in case of instrumentation failures. This can potentially induce further decrease in forecast accuracy.

Although the overall uncertainties tend to be higher within the decentralised route, the sources of uncertainty are typically easier to identify, since the coupling (hence implicit combination of the uncertainties) within the modelling chain are limited.

The *centralised route* refers to establishing a combined power forecasting model for all wind farms in a region or control zone, which can mean that in a given region there is only one forecast provider. Considerations usually include a spatio-temporal correlation of weather information at multiple locations in the region/control zone, where wind speed and wind direction at different geographic locations will be affected by the weather conditions in the upstream neighbouring locations, and will continue to affect the weather conditions in the downstream geographic locations. The advantages of a centralised forecasting route are multiple, as explained below.

- By designing a 'multi-to-multi' mapping network, which is able to reflect weather-related correlations, the forecasting accuracy of every wind farm in the target region can be improved.
- Through data sharing, more weather and power data in the area can be utilised, which increases the scales and dimension (hence generics) of the training data set. This is more suitable for exploiting the advantages of statistical methods in particular to capture the complex and internal relations of large-scale data.
- WPF systems with multiple wind farms will be incorporated into one integrated system, eliminating the need for wind farms to be maintained individually and facilitating system maintenance and management. In addition, this kind of centralised management mode helps wind farm clusters to participate in market bidding as large power generation assets, and is especially suitable for wind farm clusters or large-scale wind power bases.

Since the centralised route for WPF includes combination of several correlated atmospheric and operational parameters, the propagated uncertainties throughout the model chain are highly affected by their interdependence. In addition to a higher level of information considered, frequently observed negative covariances[5] reduces the overall uncertainties of the forecasting system. However, characterising the sources of the resulting uncertainties is also typically more challenging.

The goal of a *refined route* is to classify specific weather scenarios or wind conditions, or to classify the wind turbines into groups according to their spatial distribution, and then establish corresponding representative models towards every scenario and group. In this way, the forecasting accuracy of each scenario and each group can be improved and the uncertainty reduced. The rationale behind this *refined route* is that the two routes mentioned above aim mostly towards establishing a model with sufficient generalisation ability, which can adapt to any season and wind condition by using collected data samples that cover many weather scenarios. However, for a single forecasting model to contain complex nonlinear relationships under different scenarios, is usually difficult to obtain. I.e., obtaining a universal and accurate WPF model, for wind farms with various weather types or complex wind conditions, is quite challenging. Limited by the representativeness of the training samples, the forecasting results of a single model will be levelled out among various weather scenarios, that is, each scenario cannot achieve the corresponding optimal model design as well as the optimal forecasting result.

2.3. Algorithm and model setting

2.3.1. Model complexity and parameters

Algorithms used in the early age of wind power forecasting, e.g. the time series method [55] and Kalman filter method [56], are relatively simple in terms of modelling; they do not rely on numerical weather prediction, but the forecasting accuracy decreases rapidly with increased time horizon [57]. In order to improve the prediction accuracy and extend the usable prediction time horizon, wind power forecasting algorithms are constantly updated, using regression algorithms with stronger learning capabilities, or optimisation algorithms to optimise model parameters. There are two types of commonly used algorithms: neural network (e.g. back propagation neural network, radial basis function neural network, dynamic neural network [58]) and kernel learning (e.g. support vector machine [59,60], relevance vector machine [61], extreme learning machine [62], or Gaussian process [63]). The model complexity and required data scale of the above algorithms are relatively small, so they are widely termed as 'shallow models.' When the training samples involve multi-dimensional large-scale data,

[5] For example, more energetic flow at an upstream wind farm is likely to cause more prominent cluster wakes for downstream wind farms, inducing lower production at the waked wind farms.

complex weather, and low-correlation data, the prediction uncertainty using such algorithms can be large. Therefore, the shallow model is more suitable for the refined technical route mentioned above.

In order to supplement the deficiencies of shallow models, one possibility is to extract data features through manual experience or feature conversion methods (before establishing shallow wind power forecasting models), and then construct a nonlinear mapping relationship between input features and output (wind power). The other possibility is to use a deep learning (DL) algorithm, which is an extension of the shallow model. Deep learning uses a complex and deep network design, which enables it to have strong nonlinear mapping capabilities and better generalisation performance [64]. Due to the complexity of DL models, they can better fit multi-dimensional and large-scale training sample data than shallow models, especially under complex scenarios and/or centralised technical routes. However, it is also more difficult to adjust the parameters of a DL model. Therefore, the model training process is an important source of uncertainty from deep learning prediction.

2.3.2. Model inputs

Forecasts at different time scales might use different forecast inputs. Moreover, these input parameters will introduce different degrees or different forms of errors and uncertainties, in different ways.

Minute/hour-ahead forecasting.

1. Data for Persistence Forecasting.
   The simplest method to deal with minute- or hour-ahead forecasting is to use a persistence forecast, where the last measured value is propagated into the next forecast time step. Persistence forecasting contains a twofold uncertainty that is difficult to quantify, due to the fact that (1) the forecast is based on a time-lagged approximation of a past state and (2) the quality of the measurement defines this state. Nevertheless, it is a method that provides a reasonable estimate of the future, with time resolution of a few minutes to a few hours in steady-state weather situations. To improve upon persistence on a minute scale is a non-trivial task.
2. Time-series Data.
   The minute/hour-ahead wind power forecasting is often based on the time series characteristics and inertia of the air (advection) [65]. Time series data of wind speed or wind power are generally used as model inputs. The uncertainty of model input is in this case reflected in two aspects:

   - Uncertainty of wind resource
     Due to the chaotic and stochastic nature of the atmospheric system, it is difficult to accurately predict the future wind speed, even in a minute/seconds interval. With extension of the forecast horizon, the forecasting uncertainty increases. Meanwhile, the fluctuation amplitudes – and more generally distributions – of wind resources vary over seasons, time of day, and regions, so that the forecasting uncertainty is different.
   - Dropout and length of the time-series model inputs
     Based on information theory, data gaps and anomalies lead to the reduction of time-series information, which leads to the increase of information entropy — and correspondingly, the uncertainty. The learning essence of a time-series-based model is to extract the continuity of time-series data. Abnormal and missing data will pollute and destroy the time-series relationship among the model inputs and learning targets, which would greatly increase the forecasting uncertainty. Moreover, selection of the length of the time-series data of the model input and output will also lead to different forecasting uncertainty.

3. Data assimilation with point measurements and NWP input.
   In [20,65] it has been reviewed how classical data assimilation in NWP can be described as assimilation algorithms that adopt weather or power forecasts with measurements in time and space for individual sites in a region [66,67] (see also Section 4.2.1). Various authors [14,21,65] have shown how an inverted Ensemble-Kalman Filter can make use of power or power-related measurements and NWP model parameters, along with the uncertainty of ensemble input data, to propagate point measurements in time and space. In this case, the uncertainty distribution of the ensemble is used together with the NWP forecast, avoiding computationally expensive feedback into the NWP, to define the future state.

Day-ahead forecasting – NWP. Since the usable forecast horizon of the persistence method and time-series-based methods is very limited, NWP data is indispensable for the day-ahead time scale in wind power forecasting. Due to the power-law relation between wind speed and power, the NWP error has part in forecasting uncertainty of day-ahead wind power. Within our WPF context, uncertainty in NWP can be manifested in two main categories, as described below.

1. **Representation of wind speed fluctuations:** Time series of wind speed measurements appear much more volatile and less smooth than those from NWP. The NWP wind speed sequence always deviates from the actual wind speed sequence at any given time, and demonstrates unclear regularity for upward deviation or downward deviation as well as their deviation amplitude and phase errors, as shown in Fig. 5.
   In general, NWP models calculate internally with a higher temporal resolution (down to 20s) than what is provided to the user after post-processing the raw data. Nevertheless, the fluctuation of the parameters is damped by physics to keep the model stable. The user thus only gets an output for prediction steps of typically one to three hours. These are, strictly speaking, snapshots from the forecast calculation and not mean values. Often these are interpolated to 15-minute averages (which probably fits quite well due to the damping), to compare with power measurements. Interpretations of the associated time stamp (validity as a start, end, or middle time) can lead to phase errors. Due to low temporal resolution output, intra-hourly fluctuations can often not be mapped and the start or end of ramps cannot be determined exactly. On the other hand, higher temporal resolution combined with better representation of the fluctuations in the NWP forecast also generally leads to larger error, as predictions of significant events are mostly out of phase. In a verification this results in a so-called "double punishment" or "double penalty" [68,69] (see also Section 6.2.2 and Fig. 11). In Fig. 6, wind roses from a NWP at multiple points in a Chinese wind farm (10 km × 8 km) as well as the corresponding wind measurements are shown to illustrate the differences between NWP output and measurements.
   The calibration of the wind forecast in the NWP model is traditionally done at 10 m height, where most of the synoptic measurements are taken with met stations carrying anemometers at this pre-defined height above ground. This is often done at locations such as airports, which do not reflect the roughness of the surrounding area. Only in recent years have meteorological centres and institutes started to extract wind speeds at 100 m heights, which correspond more to today's hub heights; this has been facilitated by remote sensing, with data from LiDAR, SODAR, and satellites becoming available. NWP models are not designed for a specific area of application and operate mostly in pressure coordinates (levels), which are terrain-following levels and change according to temperature, air density and pressure. Thus any physical property at a fixed height needs to be calculated with spatial interpolation. These vertical levels depend

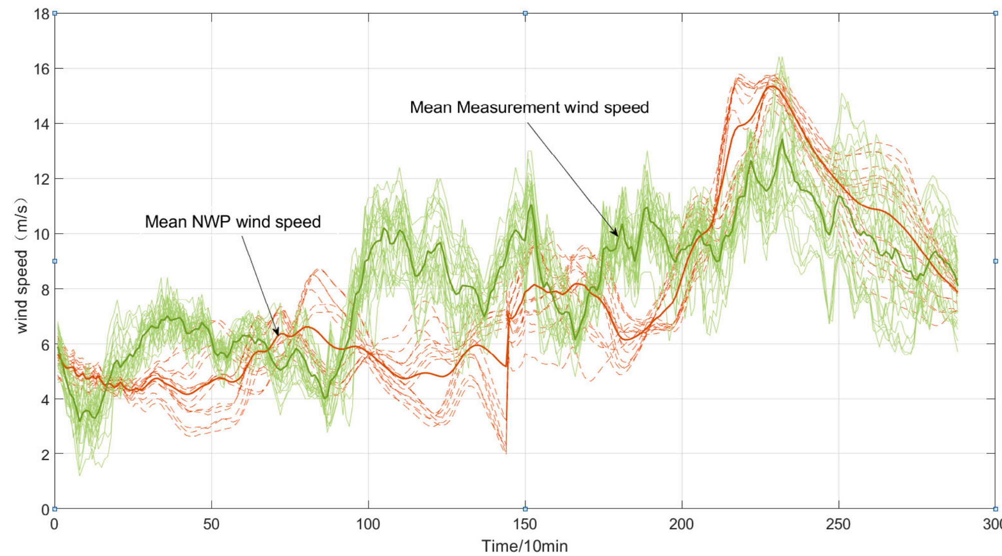


Fig. 5. Time series wind speed data from NWP and measurements.

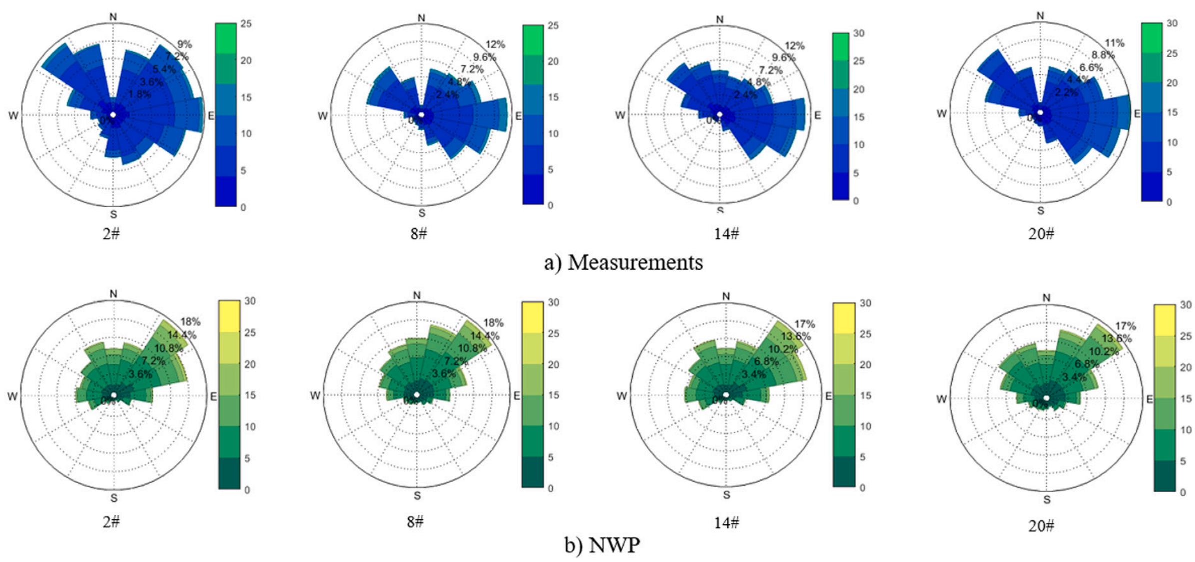


Fig. 6. Wind direction rose map of NWP and measurements.

on many aspects, but are usually dependent on the forecast length regarding the density of levels in the stratosphere and the desired resolution in the troposphere. The longer the forecast length, the larger the numbers in the stratosphere, and the higher the resolution in the boundary layer, the higher the number of levels in the troposphere. In most NWP models the vertical levels have a relatively small spacing near the ground, in order to model boundary layer processes well, with vertical resolution decreasing away from the surface. The modelled wind speed value is considered constant over a level or assumed to be valid for the centre of the level, and must be interpolated to wind turbine hub height. The fact that the rotor covers a larger area, which can extend over at least two or three levels, is often neglected. If the vertical profile of wind would always follow a logarithmic or power-law then simple interpolation would be possible, but this is generally not the case. Statistical models can compensate for this disadvantage to a certain extent by having two height levels as input to the model, where an implicit weighting will be optimised in the training process. Depending on the correction in the weather model, this might lead to systematic overestimation or underestimation of wind speeds at different heights than the calibrated heights, especially in areas with complex orography.

In addition, a weather model always outputs only one (volume average) value per grid cell and assumes a mean orography below. From this perspective it is desirable that grid cells are small, i.e. have a 'high' (fine) horizontal resolution. Enhanced resolution also creates increased degrees of freedom in the numerical equations, and hence increased uncertainty and risk of "double penalties" (see above and Section 6.2.2).

2. **NWP Ensemble Forecasting Uncertainty.** One of the difficulties in quantifying the uncertainty of a weather forecast is the lack of knowledge of the correct state of the atmosphere at any point in time. Nowadays there are millions of measurements taken and exchanged through a global telecommunication system network [70], but the atmosphere has so many degrees of freedom that even with today's computer capacities, it is not possible to describe the atmosphere on a global scale without simplifying models. Due to the lack of knowledge about the true state of the atmosphere at the outset of a forecast, there is

an inherent uncertainty in the forward propagation of weather development in NWP models. Other types of weather uncertainty arise from the assumptions made to solve the partial differential equations that govern the atmosphere, the so-called parameterisation schemes for (1) physical processes such as condensation, vertical diffusion, solar radiation, heat exchange, (2) dynamical processes such as the advection, and (3) climatic data such as vegetation, hydrological properties and terrain effects.

NWP ensemble model systems play a crucial role in the understanding of the weather uncertainty, as they can be configured to describe the uncertainty of a specific parameter over a specific time horizon. Ensemble prediction systems (EPS) such as the European Centre for Medium-Range Weather Forecasts (ECMWF) Ensemble system [71] or the National Centers for Environmental Prediction (NCEP) Ensemble [72] are designed to describe uncertainty over forecast horizons of 3–6 days ('medium' range). Meanwhile, multi-scheme ensemble methods such as the multi-scheme ensemble prediction system (MSEPS) [8,14] or multi-model ensemble methods such as the limited-area ensemble prediction system COSMO LEPS (developed in the framework of the Consortium for Small-scale Modelling [73,74]) characterise and describe the uncertainty of physical and dynamical processes over short- to medium range horizons (0–10 days). One of the difficulties with multi-model ensemble methods is that there is no well-defined uncertainty quantification from the ensemble members; this is due to the collection of deterministic models and their specific, often damping nature to perform with standard meteorological parameters. Designing a multi-model EPS is difficult and extremely expensive, as all models need to be maintained [20].

2.4. Wind resource uncertainty

Uncertainty quantification in wind resource assessment (WRA) has become important enough in the wind energy industry to merit its own IEC standard: the emerging 61400-15-2.[6] WRA can be thought of as multi-year prediction, within the WPF context. Whether based on pre-construction wind measurements or operational turbine data, in practice wind resource uncertainty is generally decomposed into observational and model uncertainties: the measurement uncertainties propagate through a model chain. Further, as shown in Section 2.5, the process of 'converting' wind into electric power also has its own observational and model components. Aspects of power production uncertainty, particularly within a context 'detangled' from wind speed uncertainty (i.e. treating power production modelling as simply propagating wind speed uncertainty) is discussed more there. Beyond energy production, yet more uncertainties are found later, connected with e.g. grid behaviour and availability.

Pre-construction uncertainty is important primarily for project financing, but also provides a basis for expected operational uncertainty, as well as potential limits on forecast uncertainty. Operational-phase uncertainty based on production data is usable for operation, valuation, and re-financing of existing farms, as well as for short-term forecasting.

As mentioned in Section 2.3.2, uncertainty of the wind resource – the *wind field itself* – can be attributed to its stochastic nature. Due especially to turbulence, which is modified by terrain, the chaotic velocity field in the lower atmosphere includes fluctuations over a large range of length and time scales (typically spanning 5–6 orders of magnitude or more, see e.g. Wyngaard, 2010 [5]) and varies randomly in a very complex manner. The strength of wind speed variations depends on the scales considered, and this is affected both by atmospheric stability (buoyancy), local surfaces, and distance from surfaces. NWP does not resolve turbulence, instead parameterising turbulence's mean effects on the mesoscale flow field and other prognostic variables (e.g. temperature, humidity). Thus in WPF, the uncertainty in the wind field can vary depending on the model resolution, as well as the forecast window. Practical methods to account for the uncertainty in the wind field, which include long-term (climatological) variation for WRA (as indicated in the emerging IEC 64100-15-2) as well as unresolved (small-scale) spatio-temporal variation in both WRA and WPF, are also mentioned earlier in this section.

2.5. Wind-to-power uncertainty

Due to the strong variability of the wind field, the random nature of turbine control systems, and non-linearity of power curves, the process of converting wind to power is characterised by randomness, non-linearity and boundedness [10]. One of the sources of uncertainty associated with these nonlinear aspects is the continuous change in the operating state of wind turbines. These are the result of both objective and non-objective effects, and their influence is reflected in the broad width of typical scatter-plots showing actual or measured power curves (see examples in Fig. 3).

2.5.1. Rotor inertia and control offsets

With the influence of rotor inertia under non-stable wind fluctuations, the same incoming wind speed may be converted into different power outputs [75]. For an incoming wind speed V, the power output is assumed to be $P(V)$, according to a static power curve. But when a wind turbine encounters an upward ramp of wind speed, the generator cannot immediately respond but spins faster, though with some lag. Although the wind speed increases to $V_{\mathrm{ramp+}}$, the corresponding power output of the wind turbine is less than $P(V_{\mathrm{ramp+}})$. Similarly, when the incoming wind speed decreases rapidly enough to some V, the power output is greater than $P(V)$. These differences in power can be taken as the uncertainty caused by rotor inertia, which is related to the complex and dynamic changes of the wind speed.

Wind turbines will inevitably have offsets or delays of actuator response. Actual power output of wind turbines deviates from the theoretical value to varying degrees, resulting in uncertainty in the relationship between wind and power output. Such uncertainties can be reduced by optimising control algorithms and upgrading communication equipment.

2.5.2. Turbine availability and performance variation

The majority of the forecasting approaches rely on constant (or at least consistent) data streams for training and validation of the WPF algorithms. Additionally, both for the individual wind farm and the cluster level, the predictions typically assume 100% turbine availability, unconscious of the local operational status and control scenarios applied. Potential sources of WPF uncertainty in terms of turbine availability and local (both temporally and spatially) operational conditions are listed below.

- Wind farm (flow) control: when the turbine(s) are controlled holistically within a wind farm (via e.g. down or up-regulation based axial induction and yaw steering based wake redirection) following several objectives (e.g. grid compliance, wake reduction for power/revenue maximisation, structural load alleviation etc.). As the forecast algorithms are typically tuned for power prediction for normal (or greedy) operation, when wind farm control is activated, it induces bias that can only be corrected by informing the forecaster of the operational set-points. The uncertainty, however, remains critical especially for short-term forecasting as the transient effects (*i.e.* during the trigger and release of the set-points) involve complex aerodynamics and fluid–structure interactions that is generally beyond the adequacy of the WPF models.

[6] The IEC 61400-15 was split in 2020 into two parts, with 61400-15-2 covering uncertainty and 61400-15-1 covering site assessment.

- Under or over-performance: unlike up- or down-*regulation*, under- or over-*performance* of a turbine(s) refers to the significant changes in the power production under maximum-power-point-tracking (MPPT) control (*i.e.* normal operation). Performance monitoring techniques that utilise normal behaviour modelling (e.g. site-specific, expected power curve) are implemented to capture this 'unintentional' deviance in active power, which commonly occurs due to short- or long-term misalignment of the turbine(s) to the incoming wind direction ('yaw misalignment') or component failure which may be temporary or permanent. Potential performance changes of the turbines have a direct impact on the power forecasts at different temporal and spatial resolutions, reducing forecast accuracy and quality, as well as influencing the risk and decision-making processes involved in market engagement.
- High-speed shutdown: as the operational status of the turbine(s) affects the forecast quality, the forecast itself also has an impact on the operation of the turbines. During high-wind periods, forecast uncertainty becomes even more critical as the WPF systems suggest shutdown above a certain threshold. The forecasts then also inform re-connection of the wind farms, which significantly impacts the grid system and its stability. Forecast uncertainty has serious consequences, especially under high wind periods and regions, which becomes increasingly important with higher wind power penetration (grid share). Accordingly, more improved techniques for reliable high wind speed forecasts are more relevant than ever (see also Sections 6.3.1 and 6.3).
- Turbine (component) or communication failure: in addition to the operational issues raised above, maintenance of wind power plants also has a direct implication for WPF accuracy. Especially with increasing wind assets offshore, the downtime for turbines tends to span over longer periods as unscheduled maintenance becomes more demanding in terms of cost and logistics. Combined with larger size and capacity turbines, these off-times result in significant WPF errors even for larger (temporal and spatial) scale forecasts. Additionally, when the data communication fails, real-time data might be unavailable as a model inputs to the forecasting task. This leads the forecasting uncertainty to rise sharply, especially for minute or hour ahead forecasting.
- Performance degradation (long term effect): detrimental changes in the physical condition of the wind turbines occur over time, with usage and external causes inducing a gradual decrease in wind farm power production. Also referred as 'ageing', the authors of [76] claim up to 1.6 ± 0.2% decrease in turbine power output per year, later in their lifetime. This also highlights the importance of the model updates over time for increased WPF quality, with higher weight on the most recent turbine/wind farm behaviour.

3. Uncertainty evaluation and mitigation in the planning phase

There are numerous contributions to uncertainty in the planning phase, as listed in Section 2. This uncertainty is often key to project financing, and can also impact wind farm design.

3.1. Evaluation of pre-construction uncertainties

The statistical basis of wind resource assessment (WRA) is also part of the foundation of uncertainty in wind power forecasting. Pre-construction calculations of the wind resource for proposed projects typically include uncertainty estimates. For existing turbines in operation, uncertainty can also be estimated, using production data [e.g. 77]; this is handled in Section 4.

As mentioned in Section 2.4, wind power uncertainty is generally comprised of observational and model components. In typical pre-construction estimates many models, each treating different effects, are used together in order to predict long-term wind statistics. Each model carries its own (sub-)component uncertainties: e.g., flow modelling over terrain accounting for different mast and turbine positions [78, 79], vertical extrapolation [80,81], and long-term correction [e.g. 82] contribute to statistical uncertainty in prediction of wind resources. Each (sub-)model is either driven by observational data, or by the output of another sub-model which is ultimately driven by observations. Thus, observational uncertainty is propagated through the model 'chain' used for WRA. For pre-construction WRA there tends to be a larger number of processes to account for and models needed, compared to production-based assessment. Following the practical scheme put forth in the IEC 61400-15-2 standard, the primary uncertainty categories in pre-construction wind resource (energy yield) assessment are

- measurement uncertainty;
- historical wind resource;
- project and evaluation period variability;
- horizontal extrapolation (flow modelling);
- vertical extrapolation (profiles and stability).

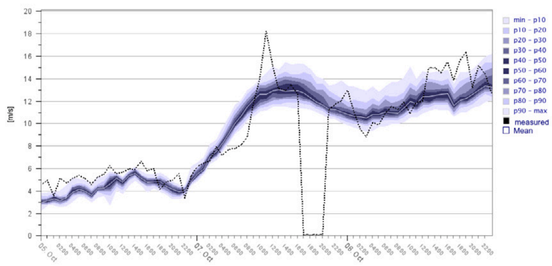
'Historical wind resource' includes uncertainties within the long-term adjustment method used to compensate for pre-construction measurements, which are usually limited to a small fraction of turbine lifetime (∼1–2 years, compared to ∼20 years), and uncertainties inherent in the long-term reference data employed (typically from reanalysis datasets and/or mesoscale models). The project/evaluation period variability includes variation in uncertainty associated with different expected operation horizons (e.g. 1 year vs. 10 or 20 years), as well as impacts of climate change.

For predicting the wind resource, each sub-model (or model component) itself also has an epistemic (systematic) uncertainty, which can be interpreted as how well the sub-model represents reality; this is usually contingent upon the values of the inputs to the sub-model (particularly for simpler parameterisations), and often depends also on other quantities describing the conditions under which the model is used (which may or may not be included as inputs). Typically microscale modelling is used for the flow-related aspects of pre-construction WRA (horizontal and vertical extrapolation), and these each have associated uncertainties, including propagated measurement uncertainty. As an example, a simpler flow model may be more robust to uncertainty in its input (measured wind statistics), but possess more epistemic uncertainty due to its limited ability to represent flow over more complex terrain; one input which may affect its uncertainty, but which is sometimes not provided, is the Obukhov length or its distribution [a metric of atmospheric stability, c.f. 80,83,84].

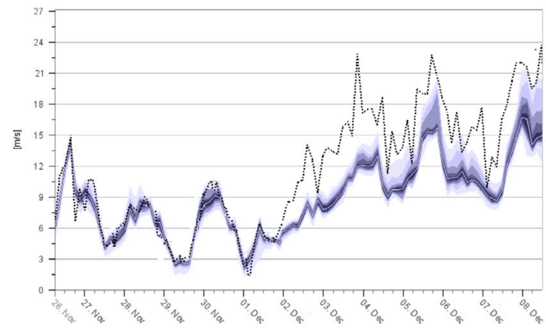
Combination of uncertainty components in WRA typically assumes independence of the various components, leading to simple root-mean-square addition. However, one may more generically combine uncertainties as in the commonly-citepd Guide to Uncertainty in Measurement ('GUM', including [42,43] with supplement [44]). Condensing the latter for convenience one can write

$$\sqrt{\sum_i \left[\sigma_i^2 + \sum_{j>i}(\rho_{ij}\sigma_i\sigma_j)\right]}, \tag{1}$$

where σ_i is the ith uncertainty component and ρ_{ij} is the correlation between components σ_i and σ_j. In the event that multiple wind measurements are used, or both production data and wind measurements are employed, then a more complicated version of (1) arises. To be practical, in WRA engineers typically assign weights to each data source (linearly, as coefficients on each σ_i^2 in Eq. (1)) depending on how much the source affects the wind prediction (or energy), and ignore the correlations. e.g. for anemometers at multiple locations, weights can be assigned according to the fraction of predicted energy corresponding to each measurement source. The latter example also hints at some of the practical difficulty of separating the uncertainty in wind resource



(a) Example 1



(b) Example 2

Fig. 7. Examples of measurement outliers and changes in functionality of an instrumentation and the possibility an ensemble forecast provides to detect such issues in an automatic way.

assessment from uncertainty in *energy yield assessment*: the models for power production prediction, or for its uncertainty, may be inseparable from those for wind prediction. To be more concrete, a related aspect is that the (uncertainty in) power production or power curve can itself depend non-linearly on the (uncertainty in) the wind resource. This becomes quite complicated in the case of interacting wakes, and the uncertainties may be modelled in a linearised way in order to separate wind and power production.

In addition to resource assessment and its associated uncertainties, there is also uncertainty within *site assessment* – the evaluation of conditions at a site in order to find an appropriate turbine class – in the planning phase of a wind farm. The IEC 61400-1 standard outlines this, requiring the following items associated with measurements: extreme winds; vertical shear of the mean wind profile (dU/dz or shear exponent $d \ln U/d \ln z$); flow inclination angle; ambient turbulence intensity; wind speed distribution; and wake-induced turbulence.

These are obtained by a combination of measurements and models, having uncertainties associated with each; but such siting uncertainty is beyond the scope of this article. However, it should be noted that uncertainties in these quantities can also affect project financing or approval, as they tend to determine the turbine class to be used and the loads expected for a given turbine.

3.2. Pre-construction mitigation: Observational uncertainty reduction via ensemble forecasts

One way to deal with uncertainty in measurement signals is to use ensemble forecasts. A well-calibrated ensemble that resembles the weather conditions can be used to reject outliers in the measurements. This method is especially useful for outliers that are still within the range of the measurement, but unrealistic.

The ensemble spread of a well-calibrated ensemble forecast can also be used to detect whether an instrument has started to malfunction, for example 'drifting' behaviour whereby an offset arises. Ideally a measurement fluctuates around the ensemble-mean value. However, if the measurement begins to fluctuate around a lower or higher percentile of the ensemble forecasts, and persists without significant weather changes, this could be due to a malfunctioning instrument. The plots in Fig. 7 show examples of such outliers and changes in functionality.

3.3. Risk vs. uncertainty, and "bankable" resource data

Uncertainty should be distinguished from *risk*, which may be evaluated in conjunction with uncertainty estimates. In pre-construction estimates, for the past decades it is not unusual to encounter WRA reports which conflate risk and uncertainty; such conflation can arise simply, as uncertainty assessments that are overestimated in an attempt to practically account for perceived risk. As the wind industry converges towards a standard practice(s) for uncertainty estimation in WRA, objective *reporting* of uncertainty in resource calculations (and forecasting) should also improve, including separation of quantified uncertainties and specific project risks. To account for such, reporting has been part of the development of the IEC 61400-15 standard. Common reporting practice – as well as methodology – allows clearer (and more usable) statistics, which are not polluted by subjective evaluations of risk. Probabilistic forecasting – especially ensemble forecasting – may therefore also enter the WRA area in the near future, as an objective means to quantify the uncertainty and associated risk for projects from a resource perspective under climate change. Monte Carlo methods and use of multiple long-term reference data are currently employed by a growing number of consultants and developers, as a step in the direction of this kind of forecasting.

Another view comes from considering that historically wind project financiers have calculated their own risk and/or uncertainty metrics associated with any particular project. These are often different than the uncertainty reported in WRA, and are used to determine e.g. the cost of loans and project financing; they are also driven to some extent by financial considerations, and (implicit) understanding of the biased or semi-quantitative nature of uncertainty estimation within WRA. As the wind industry matures and performs better and more standardised uncertainty quantification in WRA, financial evaluations of risk and uncertainty increasingly incorporate the uncertainties reported in pre-construction project assessments by wind engineers and consultants, but they cannot (yet) be assumed to be identical to such quoted WRA uncertainties.

4. Uncertainty evaluation and mitigation methods in the operation phase

When a wind farm becomes operational, the available information from the site increases. In addition to production data, this can potentially include both temporal and spatial aspects of the wind resource.

In this section we discuss the added value of such additional *operational* data, relating to potential model improvements through uncertainty mitigation, and challenges for improved data analytics.

In the operational phase of a wind farm we have to separate between the functioning of the turbines (hardware), and integration of the production into the electrical grid and market. The uncertainty of measurements that drive and control the wind turbines are generally on a different time scale than that for the forecasting. Even for high-resolution applications, forecasting uncertainties are typically addressed at minute(s)-scale [65], while turbine control and its associated uncertainties are utilised over time intervals of seconds or even milliseconds. However, with wind farm control studies and commercial applications gaining momentum [85], the timescales of the problem to be coupled with the turbine controllers are also increased to several minutes, as the flow within the wind farm is manipulated via the control handles at the turbine. Accordingly, here in this section the interaction of the forecasting approaches and turbine operational/conditional status is discussed from the uncertainty assessment perspective.

4.1. Increased information under operational conditions

4.1.1. Adding wind and turbine performance data

Many grid codes of system operators today require that wind farms deliver meteorological data; projects greater than 5–10 MW in size are usually required to submit such data, in addition to their power production output. Aside from wind speed and direction, typical variables include air density, pressure, and temperature. This data is collected through the SCADA system of the wind farm and/or through direct power line signals; it is commonly used for the short-term wind power forecast process, as well as evaluation of power production in cases of curtailment. In the forecasting process, the wind speed signal is the most important of these variables, particularly for the prediction and detection of high-speed shut down events.

Adding relevant (but not redundant) information can reduce the difficulty of model learning, and reduce the uncertainty of forecasting. This is also in line with the basic principle of information entropy. Presently we identify primary types of additional information, which are part of wind power forecasting (WPF) research:

- **Wind and power measurements at more locations, and spatial–temporal correlations.**
  Wind power forecasting often employs ‘mapping models’ for a given reference location. Many researchers have confirmed that the data at a single location is not necessarily representative of regional weather conditions, and the power generation profiles of wind turbines, and therefore limits the robustness of the forecasting model. The errors and uncertainties of minute or hour scale forecasting can be reduced by considering the temporal–spatial characteristics of wind resources or using the real-time weather data from nearby wind farms/weather stations to correct the forecasting results under sudden weather processes [20,86–88]. Data that can help reduce the forecasting uncertainty might include wind measurement and power data of multiple wind turbines in the station or other adjacent wind farms, and NWP data at multiple locations.
- **NWP from multiple sources or Ensemble Prediction System (EPS).**
  Atmospheric motions exhibit complex and chaotic characteristics. It is difficult for a single-source NWP to always make accurate predictions in various weather conditions. NWP from different sources, e.g. using different initial fields and/or different model configurations, has different adaptability and accuracy under different weather conditions and at different times. Therefore, the use of NWPs from multiple sources or a NWP ensemble prediction system (EPS) can reduce prediction uncertainty.
- **Turbine condition and status information.**
  For planned maintenance or shutdown, forecast uncertainty can be easily reduced by removing the predicted power of the wind turbine that will be under planned maintenance. But for unexpected failures and shutdowns, condition monitoring and failure warnings are beneficial for correcting forecasts in real-time and reducing the uncertainty.

4.1.2. Using feature engineering to incorporate more features

Feature conversion and extraction. As is well known, wind speed is the most relevant feature for wind power generation, and it is also the inevitable input of the power prediction model. However, using wind speed alone as a model input does not always obtain good prediction results. Multi-dimensional meteorological inputs can effectively improve the accuracy of wind power forecasting [89], by accounting for other degrees of freedom in the modelled system, and thus keeping uncertainties in the models under control.

Many WPF researchers have been working on feature conversion and extraction. Useful features include statistics (e.g. minimum, maximum, mean, median, quantiles, logarithmic averages); category (e.g. weather type); time index (e.g., hour of the day, day of the year); multi-dimensional meteorological data (e.g. barometric pressure, temperature, and stability metrics along with wind direction, time lagged inputs of the NWP model). These can effectively improve the accuracy of power forecasting, as well as (pre-)processing input data. The latter – which includes conversion and extraction – has been done using polynomial constructions [90,91], frequency domain decomposition [92], dimensionality reduction via Principal Component Analysis (PCA) [93], deep feature extraction (e.g. auto-encoders [94,95], or with convolutional neural networks (CNN) [96].

Feature selection. Feature conversion and extraction produces a large number of quantities that can be essential for WPF model training. Silva [91] believes that the square and cubic of wind speed are important features to input into WPF models. However, not all features have sufficient relevance, and may even increase the uncertainty of the result [90]. Therefore, it is necessary to filter out driving features as model inputs. Mainstream feature selection methods include filter approaches (e.g. mutual information), wrapper selection methods (methods include minimal Redundancy Maximal Relevance (mRMR) and Recursive Feature Elimination (RFE) and are used to select the optimal feature set for wind and wind power prediction in [97,98]), and embedded selection (e.g. Automatic Relevance Determination [99] and the tree method [100,101]).

4.2. Numerical weather prediction uncertainty

In [20] it was found that two chief barriers for industrial adoption of uncertainty forecasts were lack of understanding of its information content (the physical and statistical modelling), and standardisation of uncertainty forecast products. These lead to mistrust and scepticism towards uncertainty forecasts, and their applicability in practice. In [20] the authors established a reference terminology and reviewed common methods to determine, estimate, and communicate the uncertainty in weather and wind power forecasts. Their conceptual analysis of the state of the art concluded that different uncertainty representations need to be mapped to specific wind energy-related user requirements, and that the industry needs to establish multidisciplinary teams in order to implement stochastic methods into the grid operation and look-ahead planning. [20] also thoroughly analysed and described uncertainty in the NWP modelling process. But within the context of uncertainty propagation, we want to focus on the impact of uncertainty in the NWP modelling process for the next steps in the modelling chain (e.g. conversion to wind power), rather than the various methodologies to generate uncertainty forecasts from NWP models.

4.2.1. Using wind turbine measurements in the data assimilation process

Quality control is an essential part of today's real-time forecasting processes, where measurements are used to adjust the forecasts for the next few hours ahead. Such data must reflect the current weather conditions at and around the power plant. If such data are of insufficient quality, a well calibrated quality control system will reject a large portion of the delivered data, which in return leads to a lower quality forecast for the next few hours and bad decision taken by the users of the forecast, e.g. control room staff or traders. The same is true for warnings that may not be based on forecasts adjusted with trustworthy measurements, which when ignored due to mistrust, can lead to critical situations in the power grid or the physical balancing of the lack or surplus of power.

A 2019 study carried out by the Irish system operator [102] showed that without quality requirements and control, meteorological measurement data was not sufficient for use in the forecasting process. The uncertainty in the data signals, especially wind speed, was found to be so high that use of the untreated data would not only no longer improve forecasting, but in fact deteriorate short-term forecasts — especially in storm events, when it is most needed.

There have been a number of initiatives in cooperation with the power industry in recent years on that topic, such as the Wind Forecasting Improvement Project (WFIP) [67] coordinated by the National centre of Oceanographic and Atmospheric Administration (NOAA) and the EWeLINE project [66], coordinated by the Fraunhofer Institute for Energy Economics and Energy System Technology (IEE) and the German weather service (DWD). In both projects NWP modellers, wind power forecasters and system operators were involved in order to target the research towards the usefulness of such data for the weather forecasts as well as the benefits for the power industry. Both projects concluded that forecast quality can only improve if the quality of the measurements towards which the forecasts are assimilated are of a high quality, meaning that instruments are chosen that are suitable for the task and that the instrumentation undergo regular maintenance to be trustworthy. In fact, it has been identified in both projects that if there is no specific effort put into standardisation of requirements in the power industry, the benefits cannot be achieved [27,103].

A work package in the IEA Wind Task 36 Wind Power Forecasting [104] has been created in order to provide recommendations for standards regarding measurement quality control. A first IEA Wind recommended practice has been accepted by the executive committee of the IEA Wind and is going to be published in 2022 [103].

4.2.2. Correcting NWP using data-driven algorithms and wind measurements

It is difficult for mesoscale NWP models to consider microscale factors such as terrain and wakes, as the temporal and spatial resolution is limited. Therefore, many researchers have been working on using artificial intelligence algorithms and actual wind measurements to statistically downscale and correct the NWP results. We list currently popular methods below.

- Single-NWP correction method: to establish a mapping model between the NWP at a single location and wind measurements at the same location.
- Joint correction of NWPs at multiple locations: a multi-to-multi mapping structure and a deep-learning-based model for joint correction of multiple NWPs builds upon the above, using NWP results at and measurements at multiple locations. Results demonstrated NWPs adjusted by a joint correction strategy have better accuracy and spatial wind speed correlation, than those adjusted with single-location NWP correction [87].
- Correction considering features of wind speed time series: the above two correction strategies aim to revise the NWP at any 'independent time slot', i.e. with no consideration of the implicit (but real) connection among wind at consecutive time slots. A sequence-transfer-based algorithm was proposed to revise the NWP wind speed, where ML methods are used to derive a dynamic relationship between the measured wind speed at time t and $t + 1$ [105]. Considering the statistics and time-series features of wind speed measurements, a wind speed error correction model has been established based on gated recurrent unit neural networks [106], and was shown to improve NWP performance correction. The decomposition method also considers the features of wind speed time series and decomposes NWP wind speed time series into sub-sequences for feature extractions based on Principal Component Analysis (PCA). The extracted sub-sequences are then used to correct the NWP via training ML for the relationships between the NWP error and each sub-sequence [107].
- NWP correction based on weather classification: the 'weather classifier'—in the meteorological literature also known as analogues [108], analog ensemble (AnEn) [109,110] or 'Grosswetterlagen' [111,112]—has been proposed to divide NWP data into typical weather patterns. NWP correction models corresponding to different weather patterns or patter switching have shown improved performance under different weather conditions [113,114].

4.2.3. Using NWP ensemble prediction systems to handle uncertainty

Integrating NWP output that results from different model initialisation, configuration, and parameter choices (as well as different NWP running modes or even versions) is an important method to handle uncertainty in NWP output. In [20] the uncertainty of the forecasting chain was reviewed and analysed in depth for NWP ensembles and their differences in design, ranging from initial condition perturbations (category 1), stochastic and physics perturbations, to multi-scheme (category 2) and multi model (category 3) approaches. All these approaches have in common that they aim to describe the uncertainty of the weather conditions due to a lack of knowledge regarding the correct initial conditions of the atmosphere and the incorrectness or under-determination of the mathematical equations as a result of physical approximations.

4.2.4. Scenario classification and (Bayesian) model updating

Besides the classical ensemble method described in the previous Section 4.2.3, there are a number of hybrid methods that are in semi-operational mode or used for research purposes. For example, in [115], an optimised fuzzy system was proposed to evaluate the effectiveness of 12 NWP members covering three different horizontal resolutions and four different initialisations. Further, it was shown by [116] that the dimensionality of the NWP portfolios can be reduced by using e.g. Kernel Density Estimation, with salient features subsequently extracted. And, in [117] a Bayesian model was employed to make probabilistic prediction of the wind speed based on ensemble NWP.

Other examples of scenario generation from calibrated ensemble forecasts were applied by [118] using ensemble copula coupling and by [119] using a dual-ensemble copula-coupling approach. Here, the authors proposed to combine two types of information: (1) the structure of the original ensemble and (2) the auto-correlation error estimated from past data. The reason is that category 1 EPS (perturbation-based, see Section 4.2.3) provides the uncertainty in the medium range, where the spatial–temporal structures of the forecast uncertainty get lost in the statistical calibration of the ensemble forecasts for each (required) lead time and location for the short-time horizon (1–3 days). Even though non-parametric approaches such as the copula coupling with medium-range EPS or generated from deterministic forecasts [109] allow some kind of reconstruction of spatio-temporal joint probability distributions at a low computational cost, none of these methods can fully describe the forecast-dependent structures without the unrealistic assumption of a stationary error matrix [20]. In practice, such reconstructed probability distributions or scenarios only cover the known or typical 'extremes', not those that have a low predictability or are very rare (with long return rate). In other words, all statistical calibration and generation of uncertainty can only describe what has been observed in the past, i.e. what is part of the training data.

4.3. Improving data quality and representativeness

4.3.1. Padding and restoring the missing/abnormal historical data

If the data that is used in operation and maintenance is also applied in the model training, the model training can give in a misleading apparent relationship between wind speed and power — resulting in forecasting uncertainty for future events. Therefore it is necessary to separate the corresponding data from the training samples. At the same time, for planned maintenance or shutdown, the uncertainty can be easily eliminated by removing the predicted power of the wind turbine that will cease operation. For unexpected failures and shutdowns, the connection to the condition monitoring and failure warning is beneficial for the forecasting results and reduce forecast uncertainty, if corrected in a real-time manner.

Missing a small amount of wind or power data will affect the continuity of the time series, while missing a large amount of data in a period might even affect the entire data set. The following methods can help mitigate uncertainty of time-series data, be it for training purposes or for operational and maintenance purposes.

1. Interpolating and padding missing data.
   Data *imputation* (gap-filling) methods can be categorised into statistics-based and ML-based. The statistics-based methods interpolate using data from adjacent times, and/or (weighted) averages of data from adjacent wind turbines. This type of method is friendly and efficient, and suitable for situations where the missing set is not large. Methods based on machine learning find a spatio-temporal ('mapping') relationship between non-missing data in a certain time window, based on which the missing data can be simulated and restored. Widely-used machine learning algorithms for data imputation include neural networks [120, 121], random forest [122], and K-nearest neighbor (KNN) interpolation [123]. ML methods might have higher accuracy, but tend to be less efficient than statistical methods for imputation. For missing power data, the corresponding wind speeds can be converted into power based on the power curve.
2. Recognising abnormal wind turbine operation/conditions.
   In the actual operational environment, the conditions of wind turbines can be complex. Under some operational conditions, especially abnormal ones (e.g. wind turbine faults and curtailment), the actual power recorded via SCADA might deviate greatly from conventional power generation levels. Elimination of such abnormal data is the premise for obtaining a representative wind-to-power relation, which the power data restoration is based on. Scatter plots of wind speed versus power are commonly used to determine the operational conditions, if the operation logs or fault tags of wind turbines are unavailable. Three commonly-used methods are based (conditional) on power density, image recognition, and rules, respectively [75,124].
   - Applying simple rules can be an efficient way to avoid intricate algorithms or model training; e.g. ignoring ('zeroing') the reported power output for wind speeds below some threshold around cut-in. However, such rules sometimes require experience with setting criteria, due e.g. to dependence on wind turbine operational parameters such as pitch angle and generator rotational speed as well as the control system. According to the operation principles and conditions of the wind turbine in question, the range of operational parameters under different operating conditions, as shown in Fig. 8, can be set as criteria. Note that since the actual operational conditions of each wind turbine are different, the criteria corresponding to various operational conditions needs to be customised for each wind turbine, and possibly for different modes of operation as well.
   - The 'density-based' method determines the operational conditions based on the distribution of power density with wind speed, e.g. application of noise to spatial clustering (DBSCAN) [125–127].
   - The method based on image recognition regards the two-dimensional scatter plot with respect to wind speed and power as an image, and uses image segmentation to classify data categories belonging to different operational conditions [128].
3. Power data restoration.
   Once the abnormal operational conditions are detected and eliminated, the representative and normal power curve model can be fitted/learned and updated using clean wind speed and power data samples. The methods for such can be broken into parametric and non-parametric types. The parametric method fits the actual power curve as a mathematical equation, with fitting parameters [124]. This method is simple and efficient. Nevertheless, the fitting performance around the cut-in and cut-out wind speed might be poor due to the non-linear characteristics of the power curve. The non-parametric method trains an undefined (virtual) function to approximate the relationships between wind speed and power based on machine learning algorithms [124]. Due to the complex joint distribution of actual power and wind speed (as well as e.g. pitch angle), as shown in scatter plots of power curves like Fig. 8, the non-parametric method is generally better for data restoration than the parametric method.

Once the power data is restored, it can be converted from wind speed based on the fitted or ML-identified (trained) power curve. Although data restoration can introduce new uncertainties into the training samples, it improves the data continuity and tends to reduce the forecasting uncertainty.

4.3.2. Improving the representativeness of training samples

In order to reduce the prediction uncertainty caused by the representativeness of training samples and the diversity of weather, and enhance the model adaptability to specific external conditions, training sample selection is required to customise the modelling process and to update the WPF model. This may be done regularly or in real-time as external conditions change.

Training samples are the fundamental material for WPF modelling. The data characteristics of the training samples determine the mapping relationship expressed by the model. After determining the input variables of the WPF model, a reasonable training sample selection is essential to reduce the prediction uncertainty and improve the adaptability of the model. The key to the selection of training samples mainly includes two aspects:

1. **Representativeness of training samples.**
   In regards to future prediction scenarios, the representativeness of a given training sample is required to examine whether the data in the training sample can represent the weather and wind conditions at a specific prediction time. Generally, inclusion of more types of weather and wind turbine operational conditions in the training samples might imply *less* representativeness of the training samples to specific external conditions [129]. For shallow learning models, it is difficult to capture the representative features corresponding to the specific prediction points found in large-scale or complex data sets. In a case study presented by Yan [130], the prediction accuracy became worse as the duration of training samples increased. Fig. 9 shows an example of how WPF accuracy based on yearly samples is lower than that based on quarterly samples, and that based on quarterly-samples is worse than using monthly samples.

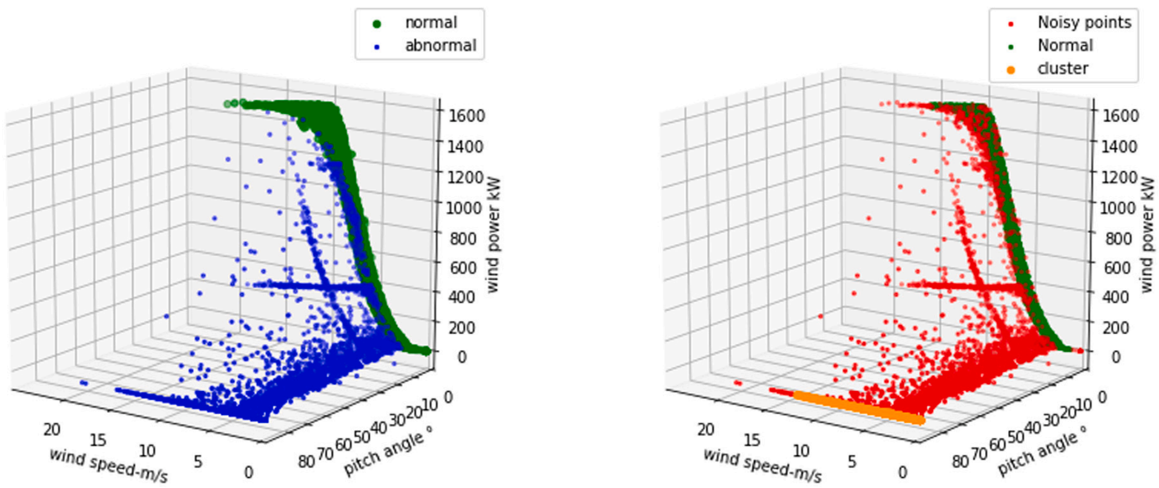


Fig. 8. Relations among wind speed, power, and pitch angle.

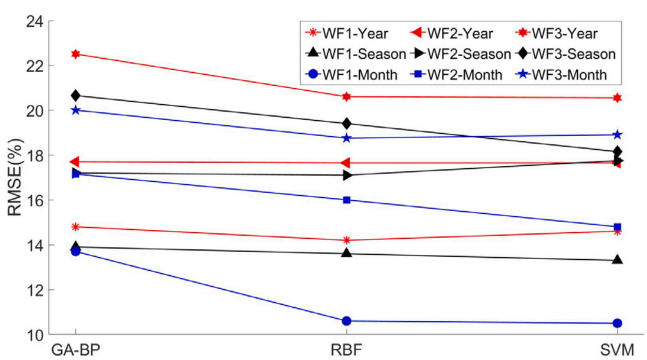


Fig. 9. Effects of time length of training samples on the WPF RMSE.

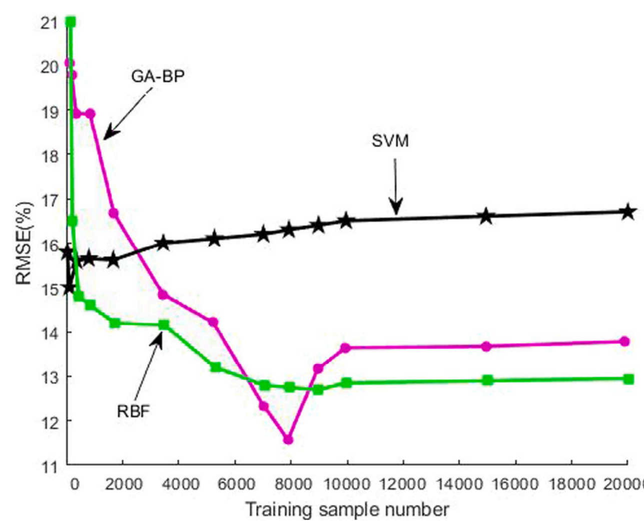


Fig. 10. Effects of training sample number on the WPF root-mean square error (RMSE) [130].

2. **Number of training samples.**
Insufficient training samples will lead to insufficient model training; however, an excess of training samples leads to over-fitting. As shown in Fig. 10 [130], there is no linear increase or decrease relationship between the number of training samples and the prediction error; the curves with respect to both variables have "sensitive changes" and "saturation" phenomena. Saturation means that no matter how the sample size changes, the RMSE hardly changes; while sensitive change indicates that the sample size is very sensitive to the RMSE. In addition, the optimal number of training samples for different WPF algorithms is different.

During the training sample selection, the representativeness of the training samples for future prediction scenarios, and the adaptability of the number of samples to the selected model, should be taken into consideration to obtain higher prediction accuracy and reduce uncertainty. Selecting the training samples according to seasons, weather, and wind fluctuation characteristics using a refined technical route is an effective means to *select representative samples*. There are three main methods of doing so, listed below.

- Select the training samples by seasons or months. For wind farms with obvious seasonal features, quarterly or monthly WPF model can reduce the predication uncertainty.
- Divide the training samples by weather categories based on parameterising the wind features, and set the valid ranges of each feature for different categories. The characteristic parameters may include: mean value, variance, (bulk) trend magnitude, maximum value, minimum value, frequency of occurrence per wind speed. This method is simple and easy, however, the weather category is difficult to distinguish in a rational way. Although the wind conditions near the threshold are not much different, they are instinctively divided into two categories, which may affect the prediction uncertainty.
- Use a clustering algorithm, such as k-means [131] or spectral clustering [132,133], to classify historical training data.

4.4. Improving data availability in real-time environments

The operation and maintenance process of wind turbines will inevitably involve abnormal power generation conditions, which increases the randomness of the actual power generation of the wind turbines, and in turn increases the uncertainty of power forecasting. For instance, routine turbine maintenance, unexpected faults, and unplanned or scheduled shutdown and curtailment fall under abnormal operation.

There are two common types of operational data anomalies. The first is observation distortion, which is common in wind speed or wind direction data, caused by anomalies in wind measurement equipment. The other is a constant data (or "dead data") problem caused by abnormal communication, which may occur with either wind or power data.

For data anomalies caused by wind measurement equipment, the 3σ criterion (assuming that data are within three standard deviations from a mean) can be used to detect abnormal data.[7] For instance, the average wind speed and variance of each wind turbine in the wind farm can be used as a baseline; or the joint distribution of wind speed of each wind turbine can also be established to detect the abnormal value exceeding the threshold. The "dead data" problem can be easily detected by searching a continuous but constant data according to simple searching rules. Once abnormal data is detected, the same imputation methods can be used as for filling in missing data, as described in Section 4.3 above.

4.5. Uncertainty of turbine models and power conversion

In addition to the uncertainty in the characterisation of the up-stream and downstream wind flow around turbines within a wind farm, another important source of uncertainty is a result of the design and simulation of wind turbines, which includes, but is not limited to the aerodynamic characteristics of the wind turbine rotor. Possible sources of uncertainty are found in the geometric properties of the blades as well as in the parameters of the air-foils, *i.e.* the aerodynamic coefficients of lift, drag and moment [134]. These coefficients in fact have direct impact on the estimated turbine performance, not only in terms of wind-to-power conversion through characterising the power coefficient (C_P), but also in terms of the interaction of the rotor with the wind flow; the latter causes velocity deficit downstream through wake effects, which are typically modelled by a thrust coefficient C_T.

The air-foil aerodynamic coefficients can be obtained by experimental and numerical techniques, both of which are challenging and lead to uncertainties of both aleatory and epistemic nature [134]. Additionally, the rotor blades accumulate dust, dirt, insects and pollen over time on their surfaces, which are also affected by weather-driven damage such as peeling and erosion, typically prominent at the leading edge. In fact, annual energy production (AEP) losses due to leading-edge erosion can be up to 4.9% in operating wind farms [135]. These changes in surface conditions during operation can result in significant uncertainties in power capture and affect forecasts directly in terms of power conversion, and indirectly in terms of the wakes generated [134].

To at least partially mitigate such uncertainties, data-augmented methods for site-specific power curves (including probabilistic power curves [75,124]) and calibrated modelling are increasingly being adopted. At a single turbine level, an example where modern deep learning approaches are deployed to mitigate the C_P (and in general turbine model) uncertainties for available power estimation and forecasting is presented in [134].

5. Uncertainty evaluation and mitigation methods in the market phase

In the market phase we deal with different types of uncertainties. Traditionally, the dimension and generating capacity of the electricity grid are limited by the instant failure of the largest power plant on the grid which at any time may need to be balanced with fast-reacting reserve power. This process has been gradually undermined by renewable power sources, namely wind and solar. The latter have effectively shifted power generation from large-plant setups near load centres, to highly dispersed smaller-scale power generation units which are often distributed across low-population areas far away from the load centres. The traditional security margin had been directly related to the generation uncertainty due to unforeseen failures of conventional power plants.

[7] Assuming the data does not have a 'fat-tailed' distribution with significant occurrence of events beyond 3σ from the mean.

Since the large-scale integration of wind (and solar) has started, the uncertainty in power generation has received a new aspect that is related to the intermittent generation pattern of such renewable resources. This type of uncertainty – highly connected to the weather conditions that fuel the renewable generation – is far smaller, but is driven by situations which occur far more often than conventional plant failure. It can be considered as a kind of 'background noise' that requires a transformation of grid operation.

Additionally, political aspects are also responsible for a portion of uncertainty related to wind (and solar) power forecasting. The following sections will illustrate in more detail how forecasting is affected by these uncertainties, and to some extent provide recommendations to mitigate some of the uncertainty in the forecast process and training.

5.1. Transparency-related uncertainty

In the European Commissions (EC) Regulation No. 543 from 2013 [136], it is stated that Transmission System Operators (TSOs) are "...required to publish data on the availability of networks, capacities of cross-border inter-connectors and generation, load and network outages as well as insider information" in accordance with EC Regulation No. 714/2009—in other words, all relevant market data. In the regulation [136] it says that "...the availability of such data is indispensable for market participants' ability to make decisions on efficient production, consumption and trading". The regulation also states that "...deeper market integration and the rapid development of intermittent renewable energy generation sources (such as wind and solar) requires disclosure of complete, timely available, high-quality and easily digestible information relating to supply and demand fundamentals. The timely availability of complete sets of data on fundamentals should also increase the security of energy supplies" (see also [137]). It should allow balance-responsible parties (BRPs) to precisely match supply and demand, to reduce the risk of blackouts. As a result, TSOs should be able to better control their networks and operate them under more predictable and secure conditions.

The information market participants should be provided with – which has the potential of increasing WPF uncertainty and subsequently uninformed decisions – can be summarised as:

- expected consumption;
- planned and unplanned unavailability of power generation and consumption units, with detailed information on where/when/why units will not be available to generate or consume, and when they are expected to return to operation [136];
- detailed information about the overall installed generation capacity, estimations about total scheduled generation, accounting separately for intermittent generation, and unit-level data about actual generation of larger production facilities [136];
- planned and unplanned unavailability of existing cross-border transmission infrastructure and plans about infrastructure developments, in order to move power from where it is available to where it is most needed and adjust portfolios accordingly [138];
- regular updates on planned and offered cross-border transfer capacities for different time horizons as well as information related to the allocation and use of capacities [136].

If these types of information are not made available and are not of high quality, the uncertainty of the production will increase together with the risk of failure or high balancing costs. The transition from few organisations and a limited number of generation plants, to large numbers of players and generation units, has shown to be a challenge for the electric grid operation and development. In a study from 2018, Hirth et al. made a review of Europe's most ambitious electricity data platform, the "ENTSO-E Transparency Platform" [139]. The authors listed the lack of announcement of data availability (i.e. which kind of data is made available) and the lack of accessible documentation as important issues of the information platform. For example, they indicated the lack of information about available capacity underlying the estimated generation, both in terms of forecasts and actual production; this directly relates to the second, third and forth items in the transparency requirement list above.

Germany is an example of where the partial privatisation of power trading for wind and solar has led to a situation where crucial information is lacking: the available capacity underlying the production data provided to the transparency platform is unknown, as is the quality of the up-scaled generation. Worse yet is the fact that the parties providing these data to the system operators could potentially have conflicts of interest, as they also provide forecasts to market participants (personal communication with the TSOs Amprion and 50Hertz). The German example is a case of problems arising from a lack of transparency in the market regarding production data, as this also influences the market, if the predictions for the expected generation from wind power may be contaminated by conflicts of interest. In other areas, such as the UK, where BRPs are accountable for the balancing of wind and solar power, it was also found that the quality of the predictions is insufficient for system operation. This can then only be resolved, if the system operator gets access to the real-time generation data from the power plants. This is again a matter of policy.

The effect of insufficient transparency on the uncertainty of the forecasting process can be summarised as follows:

- missing input data to the forecasting model;
- missing information from real-time and historic metered production;
- missing meta- or standing data[8];
- missing planned and unplanned outage information;
- missing available active power information of the power plants.

The overall effects on the power system and market operation can then be summarised as:

- risk of conflict between market and system operation;
- increased reserve requirement due to a lack of knowledge of real-time production;
- increased congestion due to a lack of transparency in policy e.g. storage.

5.2. Uncertainty from political decisions on capacity and infrastructure

The energy system is widely driven and impacted by political decisions. Decisions made in one country/region may have influence on other (neighbouring) countries/regions with respect to both capacity and infrastructure. In most countries/regions, wind farms are constructed where there is the highest wind potential. This can easily lead to a skewed geographical distribution that increases the uncertainty of both the generation itself and the forecasting in comparison to a well distributed power generation, where forecast errors often balance out geographically. Policies should therefore incorporate the geographical distribution in the integration phase to avoid high errors in the forecasting process and hence increased reserve requirement in the transition phases [140].

Defining requirements for wind farms, in regards to measurements or additional information for uncertainty reduction, is ideally complemented by incentives that are aligned with policies. Such incentives are essential, because they can increase the generation and finance the additional cost for wind farms to provide more information to a centralised forecasting system of the system operator.

From interviews in Denmark conducted as part of the IEA Wind Task 36 "Wind Power Forecasting" [18], it was found that the political decision process is considered to be a major risk factor for investment in new technologies: it is *not the market nor the operational aspects*. Due to the globalised economy, this result is not only relevant in Denmark, but world-wide.

5.2.1. Uncertainty in a high wind penetration area

The forecast uncertainty changes in areas where wind power penetration increases above approximately 30% of the demand [141]. In areas such as Denmark, Germany, Spain and Ireland, it is observed that the average width of the forecast bands (i.e., uncertainty) grows more strongly with increase in capacity, compared to the reductions in forecasting error [142]. In areas where there are power markets, such increases in uncertainty imply higher volatility, and result in potential increase of intra-day trading volume over time [139].

In such high-penetration scenarios, one way to mitigate uncertainty is to split the forecasting process into different components [25,26,143]. The most common components are: (1) smoothed out day-ahead forecasts through training or by using multiple forecasts to dampen large errors; (2) variable inta-day forecasts with adjustments from measurements to balance day-ahead errors; (3) highly variable minute-scale forecasts for the physical balancing on system operation level or after gate closure.

In most electricity markets such as the European Union, USA, Canada, Australia [139,142], the first two forecast components are common practice, also in non electricity market areas such as China, Japan and India. The last component is observed in high penetration areas, where balancing power is getting short and where wind power generation levels on system generation of 65% or more are applied [144]. These minute- or hour-scale forecasts are designed to bridge the gap between the current measurements and intra-day short-term (ST) forecasts, which are still significantly smoother than the actual generation. Such forecasts are mostly dependent on the availability and quality of the wind farm data. As the forecast horizon and frequency shorten, the need for accuracy and reliability in both forecasts and measurements increases. This is due to the fact that the time from forecast generation to deployment is much shorter, and consequently the forecast process must be nearly as reliable as e.g. the primary reserve itself [65].

5.2.2. Uncertainty from commercialised balancing of wind power

In most of Europe the balancing of wind power and other renewables is done by balance responsible parties (BRP) and not the system operator. This works especially well for such well-interconnected areas. In less interconnected areas or island grids, it is more difficult and therefore balancing is still done mostly by system operators.

A key reason behind such a 'liberalised' strategy is the desire to make the market more competitive with more actors. Denmark and Germany were pioneers in the introduction and could provide a proof of concept for reduced spot market prices in Nordpool and the German–Austrian part of European Power Exchange (EPEX) in the early years of the millennium. However, the consumer prices increased, because the total costs of running the system with a high share of wind power increased for various reasons. The time horizon is one of the main differences between the optimisation strategies of system operators and traders/BRPs. The trader is typically interested in short time horizons of maybe one month ahead, where the system operator is concerned with more than a year ahead. The shorter-sighted interest of traders could, in the long run, be a disadvantage. For example, transparency becomes an issue, if the distribution of generation data of the power plants are not regulated and the trader has more or more timely information about the generation state of turbines. This is the case in Germany, where the dependency of system operation processes to commercially sensitive data is leading to a lack of efficiency in the handling of wind and solar power [139,145].

[8] Here defined as the non-changeable data of a wind project, such as geographical data of the turbines, turbine type, amount of turbines, etc.

5.2.3. Uncertainty arising from balancing via interconnections

In some areas, such as Denmark and Germany, inter-connectors are deployed to handle renewable energy via market coupling, and special regulations apply to moderate the power flow through the inter-connectors (e.g. in cases of internal congestion). The inter-connectors help to reduce price volatility due to uncertainties. An unresolved challenge remains that it often leads to negative spot market prices during times of spatially correlated high generation [141], during night times and in weekends, when there is low demand and high wind speed [146–148].

5.3. Uncertainty resulting from future application of storage and flexible demand

Storage is a competitive balancing solution in systems with high price volatility, or where there are generation limits on the grid. Both conditions are widely present in Germany for example, which is a target area for storage in the coming decade. The market price of storage solutions will be driven by the demand, and may be an obstacle for deployment in regions where wind power generation is most cost effective. If the potential for heat pumps and electric cars is high and provides some opportunities to better utilise variable generation, such technologies are in direct competition to energy storage and can therefore become obstacles for development of large-scale storage – even though these technologies are limited in their capacity, and probably will be (in the best case) only capable of providing instantaneous, high frequency balancing tools for primary or secondary reserves [149]. With that in mind, there are still significant differences between grids with high share of wind at high levels of interconnection compared to other control zones and grids with high share of wind at moderate interconnection levels. The latter will naturally develop price volatility due to uncertainties in the production with reduced ability to export over-capacities unless (cascading) storage systems will be developed [150].

For the forecasting process, storage can increase forecast uncertainty significantly, unless the electricity from the storage is traded on the market or wind power plants are balanced intelligently (e.g. via scheduled ‘virtual power plant’ by the storage operator).

Battery storage systems for household prosumers have recently been studied by [151] with respect to the design of retail prices, grid fees and levies. There the attractiveness to prosumers of investing in household storage was evaluated, along with the impact on system-oriented grid and market operation. In particular they found that market- and system-oriented operation are not necessarily in harmony, and at times even contradict each other; they recommended caution with the term “system-oriented operation” [151]. In this regard, system-oriented operation is linked to the transparency dilemma described above (Section 5.1), if the system operator lacks control of the physical effects of power production and consumption. Accordingly, the uncertainty in forecasting can increase significantly and incur large costs. Germany is a classic showcase for this effect, where small PV systems began to include battery storage to avoid paying taxes and levies on consumed electricity. With a small percentage of such systems in the overall grid infrastructure, this is just lowering the negative consumption for the system operator. However, when the percentage increases and the consumption of e.g. a million households stops following general patterns, but instead is skewed due to available battery capacities, problems can arise in the grid operation; especially in extreme cases, where there is little power available and a sudden increase in consumption or vice versa. A recent example of such an event happened in Texas for the Electric Reliability Council of Texas (ERCOT) causing large scale black-outs. A storm with extraordinary cold weather (well below freezing) caused the consumption to rise exponentially, while power plants were not able to become available and wind turbines were shut-off due to icing [152,153]. This can happen anywhere with large amounts of households that use heat pumps with or without battery storage.

Such situations are good examples to understand why reserve forecasts need to be based on physical ensembles, when it comes to extremes. In the ERCOT case with temperatures dropping below the freezing point and heat pump efficiency dropping to ~50% of normal operation, the amount of electricity requested from the system increased dramatically. Today, events like the January 2021 ERCOT case [154] are predicted well in advance by the weather services. However, if the weather information’s uncertainty is not moved forward in the chain of applications, the risk for damages and failures increases dramatically [155].

To conclude, the uncertainty chain from weather development to weather-dependent application needs to be followed and handled correctly per the forecasting chain — and must be considered in market design as well as system operation best-practices.

5.4. Uncertainty in price forecasting

Price forecasting is not directly connected to the weather-related forecasting of wind power, but is a driver of the market mechanisms and should therefore be mentioned here as well. Ketterer [156] explained the impact of wind power generation on electricity prices, and provided a substantial amount of references for further study of price forecasting. In her study, the effect of wind electricity feed-in on level and volatility of the electricity price was evaluated, and the results confirmed that variable wind power reduces the price level, but increases its volatility, as found an earlier investigation [157]. This in turn means that the uncertainty in the profitability of electricity plants, conventional or renewable, increases [156].

6. Advanced prediction methods to mitigate uncertainty

6.1. Using advanced AI algorithms

The use of artificial intelligence techniques is increasingly popular in wind (or in general renewable) power forecasting. It is also commonly referred to as intelligent forecasting, for which the common methodologies are exemplified in Section 2.2.1. Intelligent methods or models endeavour to learn and fit underlying relationships within a set of historical weather inputs and wind power output as a “black box”. Algorithms with deep network layers, or employing advanced learning mechanisms, are expected to possess better learning ability during model training. For predictions, their accuracy and associated uncertainties ultimately depend on their generalisation ability regarding inputs beyond the training domain.

6.1.1. About shallow machine learning algorithms

So-called gradient boosting machines (GBM) are one of the most commonly used shallow learning models, including the eXtreme gradient boosting (Xgboost) [158], light gradient boosting machine (LightGBM) [159], and categorical boosting (Catboost) [160] algorithms. GBM stands out in its capability of selecting driving features from rich (multidimensional) input domains. Moreover, the GBM algorithm can achieve very efficient and good mapping performance between input and target values.

In the shallow learning field, model integration technology is also a promising research area, which includes stacking, bagging, and boosting methods. Stacking is a popular algorithm in many international prediction competitions. It is generally used to integrate multiple layers and multiple base predictors into one WPF model after meta learner to optimise the weights of all base predictors. Integrating results from different predictors (better with small deviations but large variance) can improve the robustness of the forecasting results and reduce its associated uncertainties.

6.1.2. About deep learning algorithms

Deep learning represents the frontier of machine learning in the past decade. A large number of advanced deep learning algorithms have been applied to WPF. From an algorithmic perspective, generative adversarial network (GAN), attention mechanism and graph convolution algorithms have great prospects for application. In addition, the spatiotemporal prediction architecture brings new ideas to future WPF research; this and the three other DL algorithms can be summarised as follows.

The GAN algorithm can support scenario forecasting with large-scale spatiotemporal data by mapping the joint distribution of multiple wind farms and multiple time slots [161–163].

The attention mechanism can better learn from time-series information in short-term and ultra-short-term prediction, mainly focusing on variable attention mechanism and temporal attention mechanism [164,165].

The graph convolution network (GCN) expands the convolution technology to irregular image data so that a complex spatiotemporal distribution can be displayed and quantified in the WPF algorithm. For example, the combination of GCN with RNN can better learn the temporal and spatial dependencies among wind farms and improve the forecast accuracy [166].

The deep-learning-based spatiotemporal prediction architecture enables the joint forecasting for wind farm clusters in an efficient manner. Examples include Spatial–Temporal Graph Convolutional Network (STGCN) [167], Diffusion Convolutional Recurrent Neural Network (DCRNN) [168], Adaptive Graph Convolutional Recurrent Network (AGCRN) [168], Graph WaveNet [169], and Graph Multi-Attention Network (GMAN) [170].

6.2. Uncertainty in ramp forecasting

The quantification of uncertainty in ramping forecasts is a challenge that starts with the correct definition of a ramp event. In a survey on wind power ramp forecasting, Ferreir [171] and Ouyang [172] presented four different definitions for ramps in wind power forecasting (WPF), where the first three definitions mainly consider the change of amplitude, i.e. a ramp occurs when wind power amplitude exceeds a predefined threshold value in a certain interval of time and the last definition uses the wind power rate to indicate a ramp [172]. They both concluded that each ramp definitions takes a different aspect into consideration, and will have different applications as well as advantages and disadvantages. In fact, the uncertainty of the definition of a ramp event in itself is the subject of the bulk of current literature on the topic (e.g. [171–175]), as almost every application has the potential to have its own definition. This makes comparisons of methods quite difficult, and adds a level of uncertainty to the applicability of methods when used in new or other circumstances.

There are a number of meteorological characteristics that have been described in literature [20,65,171,175]. In [173] four of the main meteorological characteristics that generate uncertainty and large variations in wind speeds that lead to wind power ramping are described:

1. Large-scale weather system passages
2. Local or meso-scale circulations (sea–land breezes, mountain–valley winds, drainage and gap flows)
3. Vertical mixing of momentum (“Dry Convection”)
4. Thunderstorms (“Moist Convection”)

[176,177] explores detection of wind speed ramps and their rate of occurrence, while [178] finds long-term joint-PDFs for the power- and load-driving characteristics of offshore ramps. The characteristics listed above are similarly also applicable to solar power ramping [179], where it is the clouds coming with fronts and moist convection that are the main generators of uncertainty and strong ramping associated with items one, two, and four in the above list [180,181].

The non-meteorological characteristics that impact the WPF uncertainty around ramping events can be described by (e.g. [27]):

1. effects of wind/solar power clusters and capacity gradients;
2. local effects related to the terrain (especially complex terrain);
3. grid topology;
4. penetration level and local demand.

These characteristics can have little to no effect, but also large effects on the uncertainty and uncertainty quantification of ramping events. Dependent on how much power an electric grid can absorb, critical ramping rates deviate strongly. Also, the dispersion level of the generating units can add to the uncertainty of ramping events, similar to the terrain characteristics adding complexity to the forecasting process and the associated uncertainty.

Hirata et al. [181] investigated causes of uncertainty from applying artificial intelligence algorithms to the problem and defined five causes of uncertainty:

1. few similar events available in the past;
2. the space spanned by neighbouring events does not behave well and they are not linearly independent;
3. the underlying system is sensitive to initial conditions;
4. the underlying system is influenced by stochastic noise;
5. the underlying system is about to change qualitatively.

The authors [181] mention that “..the first two are related to the properties of the time series prediction with a neural network approach”, which is a known phenomenon (see also 4.2.3) when using statistical methods to quantify uncertainty in cases of extreme events or rare wind speeds within a certain time period. In [20] the different methods for probabilistic forecasting and their ability to quantify the actual uncertainty are described in detail. The other three uncertainty sources originate from the properties of the underlying dynamics of the approaches and their inputs. The latter three properties are difficult to quantify, unless there is detailed knowledge about the time series origins, some of which are described in Section 4.2.3.

6.2.1. Definition of ramp forecast uncertainty

In the renewable power arena, the objective of ramp forecasting is to forecast the change in power generation between two time stamps relative to the schedule, which is defined by a wind power or solar power forecast (hereafter referred to “variable generation” abbreviated as VG). Using a deterministic forecast process for the variable generation schedule implies that the corresponding deterministic ramp forecast uncertainty is always zero. A deterministic forecast, therefore, needs to be complemented with past statistical results to produce some kind of a ramp result.

In [174] it was found that “..a single time series of wind power forecasts may not include enough information to make secure management decisions related to the potential occurrence of a ramp”. They also studied “..the extent to which NWP ensembles provide information on the forecasting of ramps, including the associated uncertainty” [174]. By applying clustering methods and regression models, they also showed that even the category 2 (see 4.2.3) ensemble forecasts were significantly better at predicting ramp events, compared to considering a single, deterministic wind power scenario or climatology.

In [23,182], ramp forecast scenarios were generated by blending forecasts from four machine learning techniques, to: (1) obtain wind power forecasts; (2) find historical forecasting uncertainty; (3) fit the probability distribution function of the forecasts with the help of a Gaussian mixture model (GMM); and (4) use an inverse transform method based on Monte Carlo sampling for the development of forecast error scenarios. They further underline the need to improve probabilistic forecasting of critical wind power ramps.

In [183] forecast uncertainty is estimated using multiple NWP inputs, statistical processing and adaptive algorithms, where they concluded that “..the temporal forecast uncertainty can be quantified and presented to indicate the likely timing and amplitude of wind energy ramp events”.

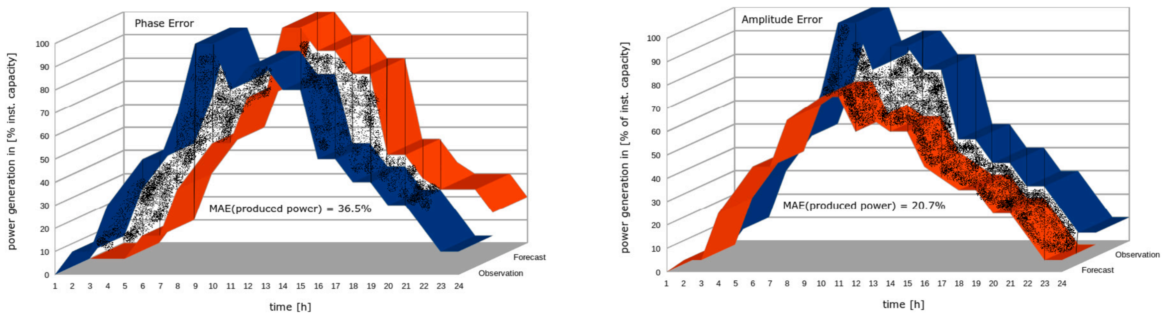


Fig. 11. Schematic of the uncertainty in terms of forecast errors in phase (left) and amplitude (right); here specifically for a ramping event. The dotted areas indicate the area of the forecast error. Here the forecast is the red curve, the observations are shown with the blue curve.

Applying a physical weather ensemble forecasting process generates high degrees of freedom, but it is to date the only proven method that is capable of generating the uncertainty associated with ramping events [27,174,184]. The schedule for power production generated by an ensemble forecasting approach is typically a soft curve derived from many ensemble members. This could be either a median, the average or a percentile (e.g. P50) of the ensemble members. In this way each ensemble member provides a ramp result relative to the schedule. In a low-uncertainty event the ramp values are close to zero and in a high-uncertainty event the ramp values increase in magnitude, both upwards and downwards. The longer the ramp lasts, the more accurate the ramp value is from each ensemble member, because the longer time horizons filter out some of the phase errors. This means that the uncertainty of an 8 h ramp allocation is normally less than the 3 h horizon, counted per hour of allocation [27].

6.2.2. Quantifying the uncertainty of ramp forecasts

When quantifying the uncertainty of ramp forecasts we often think of the predicted amplitude and timing of the ramping event. When evaluating ramp forecasts, it is therefore necessary to not only distinguish between phase and amplitude of forecast error, but also the possible timing of significant events [183]. In general statistics with mean or mean squared errors, which are common evaluation metrics in the power industry, the amplitude and phase are differently weighted. A description of how errors in predicted ramp timing (‘phase errors’) are penalised more than the amplitude errors in ramp forecasts is provided by [143], including the lack of steepness. Forecast phase errors are subject to a so-called “double punishment” (sometimes referred to as “double penalty”), where the forecast performance is penalised twice: once for not forecasting the peak value (in time) and once for generating it too late. These type of error measures are now considered to be unsuitable for ramp forecast evaluation (see also [171,172,185]). Fig. 11 gives a schematic of the double punishment/penalty, illustrating both phase and amplitude errors. When considering the uncertainty inherent in wind ramps, hence in power forecasts, the end-user must take this asymmetry into account if the amplitude of ramp events are critical. This can be done using e.g. contingency or skill score tables (which classify forecasts with hits, misses and false alarms instead of mean error statistics), where there is an incentive for the forecast provider to put a larger weight on phase and relax on the correct amplitude [183,186,187]. The same applies for timing errors or the phase of the ramp. In [183] ramp events are presented as temporal distributions in a pre-defined period around the ramp event rather than at one specific point in time. In that way, the forecast uncertainty can be mitigated.

6.2.3. Mitigation of uncertainty for critical ramps

From a power system perspective, it has been observed that the security constraints from a system operator can conflict with optimal power production from wind and solar power [181,188]. In [181] the causes for uncertainty in prediction of critical ramps are investigated

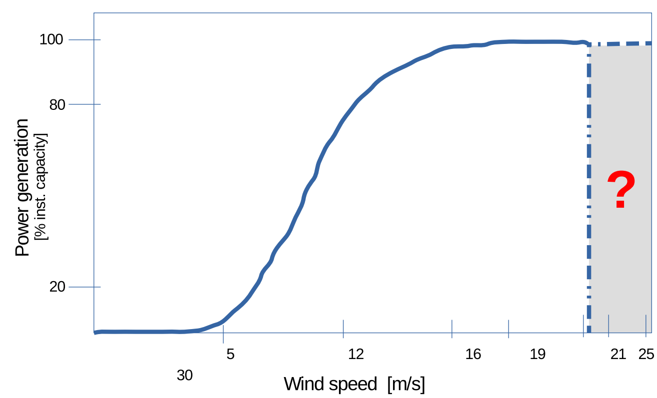


Fig. 12. Schematic of the uncertainty range of the cut-out wind and associated power drop.

and quantified in order to predict the likelihood of such rapid changes in wind speed, which cause critical situations with machine learning techniques. The aim of the quantification is then to find mitigating measures for the impact of such dangerous fluctuations. [181] also attempted to produce ramp predictions based on a set of credibility measures that quantified the associated causes of uncertainty, which remain unsolved.

6.3. Uncertainty quantification in cut-out wind speeds & high-speed shut-down

Wind turbines under normal conditions generate full (rated) power when wind speeds reach rated speed of approximately 12–15 m/s and up to where the wind farm’s high-speed shutdown (HSSD) set point is activated. One of the difficulties in quantifying the uncertainty of cut-out wind speeds is the fact that in modern wind turbines, the latter is no longer a specific point or wind speed, but a range, usually starting at about 21 m/s up to 30 m/s, where the wind turbine gradually pitches the blades out of the wind to reduce loads and consequently long-term damage. Thus from a forecasting perspective, *both* wind and power measurements are needed together to quantify the uncertainty and thereby become able to calibrate forecasts for HSSD. In Fig. 12 the uncertainty of HSSD signals in the wind range above 21 m/s is illustrated. The grey area indicates the broad range from 21 m/s up to 30 m/s where HSSD can occur. In advance of a HSSD event, there is no indication from time averaged power generation signals from wind farms to warn of such an event. In other words, HSSD warnings require wind speed forecasts and signals to be useful for forecasting. It is necessary for a HSSD forecast application to learn from locally measured wind speeds around the HSSD set-point of a wind farm. Although there may be few HSSD events to train the HSSD model, more detailed training of forecasted wind speeds in the more frequently

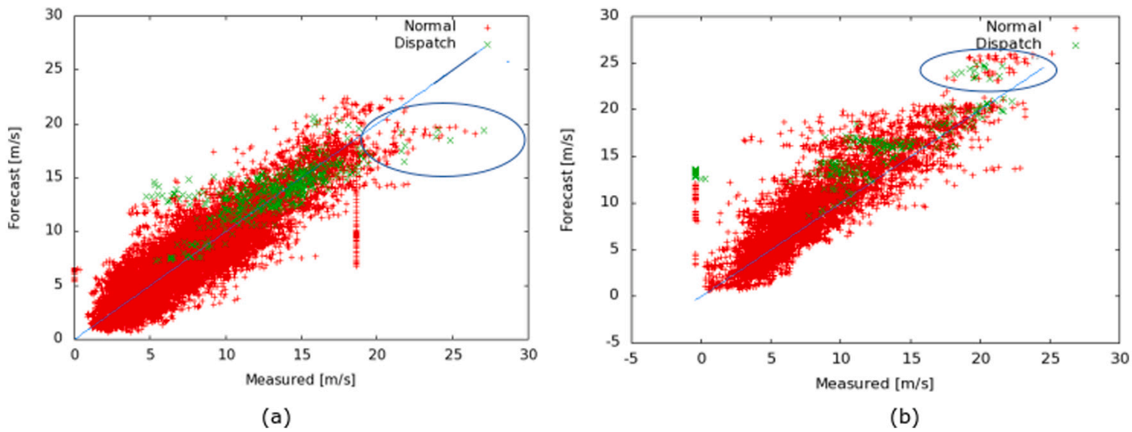


Fig. 13. Scatter plot of measured against forecasted wind speed at two wind farms where (a) the measured wind is much higher in the HSSD range than the forecasted and vice versa (b) the wind speed much higher than what is measured at the wind farm's met mast.

occurring 20–25 m/s range can reduce the uncertainty and increase forecast quality, also in the wind speed rage 25–30 m/s which is where the HSSD typically occurs.

This also points to the requirement of high-quality measurements (see 2.1.4) at the wind farms in areas, where high wind speed above 25 m/s is common. In cases where there are no online wind speed measurements, there exists always an option to use forecasted wind speed to train HSSD signals in a forecast application. An example is shown in Fig. 13, where the observed forecast errors are randomly varying by about 1–2 m/s. This would not be the case with observed wind speeds and should therefore be applied for the power curve generation in the high wind speed range, if available.

6.3.1. Use case: Mitigation of uncertainty in high-speed shut-down

In [142] a case study is described with a "[...] real-time example of two concurrent HSSD events in Ireland, in which the forecast system issued warnings when the probability for an HSSD affecting more than 20% of the wind generation capacity exceeded 25%". In the establishment of the warning system, the service provider worked closely with the system operator in order to develop a warning strategy that fit their needs without being overwhelming. Since Ireland is a country that is exposed to strong winds from the Atlantic, there are many high-speed events throughout the autumn and winter seasons.

A warning system must provide a reliable and timely warning (see also [183]) without too many false alarms, to avoid undermining the seriousness of the warnings. This has been found to be a major challenge in such systems. In the example from Ireland [142], forecast lead time, change of severity level since previous alert, initial day of the week, valid day of the week, time of day, severity of the event, actions required, as well as need and ability to call back actions, were carefully configured.

In the Irish case [142], the alert strategy had two main categories:

1. A well-defined scheme is used for the first warning.
   The scheme alerts, when the probability of an event reaching 20% or more of the running capacity exceeds 25% for more than three consecutive forecast cycles. This still allows a HSSD event to be alerted 6 days in advance.
2. Limited number of alert.s
   Alerts are only issued when the threshold is fulfilled, and not updated until a significant change is observed; e.g. the probability changes by more than 10% or the affected capacity is reduced by 10% or more.

Such an alert system prevents the system operator from acting too early or too late, allowing preparation time in case critical grid situations may occur due to such an event. By providing an alert, the uncertainty of the forecast cannot be reduced, but the costs of (in)action in terms of redundancies can be optimised to a large degree. Strategic preparation is possible by optimising the alert system to specific targets, thereby reducing the risk of failure of the system while also reducing costs.

From a human decision perspective, it has been found (see e.g. [68, 142,189] that "[...] providing probabilistic forecasts and measurements in both graphical and tabular formats is important to the end-users, as humans react differently to graphics and tabular information" [190]. So-called *weather literacy* [191], i.e. the ability to understand weather risk and be able to interpret probabilities versus deterministic forecasts and their inherent uncertainty, is an important aspect for humans to make rational decisions. In order to accommodate such differences, it is important to indicate if a forecast was/is not correct: both to strengthen the confidence in the forecast information, and to assist such operators to take the best possible actions without wasting resources. In this respect, we note that not all probabilistic forecasting methods are useful. In the above described Irish HSSD case a physically based ensemble forecast is required in order to capture events that may not be present in historical training data used by the statistical methods. The current state-of-the-art methods and their applicability have been described in e.g. [20,142,189], and are a focus of Task 36 (from 2022: Task 51) within the Technology Collaboration Programme for Wind by the International Energy Agency (IEA Wind TCP) [104].

6.3.2. Uncertainty of cut-in wind speeds

Cut-in wind speeds can become a risk factor after a high-speed cut-off event, when wind turbines are not connected to an automatic SCADA system and have to be started automatically. These risk factors have been most pronounced where wind turbines lack automatic control systems. In countries like Denmark, Germany, and Ireland, turbines made before 2000 often had no automatic control system; thus extreme events caused a lot of uncertainty in the forecasting process after frontal passages, where the wind dropped again and generation should have started. In Ireland this amounted to a significant fraction of installed capacity, significantly enough that expensive reserve power needed to be provided at times (personal communication with the Irish system operator EIRGRID). The mitigation measure for this issue was to incorporate a high-speed shutdown forecasting tool that also is trained with existing data from past events to ensure that there was enough reserve allocated when high-speed shutdown was about to come to an end [142].

6.4. Uncertainty quantification requirement for reserve forecasting

To maintain system stability, one of the ancillary services required to be provided by the wind power plants is reserve power, typically performed via curtailing the wind farm from its maximum available power [192,193]. For mandatory down-regulation (or curtailment) requested by the Transmission System Operators (TSOs), the reserves can be used to estimate the compensation made for the loss in production. The reserves can also be traded in the balancing markets. Due to their economic value and significance, the assessment of the accuracy of the reserves (or indirectly the available power) is crucial and often regulated. The requirements for qualification of reserve power provision and potential compensation are regional, i.e. market and legislation dependent. According to the European Network of Transmission System Operators for Electricity (ENTSO-E) policies, the quality assurance of the reserves are under TSOs responsibility within continental Europe [194]. A brief summary for regulations in Europe with offshore wind power in the grid, is presented in [192] for 2016. Shortly after that, Germany introduced much stricter regulations for the compensation during mandatory down-regulation [195], stating that “the available power is to be estimated/forecasted for 60-s intervals for down-regulated wind farms”. Additionally, “1-min standard deviation of the percentage error of the available power is required to be less than 3.3% (after the pilot phase)”. Göçmen et al. [192] therefore concludes that “[...] the enforced regulations are difficult to comply with, and are subject to penalty if not met” underlining the importance of uncertainty quantification for (short-term) reserve forecasting.

In Canada, [196] showed that the characteristics of the wind power forecast error distribution is the main contributor to the level of risk, which is also a function of the wind production level. Accordingly, their results suggest that (temporal and spatial) local wind conditions are a highly important factor for determination and management of risk, as well as the estimation of the required level of the balancing reserves within the system. Although this seemed to be counter-intuitive, it was found that for certain lead times, the overestimation of the anticipated generation can serve as a virtual balancing reserve, if it can be predicted as such. The results also revealed the importance of forecast error classification, for balancing reserve calculations — independent of the market structure. Their conclusion was that in addition to the standard deviation of the wind power forecast errors, the potential systematic biases introduced at each production level should also be accounted for when estimating and exchanging balancing reserves [196].

6.5. Assessment of uncertainty in the dynamic modelling of turbines

As stated earlier, the (available) power forecasting is very sensitive to the turbine models considered and their operation strategy. Physics-based approaches in particular rely on a pre-calculated power coefficient (C_P) or nominal/optimum power curve for converting wind speed forecasts to power. However, local (spatially and temporally) aspects of the meteorology (e.g. humidity, temperature), flow (e.g. turbulence), turbine characteristics (e.g. control response [197], and condition of the blades [198]) affect the C_P and power curve behaviour. These deterministic parameters lack the required dynamic representation of turbine and wind farm performance, and are subsequently important sources of uncertainty [199]. [134] presents a ‘model-free’ approach to convert wind speed to power as an attractive alternative in order to eliminate the inadequacy of turbine models to represent the dynamic power output of a turbine under turbulent inflow.

6.5.1. Use case: Mitigation of uncertainty in turbine power conversion (Wind speed to power)

In [134], a case study is described to address and reduce turbine model inadequacies and other uncertainties for reserve power forecasting, where recurrent neural networks, via long-short-term memory (LSTM) neurons, as a single-input, turbine model-free approach for 1 Hz (available) power forecasts are utilised. The changing inflow conditions and turbine control settings are captured via transfer learning, where the network is updated using the most recent observations in the data stream. Even under highly turbulent inflow, the (continuously updated) network is shown to comply with the strictest German grid code requirements mentioned above. To quantify its added value in terms of uncertainty reduction, 5% uncertainty is added to C_P, representing off-performance conditions (e.g. curtailment). As it is ‘model-free’, the LSTM network is not affected by the C_P uncertainty introduced. However, the bias observed in the novel wind speed and power predictions with Extended Kalman filter is increased significantly, indicating both under and over-estimation of the available power (with bias ranging from −5% to 6%) under turbine model uncertainty of 5%. The comparison is shown in Fig. 14.

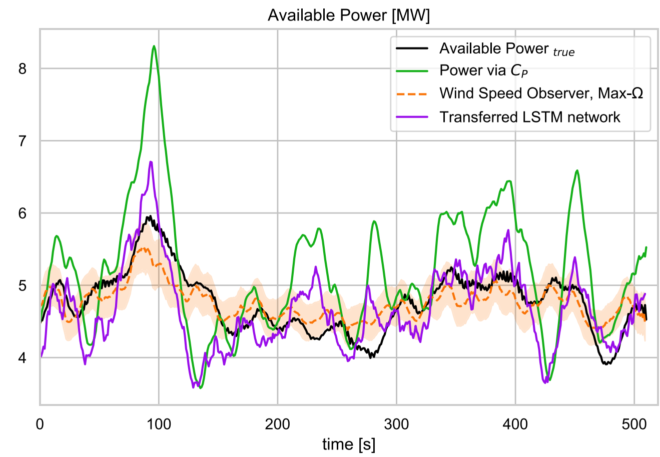


Fig. 14. The comparison of the available power forecasts under curtailment via: Direct C_P approach, Current state-of-the-art C_P based Wind Speed Observer with Kalman filter (following the maximum rotational speed control strategy for curtailment) and turbine model-free (or C_P-free) LSTM network with transfer learning. The shaded area corresponds to the effects of ±5% over and underestimation of C_P during curtailment for wind speed observer results. The LSTM network has no dependency on level of curtailment, turbine model or estimation of operational C_P, providing a robust baseline for control applications of the short-term forecasts. (Originally presented in [134]).

6.6. Trade-offs between improvements and effort needed to implement operationally

A limiting factor in the further improvement of wind power forecasts is the transition of new approaches to operational use. While some new models can still be implemented well at laboratory level, they push the limits of available resources in operational mode. This can be due to the computing power or memory needed, as well as arising from data quality limitations; per the latter, if the WPF models are not robust enough against fluctuating or noisy input data, the realisable improvement might be reduced.

The model-building process requires large amounts of historical input data, especially the ML approaches. ML models often rely on the input data being complete, i.e. if there are missing measurements/data, they have to be substituted by appropriate methods. Such substitution reduces the quality of the input data, both during training and later during operational use. The question then becomes: how robust a more complex approach is, and whether it can provide a good forecast despite losses in the quality of input data. In cases involving thousands of wind farms, the training must be automated. Is it still possible in this scenario to provide high quality training data for the models, even if the quality control is done automatically?

The cost of training an ML model often comes from the amount of computations required, known as FLOPs (floating point operations). These depend on three categories for the ML model in wind power forecasting: (a) number of parameters, (b) model size, and (c) length of data set (i.e. which time period from the past is used for training). The model size depends on the other two categories, because with more input data a more complex model is needed to represent that information. Together with model complexity the number of FLOPs increases and so does the cost of training. Depending on the algorithm, training must also be performed more often to optimise global parameters (hyperparameter tuning) or to start the model with different random initialisation. In addition, depending on the calculation, there are the costs that are caused by the training time itself. Here, in addition to the FLOPs, the hidden costs also play a role: time to load data and loss of performance due to sub-optimal parallel execution [200]. A significant increase in the complexity of network structures has been observed in recent years, especially with deep learning approaches. In general the cost per FLOP has decreased, due to adapted processors and improved training methods [201]. However, model complexity as well as training 'administration' (parallel processing, data handling and communication) continues to increase.

In operational use of a forecast model, the increased computing time for more complex models should be mentioned. For machine learning approaches this is moderate, but for NWP models it increases considerably with finer spatial resolutions. In addition, there is the memory requirement: halving the horizontal NWP grid spacing results in four times the number of calculation points. Reducing the grid spacing also increases the degrees of freedom in the numerical solutions, and potentially the uncertainty of the model results. Additionally, it has been proven that improvements in NWP forecasts as a result of higher spatial resolution are often lost in the conversion to wind power, due to increased risk of phase errors (the so-called *double punishment problem* described in Section 6.2.2 and Fig. 11). As a result, the accuracy of wind power forecasts can decrease with higher resolution, while the uncertainty can increase significantly with increased resolution [8]. From a computational perspective, doubling the resolution means more than quadrupling the required CPU time (the temporal resolution usually must increase too, and often the vertical resolution is increase as well), which must be justified by robust improvements resulting from the higher resolution.

The situation of NWP models is similar for ML approaches, with more complex ML models requiring considerably more memory. This can lead to a trade-off: do all ML models have to be persistently held in memory at the same time (due to time constraints), or is there enough time to load and execute them one after the other? The bottleneck can be either working memory or access to storage. We remind here of the operational situations, with several thousand wind farms being forecasted simultaneously, and one where the calculation must be done within only a few minutes (depending on the actuality of the forecast). In the end, a balance has to be found between the available computer capacities, maximum calculation time, and the realisable improvement of the new forecast model in operational mode. Typically the effort does not increase linearly, but rather exponentially with the improvements that may be achieved. Therefore, the potential model improvements must be assessed with a thorough cost/benefit analysis, as the gain to be expected in the operational mode might be only marginal with a high cost attached to it.

7. Uncertainty and validation

7.1. Uncertainty in verification & validation results

In a recent research study, the uncertainty of a NWP forecast and its validation has been investigated [202]. The authors verify and validate model improvements through one or more common data sets against model results; they also verify the validation methods and quantify the uncertainty of the results, dependent on the method and data set. The authors distinguish between verification and validation (V&V), with definitions found in e.g. the Sandia V&V Framework [17]. In these definitions, "..verification is concerned with checking the mechanics of the model implementation (software code) rather than checking that the model's physics are correct. Because the code mechanics may include the use of discrete equations to represent processes that were defined using partial differential equations, there is the risk that numerical errors may be introduced; verification seeks to ensure that these errors have been identified and minimised. A model may also include iterative processes to find a solution, which also requires checking. Validation, on the other hand, is determining the degree to which the model represents the real world for a particular application. Validation should be carried out after verification to ensure that the validation process identifies shortcomings in the model physics and representation. Validation of an unverified model may otherwise just identify numerical or coding problems" [202]. [202] proposed in their V&V strategy to use well-defined (i.e., given model data and observations) test cases that users can check their validation code on. They found that different validation results were seen due to different interpretations of time stamps (e.g., values of 10-min averages interpreted at the beginning or middle of the averaging period), and due to users applying different averaging techniques of model simulations in the horizontal and vertical (i.e., matching observation and model locations) [203].

7.1.1. Ad-hoc versus formal validation

Ad-hoc or quantitative validation is a useful tool for both modellers/forecast providers and instrumentation manufacturers. Typically a few instruments are used for a limited period of time, and compared to model results. Conversely, modellers/forecast providers that want to test their improvements quickly for critical cases use such short periods to compare qualitatively their simulation with observations at a few sites. Such ad-hoc validation is useful in a development phase, but contains a lot of uncertainty in the results (as to whether or not a certain improvement is only valid locally, for specific cases, or more generally). In order to validate an improvement in general terms, a formal validation is required. Formal, or quantitative validation is a bigger effort, where maybe a few or hundreds of different observations and parameters are compared to a forecast output using formal agreed-upon metrics. Such validation is also providing information about the forecast's or model's performance [202]. The same syntax of validation, in the form of building validation frameworks, has been a subject in the Recommended Practices for Forecast Solution Selection (Part 3) of IEA Wind Task 36 [186]. Such formal, qualitative validation is recommended to reduce and mitigate the uncertainty associated with single metrics that may contain quantitative information about a specific forecast product, but which are not qualitative performance measures.

7.1.2. Representativeness, significance and relevance

In the IEA Wind TCP Task 36 guideline [186], it is mentioned that "uncertainty is an inherent characteristic of the forecast evaluation process. An objective of the design and execution of a forecast validation is to minimise the uncertainty, and thereby reduce its impact on the decisions associated with forecast selection or optimisation". In order to minimise forecast evaluation uncertainty, the guideline [186] names the main sources of uncertainty in the validation step in terms of three key attributes: (1) representativeness; (2) significance; and (3) relevance. The guideline goes so far as to suggest that if any one of these attributes are not satisfactorily addressed, than a forecast validation will not provide meaningful information to the forecast solution decision process [186].

7.1.3. Forecast value versus quality

One topic that is rarely discussed in journal publications is the difference between quality of a forecast and its value to the user [204]. In academia the quality of a forecast, in terms of some specific or accepted (standard) quality measures, is the driving factor for the development of new algorithms and methodologies. The quality of a method or algorithm however does not a priori also provide value to a specific user; i.e. in advance it is not known if a specific forecast will benefit the end-user while satisfying the user's requirements.

Validation has an inherent uncertainty, and being at the 'end' of the forecasting chain, is highly underrated in its value for the end user. Bessa et al. [188] investigated validation of forecasts on their goodness of fit, and how to build in compromises regarding conflicts in validation criteria. Such conflicts typically arise when the skewness of a forecast's probability distribution benefits one party, but where this particular skewness implies a degradation of value to another party. As an example they mention a typical situation: allowing a systematic forecast bias in order to attain a certain quality measure based on an average statistical metric, which causes system security issues for a system operator that has to physically keep the electric grid in balance. If the system operator receives a significant number of biased forecasts, wrong decisions will likely be made regarding other generating units, specifically expensive reserve allocations. Such conflicts of interest are not limited to different external parties. It is not uncommon that such conflicts are observed within organisations, when the evaluation process lacks the priority (or authority) it needs in order to establish an incentivisation, with an evaluation for the forecast provider to train and tune forecasts to serve a specific purpose.

7.2. Usage case: Uncertainty validation in ramp forecasting

As described above, it is well known (e.g. Part 3 of [186]) that the commonly applied metrics of mean absolute error (MAE), mean squared error (MSE), and root-mean square error (RMSE) – which are still used to a large extent in the industry to validate wind power forecasts – alone are only capable of evaluating forecasts for specific applications under specific circumstances. For ramp forecasts, this is even more pronounced (see also Section 6.2). For example, Potter et al. [185] observed that MSE-based metrics tend to over-penalise large errors and are hence not appropriate for ramp forecasting assessment. They suggest using three metrics in concert, to verify both absolute errors (MAE and bias) and the standard deviation around the absolute error (StDev) [171,172,175,185].

7.2.1. Validation with incentivisation

Standard error metrics may provide some indication of the goodness of a ramp forecast. However, they do not account for the uncertainty in ramp occurrence. If this is the aim, we have to define a set of measures that count the hits, misses and false alarms of a ramp event and distinguish between economic value and system security by also counting the reserve spill, defined as non-allocated reserve due to forecast inaccuracies. Contingency tables are useful tools to get a first estimate of the uncertainty range over which a ramp forecast should be accountable. Such a procedure then becomes an optimisation or incentivisation for the forecast provider, to tune forecasts towards the ranges in which a hit or miss is defined. The word "incentivisation" may indicate that economic value counts more than system security. In this context, we use incentivisation as a kind of optimisation criterion that could essentially be used for incentivisation of a ramp forecast product (see Fig. 15).

Beyond contingency tables, one type of metric which can be used for the user-targeted uncertainty optimisation mentioned above is a cost function. To demonstrate by means of example: we start by defining a cost function that penalises misses, and especially particularly large misses; this would typically be used to ensure system security, and would not necessarily impact any economic value. This cost function illustrates the particular importance of sufficient penalty for large misses, to ensure that the ramp forecast is sufficiently optimised towards grid security. A cost function serving grid security does not need to be more complicated than the following equation:

$$\text{cost} = 4|P_{\text{missing}}| + \frac{1}{2}|P_{\text{spill}}|, \tag{2}$$

where P is power (in MW). The lowest in (2) is achieved if no spill nor missing power are forecasted, but it is 8 times more expensive to miss than spill. Spill may or may not include cases where there is excess wind or solar power. In other words, the coefficients can differ between different reserve products, accounting for the ramping.

There is reason to believe that a simple cost function like the above will work well in extreme cases [27]. Any high speed shut-down (HSSD) event should be covered, when misses are penalised sufficiently. If that would not be the case, then it is because the forecaster does not (try to) predict the HSSD, due e.g. to it being perceived as a rare event with low likelihood of forecast success.

7.2.2. Index validation with categorical statistics

Hirata et al. [181] introduce an index validation with an error or so-called 'confusion' matrix, for evaluation of their ramp forecasts. It is similar to a contingency table for categorical statistics in meteorology, but mostly used for classification of errors in machine learning techniques. In comparison to contingency tables, the hits in the error matrix are named true positives (TP), the correct negatives are named true negatives (TN), the misses of the contingency table are now false positives (FP) and false alarms are false negatives (FN). Hirata et al. [181] overall goal is to "maximise the true positives and true negatives and to minimise the false positives and false negatives; in other words, a perfect forecast would produce only hits and correct negatives, and no misses or false alarms".

In the statistical machine learning matrix, the *sensitivity* or *probability of detection* (POD or hit rate) indicates how good the forecast is to predict ramps and was defined by Hirata et al. [181] as

$$SV = POD = \frac{TP}{(FN + TP)} \tag{3}$$

where TP is true positive, FN is false negative. The *specificity* or probability of false detection (false alarm rate) indicates how good the forecast is to predict that no ramp has taken place, and was defined as

$$SC = \frac{TN}{(FN + TN)} \tag{4}$$

where TN is the true negative and FN the false negative. The SC (FAR) score is sensitive to false alarms, but ignores misses and can therefore be improved by issuing fewer "yes" forecasts to reduce the number of false alarms [205]. However, it is an important component of the relative operating characteristic (ROC), which is also widely used in probabilistic forecast verification. The *precision* or *false alarm ratio* shows how accurate the prediction for negatives agrees with the true outcome of false alarms and was defined by Hirata et al. [181] as

$$FAR = \frac{TN}{(FN + TN)}. \tag{5}$$

Lastly Hirata et al. [181] looked at the odds ratio

$$OR = \frac{TP * TN}{(FP * FN)}. \tag{6}$$

This measures the ratio of the odds of making a hit to the probability of a false alarm and takes prior probabilities into account, but gives better scores for rarer events [206]. For this reason it is usually not used in meteorological validation [207]. Instead the odds ratio skill score ($ORSS$) is often applied. It looks at the improvement of the forecast over random chance, and is independent of the marginal totals, i.e. it is difficult to hedge [208]. The odds ratio skill score is defined as [207]:

$$ORSS = \frac{(TP * TN) - (FP * FN)}{(TP * TN) + (FP * FN)} \tag{7}$$

One other interesting metric is the *threat score* or *critical success index*, as it considers overall how well forecasted ramp events correspond to observed events. In [209] it is stated that "..with this score, the fraction of forecast events that were correctly predicted is measured. It is a kind of accuracy measure, where correct negatives have been removed from consideration, i.e. it is only concerned with forecasts that were in a pre-defined range". In [210] it is concluded that "..while it is sensitive to hits, it penalises both misses and false alarms. There is however no distinguishing of the forecast error sources, so it, in general, will score more poorly for rare events since some hits can occur purely due to random chance".

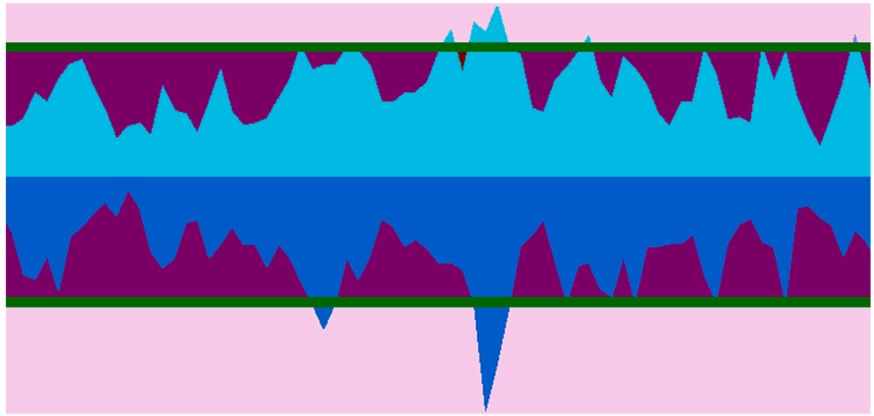


Fig. 15. Schematic plot of a typical uncertainty time series of reserve requirements due to power ramps with zero (difference between schedule and forecast equals zero) as the basis and positive ramp uncertainty in light blue, negative ramp uncertainty in blue, spill in purple and outliers not fulfilling the reserve requirement reaching out into the magenta area.

7.2.3. NWP uncertainty quantification with NOAA's "ramping tool & metric"

The ramping tool and metric (RT&M) described by Bianco et al. [187] can be employed as a use case for the validation of ramping and UQ in the wind ramp forecast process coming from the NWP side. The tool has three components:

1. identification of wind ramp events;
2. matching of wind forecast and observed power ramp;
3. calculation of a skill score for the forecast.

The tool was developed to test the ability of a model to forecast ramp events for wind energy [187]. The users can compare forecasts from a NWP model to observations from a mast for specifically selected cases. The cases are selected by defining a critical ramp rate in the power observations. The accompanying validation code then computes the skill score of the forecast's phase, duration, and amplitude in capturing the ramp event. Users of the tool can integrate skill definitions on changing the range of ramping and the associated time interval that may cause a critical situation for the user. The tool has been designed in a flexible way to accommodate other users' definitions for critical ramping, and lets them modify all ranges for their specific purpose(s). If more extreme events are an issue for the user, a weighting matrix can be applied for the skill score validation. The tool, its features, and a use case has been described in detail by Bianco et al. [187] in an open-access article. The ramp tool itself can be found on the NOAA Earth System Research Laboratory (ESRL) website.[9] The tool is run through a graphical user interface (GUI) in which one selects from several options to customise it to specific needs [202].

8. Summary and trends

8.1. Summary and highlights

This review has given a detailed look into the different sources of uncertainty, propagation of these uncertainties towards the final forecast, and possible uncertainty mitigation methods for wind power forecasting. Here we summarise the key points.

Uncertainty sources in Wind Power Forecasting.

In order to propagate uncertainty through the forecasting model chain, it is necessary to define the sources of uncertainty in each of the steps — from planning to building and operating a wind farm, and then transmission and sale on markets. We define and discuss the various aspects of data, such as data quality, availability and variety, sample size and representativeness, data contamination, as well as measurement uncertainty stemming from instrumentation. Thereafter we dove into the modelling approaches and technical routes of wind power forecasting and the associated uncertainty, as well as typical algorithms and model settings; the latter included model complexity, parameters and input uncertainty. We also defined uncertainty associated with wind resource, with wind-to-power conversion (covering rotor inertia and control), as well as uncertainty connected with turbine availability and performance variations.

Uncertainty evaluation and mitigation in the planning phase.

The planning phase of a wind power project has to deal with many uncertainties. In this work we limited our review to the evaluation of pre-construction uncertainties and how to mitigate such uncertainty by applying ensemble weather forecasts in order to design the resource assessment for the production of "bankable" data weighing risk against uncertainties.

Uncertainty evaluation and mitigation in the Operation Phase.

The operation phase, in which a wind farm generates the electricity to the grid, produces the most uncertainties during the process of both building and using a wind power forecasting system. In this phase, DATA would be the first priority to look into for improving the certainty and reliability of the forecasts. Here we define DATA to have a wide concept including NWP, historical training sample, real-time data flow, turbine condition and so on. Therefore, we firstly reviewed how "added" information about weather and wind turbine as well as more features in model inputs can help the uncertainty mitigation. In order to reduce the uncertainty of NWP, which generates the most uncertainty to day-ahead forecasting, we reviewed the methods of data assimilation and data-driven correction using wind turbine measurements, and the ensemble approach. Historical training samples also have to be carefully cleaned and selected to ensure its authenticity and representativeness. During the operation and maintenance, availability of real-time data flow should be improve by detecting abnormal data and filling missing data; moreover wind turbine performance and models should be calibrated to mitigate uncertainties.

Uncertainty evaluation and mitigation methods in the Market Phase.

The market phase, or the phase in which wind power starts to be traded in a power exchange is characterised by adding one more level of uncertainty into the forecasting process for wind power production. In this phase, incentives and policies need to consider uncertainties that may arise in the case of missing transparency, as well as the distribution of capacity and infrastructure available to distribute the generation. The uncertainties forecasting processes have to deal with in a system

[9] The ramp tool may be found at https://www.esrl.noaa.gov/psd/products/ramp_tool/, and is available as either an executable file or a set of Matlab user files.

with a market have been described and analysed with focus on state of the art challenges and already existing mitigation measures for high penetration areas, commercialised balancing and balancing over interconnections. Some future aspects on the impacts of storage on small and larger scales as well as price forecasting has been touched and references provided for further study of these subjects.

Advanced prediction methods to mitigate uncertainty.

Uncertainty in weather and wind power forecasting has been researched for many years and a number of methods to mitigate unwanted uncertainty has been developed. We reviewed in this work advanced artificial intelligence (AI) methods and worked through a number of forecasting approaches such as ramp forecasting, cut-out wind speeds and high-speed shutdown, looked into the quantification of uncertainty for reserve forecasting, assessed the uncertainty in the dynamic modelling of wind turbines and discussed typical trade-off uncertainties of forecast improvements versus required resources for improvement implementations.

Uncertainty in Verification and Validation Results.

Validation and verification is a crucial part for any project and subject to different types of uncertainty. This review has been limited to the description of uncertainty in validation results, where we distinguished between ad-hoc versus formal validation, described the importance of representativeness, significance and relevance of verification results; consideration of different attributes of forecast value and forecast quality was also made. Additionally, we presented some use cases as examples for validation with incentivisation, index validation with categorical statistics, and NOAA's Ramping Tool & Metric for quantification of weather forecast uncertainty.

8.2. Trends and future work

Current topics in the forecasting community's conferences, workshops and publications, both in weather forecasting and wind power forecasting, show that uncertainty in forecasting has been recognised as a part of the forecasting chain that needs attention. The question is no longer *whether* there are uncertainties to deal with, but rather how to deal with, quantify, and communicate recognised uncertainties in the forecasting processes in a way that can address end-users' applications. There are a number of gaps in the use of uncertainty information in end-user applications, and in the way uncertainties can be communicated. In other words, a paradigm change has happened regarding the recognition of uncertainties, while the integration and implementation of methods to deal with the new information is still an outstanding task for researchers, developers and end-users—with many unanswered questions. One of those questions is the quantification of the various uncertainty sources, to support this more qualitative review with numbers.

One of the trends that can be observed is the increasing level of interdisciplinary work across various areas, such as meteorology, (wind) power engineering, applied mathematics, physics, sociology, and public policy. The interconnection of model input/output and political inter-dependencies in the integration of wind power into the electrical grid, and renewables in general, points to the necessity of a new definition for forecasting as "integrated forecasting"; this is in line with other disciplines, such as integrated design of wind power systems and integrated wind farm control.

Another trend is the extended horizon of forecasts from short to medium-term down to minutes and second-ahead forecasting, as well as seasonal or long-term forecasting; the latter includes quantification of uncertainties in future production patterns, and the impact of climate change. In both new forecast horizons, but especially in the minute or second-term horizons, we can see a trend to better interpret and use different types of measurements, as well as observations from volume based instrumentation, for better UQ in wind power forecasting.

As a result of this review and discussions about the state-of-the-art in uncertainty quantification through the modelling chain of wind and wind power forecasting, we have identified a number of outstanding topics. These span the areas of resource assessment, wind power forecasting, and validation, as well as market instruments, and are listed below.

- **Resource Assessment**
  - Uncertainty in wake modelling over terrain, and wake-associated inter-farm effects;
  - top-down/entrainment effects on wind farms;
  - wind farm blockage, its stability dependence, and interaction between farms;
  - uncertainty quantification (UQ) for advanced physical flow modelling, i.e. RANS-CFD, over complex terrain.
- **Wind Power Forecasting and Uncertainty**
  - Development of a standardised guideline to validate the uncertainty in wind power forecasting;
  - development of an 'exemplary uncertainty chain calculation platform', illustrating the uncertainty chain using data from a global set of wind farms to build up standardised and public data sets for comparisons of different WPF and uncertainty analysis methods — for the forecast developers, providers and end-users.
- **Market Instruments**
  - Development of a guideline on uncertainties resulting from transparency and omissions, particularly for emerging markets.

CRediT authorship contribution statement

Jie Yan: Lead of the paper, Wrote Sections 1, 2, 4 and subsection 6.1, Contributed to Sections 8. **Corinna Möhrlen:** Co-lead of the paper, Performed editing led and wrote section 5, 7, wrote section 8 and subsections 6.2,6.3, 6.4, contributed to sections 1, subsections 2.1.4 and 2.3.2 in 2, subsection 3.2 in 3, section 4.1 in section 4. **Tuhfe Göçmen:** Initiated the framework for uncertainty propagation within wind power forecasting, Contributed to Sections 1, 2, 4, 6 and 8, Wrote Sections 2.5, 4.5 and 6.5. **Mark Kelly:** Wrote Sections 2.1.4, 2.4, 3.1, and 3.3, Writing text for the Section 2 intro, Sections 2.3.2 and 8.2, with contributions to Sections 1, 2, 4, and 8, Performed final editing for language (native-speaker level), flow, and content, Adding references. **Arne Wessel:** Wrote Section 6.6, Contributed to Section 2. **Gregor Giebel:** Contributed to Section 1, Provided supervision in all phases from writing – review and funding acquisition, English language checks.

Declaration of competing interest

All authors in this paper declare no conflict of interest.

Acknowledgements

This paper belongs to subtask 2.3 of *IEA Wind Task 36* on *Wind Power Forecasting: Uncovering uncertainty origins and development through the whole modelling chain*. The authors want to thank the IEA Wind Task 36 community for their valuable input in many discussions around this topic. Special thanks to Caroline Draxl for the valuable input to the validation Section 7.

References

[1] Brown BG, Katz RW, Murphy AH. Time series models to simulate and forecast wind speed and wind power. J Appl Meteorol Climatol 1984;23(8):1184–95.

[2] Giebel G, Brownsword R, Kariniotakis G, Denhard M, Draxl C. The state-of-the-art in short-term prediction of wind power: A literature overview. 2nd ed.. ANEMOS.plus; 2011, http://dx.doi.org/10.11581/DTU:00000017, Project funded by the European Commission under the 6th Framework Program, Priority 6.1: Sustainable Energy Systems.

[3] Lorenz EN. The predictability of a flow which possesses many scales of motion. Tellus 1969;21(3):289–307.

[4] Slingo J, Palmer T. Uncertainty in weather and climate prediction. Phil Trans R Soc A 2011;369(1956):4751–67.

[5] Wyngaard JC. Turbulence in the atmosphere. Cambridge University Press; 2010, p. 393.

[6] Foley AM, Leahy PG, Marvuglia A, McKeogh EJ. Current methods and advances in forecasting of wind power generation. Renew Energy 2012;37(1):1–8.

[7] Gill J, Rubiera J, Martin C, Cacic I, Mylne K, Dehui C, Jiafeng G, Xu T, Yamaguchi M, Foamouhoue AK, et al. Guidelines on communicating forecast uncertainty. Technical Report WMO/TD No. 4122, Geneve, Switzerland: World Meteorological Organization; 2008.

[8] Möhrlen C. Uncertainty in wind energy forecasting (Ph.D. thesis), Cork, Ireland: National University of Ireland; 2004, URL: http://library.ucc.ie/record=b1501384.

[9] Pinson P, Nielsen H, Madsen H. Methods for the estimation of the uncertainty of wind power forecasts. Technical Report, Informatics and Mathematical Modeling, Technical University of Denmark; 2007.

[10] Yan J, Liu Y, Han S, Wang Y, Feng S. Reviews on uncertainty analysis of wind power forecasting. Renew Sustain Energy Rev 2015;52:1322–30.

[11] Huang C-M, Kuo C-J, Huang Y-C. Short-term wind power forecasting and uncertainty analysis using a hybrid intelligent method. IET Renew Power Gener 2017;11(5):678–87.

[12] Juban J, Fugon L, Kariniotakis G. Uncertainty estimation of wind power forecasts: comparison of probabilistic modelling approaches. In: European wind energy conference and exhibition 2008. Brussels, Belgium: European Wind Energy Association (EWEA); 2008.

[13] Zhang X, Zhang X, Liao S, Chen W, Wang X. Prediction error analysis of wind power based on clustering and non-parametric kernel density estimation. Acta Energ Solaris Sin 2019;040(012):3594–604.

[14] Pahlow M, Möhrlen C, Jørgensen J. Application of cost functions for large-scale integration of wind power using a multi-scheme ensemble prediction technique. In: Optimization advances in electric power systems. UK: Nova Science Publishers; 2009, p. 151–80.

[15] Bremnes JB. Probabilistic wind power forecasts using local quantile regression. Wind Energy 2004;7(1):47–54.

[16] Nielsen HA, Madsen H, Nielsen TS. Using quantile regression to extend an existing wind power forecasting system with probabilistic forecasts. Wind Energy 2006;9(1–2):95–108.

[17] Hills RG, Maniaci DC, Naughton JW. V & V framework. Technical Report SAND2015-7455, Albequerque & Livermore, USA: Sandia National Laboratories; 2015, http://dx.doi.org/10.2172/1214246, URL: https://www.osti.gov/biblio/1214246.

[18] Giebel G, Cline J, Frank H, Shaw W, Pinson P, Hodge B-M, Kariniotakis G, Madsen J, Möhrlen C. Wind power forecasting: IEA Wind Task 36 and future research issues. J Phys Conf Ser 2016;753:032042.

[19] Zhang Y, Wang J, Wang X. Review on probabilistic forecasting of wind power generation. Renew Sustain Energy Rev 2014;32:255–70.

[20] Bessa RJ, Möhrlen C, Fundel V, Siefert M, Browell J, Haglund El Gaidi S, Hodge B-M, Cali U, Kariniotakis G. Towards improved understanding of the applicability of uncertainty forecasts in the electric power industry. Energies 2017;10(9).

[21] Möhrlen C, Jørgensen J. The role of ensemble forecasting in integrating renewables into power systems: From theory to real-time applications. In: Integration of large-scale renewable energy into bulk power systems - from planning to operation. Power electronics and power systems, vol. 1, Springer International Publishing; 2017, p. 79–134. http://dx.doi.org/10.1007/978-3-319-55581-2.

[22] Fundel VJ, Fleischhut N, Herzog SM, Göber M, Hagedorn R. Promoting the use of probabilistic weather forecasts through a dialogue between scientists, developers and end-users. Q J R Meteorol Soc 2019;145:210–31.

[23] Cui M, Feng C, Wang Z, Zhang J, Wang Q, Florita A, Krishnan V, Hodge B-M. Probabilistic wind power ramp forecasting based on a scenario generation method. In: 2017 IEEE power energy society general meeting. 2017, p. 1. http://dx.doi.org/10.1109/PESGM.2017.8274394.

[24] Möhrlen C, Bessa J, Fleischhut N. A decision-making experiment under wind power forecast uncertainty. Meteorol Appl 2022. (accepted).

[25] Möhrlen C, Pahlow M, Jørgensen J. Untersuchung verschiedener handelsstrategien für wind-und solarenergie unter berücksichtigung der EEG 2012 novellierung. Z Energiewirtsch 2012;36(1).

[26] Möhrlen C, Pahlow M, Jørgensen J. Investigation of various trading strategies for wind and solar power developed for the new EEG 2012 rules – English author translation of [25]. 2012. URL: http://download.weprog.com/WEPROG_Trading_strategies_EEG2012_ZEFE_71-2012-01_en.pdf.

[27] Möhrlen C, Vestergaard U, J. Ryan KC, Griffin S. Probabilistic forecasting tools for high-wind penetration areas: an Irish case study. In: Proceedings of the 18th international workshop on large-scale integration of wind power into power systems as well as on transmission networks for offshore wind power plant. 2019.

[28] Murcia Leon JP. Uncertainty quantification in wind farm flow models (Ph.D. thesis), Denmark: DTU Wind Energy; 2017, URL: https://orbit.dtu.dk/en/publications/uncertainty-quantification-in-wind-farm-flow-models.

[29] Foti D, Yang X, Sotiropoulos F. Uncertainty quantification of infinite aligned wind farm performance using non-intrusive polynomial chaos and a distributed roughness model. Wind Energy 2017;20(6):945–58.

[30] Craig A, Optis M, Fields MJ, Moriarty P. Uncertainty quantification in the analyses of operational wind power plant performance. J Phys Conf Ser 2018;1037:052021.

[31] Khosravi A, Nahavandi S. An optimized mean variance estimation method for uncertainty quantification of wind power forecasts. Int J Electr Power Energy Syst 2014;61:446–54.

[32] Fernández-González S, Martín ML, Merino A, Sánchez JL, Valero F. Uncertainty quantification and predictability of wind speed over the Iberian Peninsula. J Geophys Res: Atmos 2017;122(7):3877–90.

[33] Jin J, Che Y, Zheng J, Xiao F. Uncertainty quantification of a coupled model for wind prediction at a wind farm in Japan. Energies 2019;12(8):1505.

[34] Quan H, Khosravi A, Yang D, Srinivasan D. A survey of computational intelligence techniques for wind power uncertainty quantification in smart grids. IEEE Trans Neural Netw Learn Syst 2020;31(11):4582–99.

[35] Wang J, AlShelahi A, You M, Byon E, Saigal R. Integrative density forecast and uncertainty quantification of wind power generation. IEEE Trans Sustain Energy 2021;12(4):1864–75.

[36] Richter P, Wolters J, Frank M. Uncertainty quantification of offshore wind farms using Monte Carlo and sparse grid. Energy Sources B 2021;1–27.

[37] Wei W, Wu J, Yu Y, Niu T, Deng X. Uncertainty quantification analysis of wind power: A data-driven monitoring-forecasting framework. IEEE Access 2021;9:84403–16.

[38] Kiureghian AD, Ditlevsen O. Aleatory or epistemic? Does it matter? Struct Saf 2009;31(2):105–12.

[39] IEC 61400-12-1 Wind turbines – Part 12-1: Power performance measurements of electricity producing wind turbines. Technical Report, International Electrotechnical Commission; 2017.

[40] Friis Pedersen T, Dahlberg J-Å, Busche P. ACCUWIND - Classification of five cup anemometers according to IEC 61400-12-1. Denmark. Forskningscenter Risoe. Risoe-R, vol. 1556(EN), 2006.

[41] Eecen P, Mouzakis F, Cuerva A. ACCUWIND, Work package 3 Final report ECN-C-06-047. Technical Report, the Netherlands: Energy Research Centre of the Netherlands, ECN; 2006.

[42] JCGM 100:2008 Evaluation of measurement data – Guide to the expression of uncertainty in measurement. Technical Report, Joint Committee for Guides in Meteorology; 2008, URL: https://www.bipm.org/documents/20126/2071204/JCGM_100_2008_E.pdf.

[43] JCGM 106:2012 Evaluation of measurement data – The role of measurement uncertainty in conformity assessment. Technical Report, Joint Committee for Guides in Meteorology; 2012, URL: https://www.bipm.org/documents/20126/2071204/JCGM_106_2012_E.pdf.

[44] JCGM 102:2011 Evaluation of measurement data – Supplement 1 to the "Guide to the expression of uncertainty in measurement. Technical Report, Joint Committee for Guides in Meteorology; 2011, URL: https://www.bipm.org/documents/20126/2071204/JCGM_102_2011_E.pdf.

[45] Sobri S, Koohi-Kamali S, Rahim NA. Solar photovoltaic generation forecasting methods: A review. Energy Convers Manage 2018;156:459–97.

[46] zhi Wang H, qiang Li G, bin Wang G, chun Peng J, Jiang H, tao Liu Y. Deep learning based ensemble approach for probabilistic wind power forecasting. Appl Energy 2017;188:56–70.

[47] Corotis RB, Sigl AB, Cohen MP. Variance analysis of wind characteristics for energy conversion. J Appl Meteorol 1977;16(11):1149–57.

[48] Kavasseri RG, Seetharaman K. Day-ahead wind speed forecasting using f-ARIMA models. Renew Energy 2009;34(5):1388–93.

[49] Box GE, Jenkins GM, Reinsel GC. Time series analysis: Forecasting and control (Fourth edition). Wiley; 2013, p. 746. http://dx.doi.org/10.1002/9781118619193.

[50] Zhu X, Genton MG. Short-term wind speed forecasting for power system operations. Internat Statist Rev 2012;80(1):2–23.

[51] Pinson P, Madsen H, Nielsen HA, Papaefthymiou G, Klöckl B. From probabilistic forecasts to statistical scenarios of short-term wind power production. Wind Energy 2009;12(1):51–62.

[52] Rohrig K, Ernst B, Hoppe-Kilpper M, Schlögl F. Online-monitoring and prediction of wind power in german transmission system operation centres. In: World wind energy conference, Cape Town, South Africa. 2003.

[53] Aggarwal CC. Neural networks and deep learning. Springer; 2018, p. 174–7.
[54] Li L, Liu Y, Yang Y, Han S. Short-term wind speed forecasting based on CFD pre-calculated flow fields. Proc CSEE 2013;33(7):27–30.
[55] Chen P, Pedersen T, Bak-Jensen B, Chen Z. ARIMA-based time series model of stochastic wind power generation. IEEE Trans Power Syst 2010;25(2):667–76.
[56] Zuluaga CD, Alvarez MA, Giraldo E. Short-term wind speed prediction based on robust Kalman filtering: An experimental comparison. Appl Energy 2015;156(OCT.15):321–30.
[57] Peng X, Xiong L, Wen J, Chen S, Deng D, Fen S, Wang B. A summary of the state of the art for short-term and ultra-short-term wind power prediction of regions. Proc CSEE 2016;36(023):6315–26.
[58] Liu R, Huang L. Wnd power forecasting based on dynamic neural networks. Autom Electr Power Syst 2012;36(11):19–22.
[59] Liu Y, Shi J, Yang Y, Han S. Piecewise support vector machine model for short-term wind-power prediction. Int J Green Energy 2009;6(5):479–89.
[60] Zeng J, Qiao W. Short-term wind power prediction using a wavelet support vector machine. IEEE Trans Sustain Energy 2012;3(2):255–64.
[61] Yan J, Liu Y, Han S, Qiu M. Wind power grouping forecasts and its uncertainty analysis using optimized relevance vector machine. Renew Sustain Energy Rev 2013;27(nov.):613–21.
[62] Abdoos AA. A new intelligent method based on combination of VMD and ELM for short term wind power forecasting. Neurocomputing 2016;203(Aug.26):111–20.
[63] Lee D, Baldick R. Short-term wind power ensemble prediction based on Gaussian processes and neural networks. IEEE Trans Smart Grid 2014;5(1):501–10.
[64] Yang M, Zhang L. Review on ultra-short term wind power forecasting based on data-driven approach. Power Syst Prot Control 2019;47(13):171–86.
[65] Würth I, Valldecabres L, Simon E, Möhrlen C, Uzunoglu B, Gilbert C, Giebel G, Schlipf D, Kaifel A. Minute-scale forecasting of wind power - Results from the collaborative workshop of IEA wind task 32 and 36. Energies 2019;12(4):712.
[66] EWeLINE project: Development of innovative weather and power forecast models for the grid integration of weather dependent energy sources. Fraunhofer Institute of Energy and Economics and Energy System Technology IEE; 2011, URL: http://www.projekt-eweline.de/en/project.html.
[67] Marquis M, Wilczak J, Finley C, Freedman J. Wind Forecasting Improvement Project (WFIP) final report. Technical Report, National Oceanographic Atmosphere Administration; 2014, URL: http://energy.gov/sites/prod/files/2014/05/f15/wfipandnoaafinalreport.pdf.
[68] Ebert E, Wilson L, Weigel A, Mittermaier M, Nurmi P, Gill P, Göber M, Joslyn S, Brown B, Fowler T, Watkins A. Progress and challenges in forecast verification. Meteorol Appl 2013;20(2):130–9.
[69] van der Plas E, Schmeits M, Hooijman N, Kok K. A comparative verification of high-resolution precipitation forecasts using model output statistics. Mon Weather Rev 2017;145(10):4037–54.
[70] Global telecommunication System, the co-ordinated global system of telecommunication facilities and arrangements for the rapid collection, exchange and distribution of observations and processed information within the framework of the World Weather Watch. World Meteorological Organisation; 2021, URL: https://community.wmo.int/activity-areas/global-telecommunication-system-gts, [Accessed 18-February-2021].
[71] Molteni F, Buizza R, Palmer TN, Petroliagis T. The ECMWF ensemble prediction system: Methodology and validation. Q J R Meteorol Soc 1996;122(529):73–119.
[72] Toth Z, Kalnay E. Ensemble forecasting at NMC: The generation of perturbations. Bull Am Meteorol Soc 1993;74(12):2317–30.
[73] Montani A, Cesari D, Marsigli C, Paccagnella T. Seven years of activity in the field of mesoscale ensemble forecasting by the COSMO-LEPS system: main achievements and open challenges. Tellus A 2011;63(3):605–24.
[74] Alessandrini S, Sperati S, Pinson P. A comparison between the ECMWF and COSMO ensemble prediction systems applied to short-term wind power forecasting on real data. Appl Energy 2013;107:271–80.
[75] Xu K, Yan J, Zhang H, Zhang H, Han S, Liu Y. Quantile based probabilistic wind turbine power curve model. Appl Energy 2021;296:116913.
[76] Staffell I, Green R. How does wind farm performance decline with age? Renew Energy 2014;66:775–86.
[77] Lee JC, Jason Fields M. An overview of wind-energy-production prediction bias, losses, and uncertainties. Wind Energy Sci 2021;6(2):311–65.
[78] Clerc A, Anderson M, Stuart P. A systematic method for quantifying wind flow modelling uncertainty in wind resource assessment. J Wind Eng Ind Aerodyn 2012;111:85–94.
[79] Kelly M, Jørgensen HE. Statistical characterization of roughness uncertainty and impact on wind resource estimation. Wind Energy Sci 2017;2(1):189–209.
[80] Kelly M, Troen I. Probabilistic stability and "tall" wind profiles: theory and method for use in wind resource assessment. Wind Energy 2016;19(2):227–41.
[81] Kelly M, Kersting G, Mazoyer P, Yang C, Fillols FH, Clark S, Matos JC. Uncertainty in vertical extrapolation of measured wind speed via shear. Technical Report DTU Wind Energy E-0195(EN), Roskilde, Denmark: Technical University of Denmark; 2019, p. 25.
[82] Früh WG. Long-term wind resource and uncertainty estimation using wind records from Scotland as example. Renew Energy 2013;50:1014–26.
[83] Foken T. 50 Years of the monin–obukhov similarity theory. Bound-Layer Meteorol 2006;1(119):431–47.
[84] Kelly M, Gryning S-E. Long-term mean wind profiles based on similarity theory. Bound Layer Meteorol 2010;136(3):377–90.
[85] van Wingerden JW, Fleming PA, Göçmen T, Eguinoa I, Doekemeijer BM, Dykes K, Lawson M, Simley E, King J, Astrain D, Iribas M, Bottasso CL, Meyers J, Raach S, Kölle K, Giebel G. Expert elicitation on wind farm control. In: TORQUE 2020 ; Conference date: 28-09-2020 through 02-10-2020, Vol. 1618/2. IOP Publishing; 2020, http://dx.doi.org/10.1088/1742-6596/1618/2/022025.
[86] Zhang Y, Wang J. A distributed approach for wind power probabilistic forecasting considering spatiotemporal correlation without direct access to off-site information. In: 2020 IEEE power energy society general meeting (PESGM). 2020, p. 1. http://dx.doi.org/10.1109/PESGM41954.2020.9281985.
[87] Yan J, Zhang H, Liu Y, Han S, Li L, Lu Z. Forecasting the high penetration of wind power on multiple scales using multi-to-multi mapping. IEEE Trans Power Syst 2018;33(3):3276–84.
[88] Andrade JR, Bessa RJ. Improving renewable energy forecasting with a grid of numerical weather predictions. IEEE Trans Sustain Energy 2017;8(4):1571–80.
[89] Yesilbudak M, Sagiroglu S, Colak I. A novel implementation of kNN classifier based on multi-tupled meteorological input data for wind power prediction. Energy Convers Manage 2017;135(MAR.):434–44.
[90] Landry M, Edinger TP, Patschke D, Varrichio C. Probabilistic gradient boosting machines for GEFCom2014 wind forecasting. Int J Forecast 2016;32(3):1061–6.
[91] Silva, Lucas. A feature engineering approach to wind power forecasting. Int J Forecast 2014;30(2):395–401.
[92] Bhaskar K, Singh SN. AWNN-assisted wind power forecasting using feed-forward neural network. IEEE Trans Sustain Energy 2012;3(2):306–15.
[93] Davo F, Alessandrini S, Sperati S, Delle Monache L, Airoldi D, Vespucci MT. Post-processing techniques and principal component analysis for regional wind power and solar irradiance forecasting. Sol Energy 2016;134:327–38.
[94] Wu Y, Wu Q, Zhu J. Data-driven wind speed forecasting using deep feature extraction and LSTM. Renew Power Gener IET 2019;13(12):2062–9.
[95] Peng X, Wang H, Lang J, Li W, Li C. EALSTM-QR: Interval wind-power prediction model based on numerical weather prediction and deep learning. Energy 2020;220:119692.
[96] Kazutoshi, Higashiyama, Yu, Fujimoto, Yasuhiro, Hayashi. Feature extraction of NWP data for wind power forecasting using 3D-convolutional neural networks - ScienceDirect. Energy Procedia 2018;155:350–8.
[97] Zhao Y, Ye L. A numerical weather prediction feature selection approach based on minimal-redundancy-maximal-relevance strategy for short-term regional wind power prediction. Proc CSEE 2015.
[98] Feng C, Cui M, Hodge BM, Zhang J. A data-driven multi-model methodology with deep feature selection for short-term wind forecasting. Appl Energy 2017;190(MAR.15):1245–57.
[99] Chen N, Qian Z, Nabney IT, Meng X. Wind power forecasts using Gaussian processes and numerical weather prediction. IEEE Trans Power Syst 2014;29(2):656–65.
[100] Li S, Wang P, Goel L. Wind power forecasting using neural network ensembles with feature selection. IEEE Trans Sustain Energy 2015;6(4):1447–56.
[101] Shi K, Qiao Y, Zhao W, Wang Q, Liu M, Lu Z. An improved random forest model of short-term wind-power forecasting to enhance accuracy, efficiency, and robustness. Wind Energy 2018;21(12):1383–94.
[102] Möhrlen C, Vestergaard U. Met mast and alternative study. Technical Report, EIRGRID Group; 2019, URL: http://www.eirgridgroup.com/site-files/library/EirGrid/EIRGRID-Met-Mast-and-Alternatives-Study-Version-2.pdf.
[103] IEAWind Task 36. In: Möhrlen C, Zack JW, Giebel G, editors. IEA Wind Task 36 Recommended practice on renewable energy forecast solution selection part 4: Meteorological and power data requirements for real-time forecasting applications. Technical Report, International Energy Agency Technical Cooperation Programme Wind; 2022, URL: https://iea-wind.org/task-36/task-36-publications/recommended-practice/.
[104] International energy agency technical coordination program, IEA wind task 36: Wind energy forecasting. 2016-2021, URL: https://iea-wind.org/task-36.
[105] Wang H, Han S, Liu Y, Yan J, Li L. Sequence transfer correction algorithm for numerical weather prediction wind speed and its application in a wind power forecasting system. Appl Energy 2019;237(1):1–10.
[106] Ding M, Zhou H, Xie H, Wu M, Nakanishi Y, Yokoyama R. A gated recurrent unit neural networks based wind speed error correction model for short-term wind power forecasting. Neurocomputing 2019;365:54–61.
[107] Xu W, Liu P, Cheng L, Zhou Y, Liu Y. Multi-step wind speed prediction by combining a WRF simulation and an error correction strategy. Renew Energy 2021;163:772–82.
[108] Lorenz E. Atmospheric predictability as revealed by naturally occurring analogues. J Atmos Sci 1969;26(4):636ff.
[109] Delle Monache L, Eckel F, Rife DL, Nagarajan B, Searight K. Probabilistic weather prediction with an analog ensemble. Mon Weather Rev 2013;141(10):3498–516.
[110] Alessandrini S, Delle Monache L, Sperati S, Cervone G. An analog ensemble for short-term probabilistic solar power forecast. Appl Energy 2015;157:95–110.

[111] Hess P, Brezowsky H. Katalog der Grosswetterlagen Europas. Dt. Wetterdienst; 1969.
[112] Drücke J, Borsche M, James P, Kaspar F, Pfeifroth U, Ahrens B, Trentmann J. Climatological analysis of solar and wind energy in Germany using the Grosswetterlagen classification. Renew Energy 2021;164:1254–66.
[113] Pinson P, Christensen L, Madsen H, Sørensen P, Donovan M, Jensen L. Regime-switching modelling of the fluctuations of offshore wind generation. J Wind Eng Ind Aerodyn 2008;96(12):2327–47.
[114] Yang J, Wang X, Luo X, Jiang C. Intelligent combined prediction of wind power based on numerical weather prediction and fuzzy clustering. IFAC Pap 2015;48(28):538–43.
[115] Zhao J, Guo ZH, Su ZY, Zhao ZY, Xiao X, Liu F. An improved multi-step forecasting model based on WRF ensembles and creative fuzzy systems for wind speed. Appl Energy 2016;162(15):808–26.
[116] Wu Y, Su P, Wu T, Hong J, Hassan MY. Probabilistic wind-power forecasting using weather ensemble models. IEEE Trans Ind Appl 2018;54(6):5609–20.
[117] Eide SS, Bremnes JB, Steinsland I. Bayesian model averaging for wind speed ensemble forecasts using wind speed and direction. Weather Forecast 2017;32(6):2217–27.
[118] Schefzik R, Thorarinsdottir TL, Gneiting T. Uncertainty quantification in complex simulation models using ensemble copula coupling. Statist Sci 2013;28(4):616–40.
[119] Zied BB, Heppelmann T, Theis SE, Pinson P. Generation of scenarios from calibrated ensemble forecasts with a dual-ensemble copula-coupling approach. Mon Weather Rev 2016;144(12):4737–50.
[120] Duan Y, Lv Y, Liu Y-L, Wang F-Y. An efficient realization of deep learning for traffic data imputation. Transp Res C 2016;72:168–81.
[121] Miranda V, Krstulovic J, Keko H, Moreira C, Pereira J. Reconstructing missing data in state estimation with autoencoders. IEEE Trans Power Syst 2012;27(2):604–11.
[122] Zhang S. Nearest neighbor selection for iteratively kNN imputation. J Syst Softw 2012;85(11):2541–52.
[123] Tang F, Ishwaran H. Random forest missing data algorithms. Stat Anal Data Min: ASA Data Sci J 2017;10(6):363–77.
[124] Yan J, Zhang H, Liu Y, Han S, Li L. Uncertainty estimation for wind energy conversion by probabilistic wind turbine power curve modelling. Appl Energy 2019;239:1356–70.
[125] Wang Y, Infield DG, Stephen B, Galloway SJ. Copula-based model for wind turbine power curve outlier rejection. Wind Energy 2013;17(11):1677–88.
[126] Ye X, Lu Z, Qiao Y, Min Y, O'Malley M. Identification and correction of outliers in wind farm time series power data. IEEE Trans Power Syst 2016;31(6):4197–205.
[127] Zhao Y, Ye L, Wang W, Sun H, Ju Y, Tang Y. Data-driven correction approach to refine power curve of wind farm under wind curtailment. IEEE Trans Sustain Energy 2018;9(1):95–105.
[128] Long H, Sang L, Wu Z, Gu W. Image-based abnormal data detection and cleaning algorithm via wind power curve. IEEE Trans Sustain Energy 2020;11(2):938–46.
[129] Yan J, Gao X, Liu Y, Han S, Li L, Ma X, Gu C, Bhakar R, Li F. Adaptabilities of three mainstream short-term wind power forecasting methods. J Renew Sustain Energy 2015;7(5):1298–311.
[130] Yan J, Gao X, Liu Y, Han S, Li L, Ma X, Gu C, Bhakar R, Li F. Adaptabilities of three mainstream short-term wind power forecasting methods. J Renew Sustain Energy 2015;7(5):1298–311.
[131] Yesilbudak M. Clustering analysis of multidimensional wind speed data using k-means approach. In: 2016 IEEE international conference on renewable energy research and applications (ICRERA). 2016, p. 961–5. http://dx.doi.org/10.1109/ICRERA.2016.7884477.
[132] Shi J, Malik J. Normalized cuts and image segmentation. IEEE Trans Pattern Anal Mach Intell 2000;22(8):888–905.
[133] Ng AY, Jordan MI, Weiss Y. On spectral clustering: Analysis and an algorithm. In: Advances in neural information processing systems. MIT Press; 2001, p. 849–56.
[134] Göçmen T, Meseguer Urbán A, Liew J, Lio AWH. Model-free estimation of available power using deep learning. Wind Energy Sci 2021;6(1):111–29.
[135] Law H, Koutsos V. Leading edge erosion of wind turbines: Effect of solid airborne particles and rain on operational wind farms. Wind Energy 2020;23(10):1955–65.
[136] Commission Regulation on submission and publication of data in electricity markets and amending Annex I to Regulation (EC) No 714/2009 of the European Parliament and of the Council. Technical Report 543, The European Commission; 2013, URL: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:163:0001:0012:EN:PDF.
[137] National Grid. ETR compliance summary of European Transparency Regulation. Technical Report, National Grid; 2014.
[138] Von der Fehr N-HM. Transparency in electricity markets. Econ Energy Environ Policy 2013;2(2):87–106.
[139] Hirth L, Mühlenpfordt J, Bulkeley M. The ENTSO-E Transparency Platform – A review of Europe's most ambitious electricity data platform. Appl Energy 2018;225:1054–67.
[140] Gaffney F, Deane J, Gallachóir B. Reconciling high renewable electricity ambitions with market economics and system operation: Lessons from Ireland's power system. Energy Strategy Rev 2019;26:100381.
[141] Ackermann T, Ancell G, Borup LD, Eriksen PB, Ernst B, Groome F, Lange M, Möhrlen C, Orths AG, O'Sullivan J, de la Torre M. Where the wind blows. IEEE Power Energy Mag 2009;7(6):65–75.
[142] Haupt SE, Garcia Casado M, Davidson M, Dobschinski J, Du P, Lange M, Miller T, Möhrlen C, Motley A, Pestana R, Zack J. The use of probabilistic forecasts: Applying them in theory and practice. IEEE Power Energy Mag 2019;17(6):46–57.
[143] Dobschinski J, Bessak R, Du P, Geisler K, Haupt S, Lange M, Möhrlen C, Nakafuji D, de la Torre Rodriguez M. Uncertainty forecasting in a nutshell: Prediction models designed to prevent significant errors. IEEE Power Energy Mag 2017;15(6).
[144] Howley M. Wind generation in Ireland february 2021. 2021, URL: https://www.linkedin.com/pulse/wind-generation-ireland-february-2021-martin-howley/.
[145] Gawel E, Purkus A. Promoting the market and system integration of renewable energies through premium schemes – A case study of the German market premium. Energy Policy 2013;61:599–609.
[146] Genoese F, Genoese M, Wietschel M. Occurrence of negative prices on the german spot market for electricity and their influence on balancing power markets. In: 2010 7th international conference on the european energy market. 2010, p. 1–6. http://dx.doi.org/10.1109/EEM.2010.5558703.
[147] Zarco J. Production begins at 500 MW mexican solar panel fab. 2020, URL: https://www.pv-magazine.com/2020/11/25/production-begins-at-500-mw-mexican-solar-panel-fab.
[148] On Sunday, april 5, low, some negative, prices were registered in the European electricity markets. 2020, URL: https://aleasoft.com/sunday-april-5-low-some-negative-prices-were-registered-european-electricity-markets/, [Accessed 12-February-2021].
[149] Hansen S. Klimaplan 2030 - Realisme Eller Utopi?. Veterania; 2020.
[150] International Renewable Energy Agency. Innovation landscape brief: Utility-scale batteries. Technical Report, Abu Dhabi: International Renewable Energy Agency; 2019.
[151] Thomsen J, Weber C. Getting prices for prosumers right? Incentivizing investment and operation of small-scale PV-battery storage systems through prices, charges and levies. Z Energiewirtsch 2021;46:35–59.
[152] Traywick C, Chediak M, Malik NS, Saul J. The two hours that nearly destroyed texas's electric grid. Bloomberg; 2021.
[153] Giberson M. The 2021 texas power crisis: What happened and what can be done to avoid another one? 2021, URL: https://reason.org/policy-brief/the-2021-texas-power-crisis.
[154] Monika O'Shea RM. Breaking down the Texas winter blackouts: what went wrong? 2021, URL: https://www.woodmac.com/news/editorial/Breaking-down-the-texas-winter-blackouts/full-report/.
[155] Chang A. Why the cold weather caused huge Texas blackouts – a visual explainer. 2021, URL: https://www.theguardian.com/us-news/2021/feb/20/texas-power-grid-explainer-winter-weather.
[156] Ketterer J. The impact of wind power generation on the electricity price in Germany. Technical Report, Leibniz-Institut für Wirtschaftsforschung an der Universität München e. V.; 2012, URL: https://www.ifo.de/publikationen/2012/working-paper/impact-wind-power-generation-electricity-price-germany.
[157] Spisto A. Impact of flexible technologies on end-users' price under different levels of renewable generation. IEEE; 2016, p. 1–5. http://dx.doi.org/10.1109/EEM.2016.7521180, URL: http://ieeexplore.ieee.org/document/7521180/.
[158] Chen T, Guestrin C. XGBoost: A scalable tree boosting system. In: The 22nd ACM SIGKDD international conference. 2016.
[159] Ke G, Meng Q, Finley T, Wang T, Chen W, Ma1 W, Ye1 Q, Liu1 T-Y. LightGBM: A highly efficient gradient boosting decision tree. In: 31st conference on neural information processing systems (NIPS 2017), Long Beach, CA, USA.. 2017.
[160] Dorogush AV, Ershov V, Gulin A. CatBoost: gradient boosting with categorical features support. In: The sixth international conference on learning representations. 2018, arXiv:1810.11363.
[161] Zhang H, Liu Y, Yan J, Han S, Long Q. Improved deep mixture density network for regional wind power probabilistic forecasting. IEEE Trans Power Syst 2020;PP(99):1.
[162] Zhou B, Duan H, Wu Q, Wang H, Meng Y. Short-term prediction of wind power and its ramp events based on semi-supervised generative adversarial network. Int J Electr Power Energy Syst 2021;125:106411.
[163] Liang J, Tang W. Sequence generative adversarial networks for wind power scenario generation. IEEE J Sel Areas Commun 2020;38(1):110–8.
[164] Niu Z, Yu Z, Tang W, Wu Q, Marek. Wind power forecasting using attention-based gated recurrent unit network. Energy 2020;196.
[165] Wang X, Li P, Yang J. Short-term wind power forecasting based on two-stage attention mechanism. IET Renew Power Gener 2019;14(2):297–314.
[166] Khodayar M, Wang J. Spatio-temporal graph deep neural network for short-term wind speed forecasting. IEEE Trans Sustain Energy 2019;10(2):670–81.
[167] Yu B, Yin H, Zhu Z. Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting. In: Proceedings of the twenty-seventh international joint conference on artificial intelligence. 2017, p. 3634–49.

[168] Li Y, Yu R, Shahabi C, Liu Y. Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. In: International conference on learning representations. 2018, URL: https://openreview.net/forum?id=SJiHXGWAZ.
[169] Wu Z, Pan S, Long G, Jiang J, Zhang C. Graph WaveNet for deep spatial-temporal graph modeling. In: Proceedings of the twenty-eighth international joint conference on artificial intelligence. 2019, p. 1907–13.
[170] Zheng C, Fan X, Wang C, Qi J. GMAN: A graph multi-attention network for traffic prediction. In: Proceedings of the AAAI conference on artificial intelligence, Vol. 34. 2020, p. 1234–41.
[171] Ferreira C, Gama J, Matias L, Botterud A, Wang J. A survey on wind power ramp forecasting. Technical Report, USA: Argonne National Laboratory; 2010, URL: https://ceeesa.es.anl.gov/pubs/69166.pdf.
[172] Ouyang T, Zha X, Qin L. A survey on wind power ramp forecasting. Energy Power Eng 2013;5:368–72.
[173] Zack JW. Optimization of wind power production forecast performance during critical periods for grid management. Technical Report, UL, AWS Truewind; 2007, URL: https://aws-dewi.ul.com/assets/AWEA_Windpower_2007_Forecasting.pdf.
[174] Bossavy A, Girard R, Kariniotakis G. Forecasting ramps of wind power production with numerical weather prediction ensembles. Wind Energy 2013;16(1):51–63.
[175] Gallego-Castillo C, Cuerva-Tejero A, Lopez-Garcia O. A review on the recent history of wind power ramp forecasting. Renew Sustain Energy Rev 2015;52:1148–57.
[176] Hannesdóttir A, Kelly M. Detection and characterization of extreme wind speed ramps. Wind Energy Sci 2019;4(3):385–96.
[177] Hannesdóttir A, Kelly M, Dimitrov N. Extreme wind fluctuations: joint statistics, extreme turbulence, and impact on wind turbine loads. Wind Energy Sci 2019;4(2):325–42.
[178] Kelly M, Andersen SJ, Hannesdóttir A. Statistical impact of wind-speed ramp events on turbines, via observations and coupled fluid-dynamic and aeroelastic simulations. Wind Energy Sci Discuss 2021;2021:1–33.
[179] Abuella M, Chowdhury B. Forecasting of solar power ramp events: A post-processing approach. Renew Energy 2019;133:1380–92.
[180] Alam MJE, Muttaqi KM, Sutanto D. A novel approach for ramp-rate control of solar PV using energy storage to mitigate output fluctuations caused by cloud passing. IEEE Trans Energy Convers 2014;29(2):507–18.
[181] Hirata Y, Aihara K. Predicting ramps by integrating different sorts of information. Eur Phys J Spec Top 2016;225:513–25.
[182] Cui M, Zhang J, Florita AR, Hodge B-M, Ke D, Sun Y. An optimized swinging door algorithm for identifying wind ramping events. IEEE Trans Sustain Energy 2016;7(1):150–62.
[183] Greaves B, Collins J, Parkes J, Tindal A. Temporal forecast uncertainty for ramp events. Wind Eng 2009;33(4):309–19.
[184] Ferreira C, Gama J, Botterud A, Wnag J. A survey on wind power ramp forecasting. Technical Report, Argonne National Lab; 2010.
[185] Potter CW, Negnevitsky M. Very short-term wind forecasting for Tasmanian power generation. IEEE Trans Power Syst 2006;21(2):965–72.
[186] IEAWind Task 36. In: C. M, J.W. Z, G. G, editors. IEA Wind Task 36 Recommended Practice on Renewable Energy Forecast Solution Selection. Technical Report, International Energy Agency Technical Cooperation Programm Wind; 2019, URL: https://iea-wind.org/task-36/task-36-publications/recommended-practice/.
[187] Bianco L, Djalalova IV, Wilczak JM, Cline J, Calvert S, Konopleva-Akish E, Finley C, Freedman J. A wind energy ramp tool and metric for measuring the skill of numerical weather prediction models. Weather Forecast 2016;31(4):1137–56.
[188] Bessa RJ, Miranda V, Botterud A, Wang J. 'Good' or 'bad' wind power forecasts: a relative concept. Wind Energy 2011;14(5):625–36.
[189] Möhrlen C, Bessa RJ. Understanding Uncertainty: the difficult move from a deterministic to a probabilistic world. In: Proc. of the 17th international workshop on large-scale integration of wind power into power systems, Stockholm. 2018.
[190] Möhrlen C, Fleischhut N, Bessa J. Insight on human decision-making fromprobabilistic forecast games and experience: an IEA Wind Task 36 initiative. In: Proc. of the 19th intl. workshop on large-scale integration of wind power into power systems as well as on transmission networks for offshore wind farms. Online, 2020.
[191] Fleischhut N, Herzog SM, Hertwig R. Weather literacy in times of climate change. Weather Clim Soc 2020;12(3):435–52.
[192] Göçmen T, Giebel G, Réthoré P-E, Murcia Leon J. Uncertainty quantification of the real-time reserves for offshore wind power plants. In: 15th international workshop on large-scale integration of wind power into power systems as well as on transmission networks for offshore wind power plants, 15th wind integration workshop ; Conference date: 15-11-2016 through 17-11-2016. 2016, URL: http://windintegrationworkshop.org/vienna2016/.
[193] Tsangaris P. Capacity withdrawals in the electricity wholesale market. Springer; 2017, http://dx.doi.org/10.1007/978-3-662-55513-2.
[194] Plate E. P1 – Policy 1: Load-frequency control and performance [C]. Technical Report, European Network of Transmission System Operators for Electricity (ENTSO-E); 2009.
[195] 50Hertz, Amprion, Tennet, TransnetBW. Leitfaden zur Präqualifikation von Windenergieanlagen zur Erbringung von Minutenreserveleistung im Rahmen einer Pilotphase / Guidelines for prequalification of wind parks to provide minutenreserveleistung (MRL) during a pilot phase. Technical Report, German Tranmission System Operators; 2016.
[196] Menemenlis N, Huneault M, Robitaille A. Computation of dynamic operating balancing reserve for wind power integration for the time-horizon 1—48 hours. IEEE Trans Sustain Energy 2012;3(4):692–702.
[197] Murcia JP, Réthoré PE, Dimitrov N, Natarajan A, Sørensen JD, Graf P, Kim T. Uncertainty propagation through an aeroelastic wind turbine model using polynomial surrogates. Renew Energy 2018.
[198] Maniaci DC, Westergaard C, Hsieh A, Paquette JA. Uncertainty quantification of leading edge erosion impacts on wind turbine performance. In: Journal of physics: Conference series, Vol. 1618/5. IOP Publishing; 2020, 052082.
[199] Lange M. On the uncertainty of wind power predictions —- Analysis of the forecast accuracy and statistical distribution of errors. J Solar Energy Eng 2005;127(2):177–84.
[200] Justus D, Brennan J, Bonner S, McGough AS. Predicting the computational cost of deep learning models. In: 2018 IEEE international conference on big data (Big data). IEEE; 2018, p. 3873–82. http://dx.doi.org/10.1109/BigData.2018.8622396.
[201] Sharir O, Peleg B, Shoham Y. The cost of training NLP models: A concise overview. 2020, arXiv preprint arXiv:2004.08900.
[202] Draxl C, Berg LK, Bianco L, Bonin TA, Choukulkar A, Clifton A, Cline JW, et al. The verification and validation strategy within the second wind forecast improvement project (WFIP 2). Technical Report NREL/TP-5000-72553, Golden, CO, USA: National Renewable Energy Laboratory; 2019, URL: https://www.nrel.gov/docs/fy20osti/72553.pdf.
[203] Oreskes N, Shrader-Frechette K, Belitz K. Verification, validation, and confirmation of numerical models in the earth sciences. Science 1994;263(5147):641–6.
[204] Mylne KR. Decision-making from probability forecasts based on forecast value. Meteorol Appl 2002;9(3):307–15.
[205] Asensio MI, Oliver A, Sarrate J. Applied mathematics for environmental problems. Springer; 2021, http://dx.doi.org/10.1007/978-3-030-61795-0.
[206] Frattini P, Crosta G, Carrara A. Techniques for evaluating the performance of landslide susceptibility models. Eng Geol 2010;111(1):62–72.
[207] Stephenson DB. Use of the "Odds Ratio" for diagnosing forecast skill. Weather Forecast 2000;15(2):221–32.
[208] Kumar A, Islam T, Sekimoto Y, Mattmann C, Wilson B. Convcast: An embedded convolutional LSTM based architecture for precipitation nowcasting using satellite data. PLOS ONE 2020;15(3):1–18.
[209] Ashrit R, Mohandas S. Mesoscale model forecast verification during monsoon 2008. J Earth Syst Sci 2010;119(4):417–46.
[210] Prodhomme C, Materia S, Ardilouze C. Seasonal prediction of European summer heatwaves. J Earth Syst Sci 2021.