Overview of the Recently Released Second Version of IEA Wind's Recommended Practices for the Implementation of Renewable Energy Forecasting Solutions

John W. Zack
MESO, Inc.
Troy, NY USA
jzack@meso.com

Corinna Mohrlen
WEPROG
Assens, Denmark
com@weprog.com

Gregor Giebel
DTU Wind Energy
Risø, Denmark
grgi@dtu.dk

***Abstract*: A 4-part document series entitled "Recommended Practices for the Implementation of Renewable Energy Forecasting Solutions" (RP-FSI) was compiled through a collaboration of an international group of experts under Task 36 of the IEA Wind Technology Collaboration Programme (IEA Wind TCP). The ultimate objective of the RP is to assist the forecast users in the selection of an optimal forecast solution for their specific set of applications. The current set of documents, approved by the executive committee of the IEA Wind, is the second version of the RP-FSI and represents an update to the first version that was released in 2019 as an IEA Wind Task 36 Recommended Practice guideline [1]. The need for the RP was based on considerable evidence that even though renewable energy forecasts are now widely employed for operational decision-making the full potential value of the forecasts is frequently not realized for many applications. The group identified four key contributing factors: (1) the specification of the wrong forecast performance objectives in the forecast solution selection process, (2) the use of poorly designed benchmarks or trials to select a forecast solution that is optimal for the user's application, (3) the use of non-optimal evaluation metrics to assess the performance of candidates or existing forecast solutions, and (4) issues with the quality and timeliness of the generation-related and meteorological data provided by the facilities that are the target of the forecasts.**

This paper provides background information about the key issues that motivated the preparation of the RP-FSI and an overview of the content of each of the four parts.

Keywords—forecast solution selection, meteorological sensors, meteorological data quality control, data communication

I. Introduction

The operational use of wind and solar power production forecasts has become widespread in the electric power industry and their benefits for the management of the variability of the generation associated with these renewable energy technologies have been documented in a number of studies (e.g., [2], [3]). However, while the operational use of forecasts has substantially grown over the past decade, there is considerable evidence that the full potential value of the wind and solar forecasts is often not realized in many applications. This is related to a number of factors.

One of these is the specification of the wrong forecast objectives in the forecast solution selection process. For example, a user with a still small, but growing portfolio may need a forecasting tool that helps to get control of the units and to start balance the generation. Making a trial or benchmark with a lot of vendors without the appropriate IT infrastructure and experienced staff may be more expensive and lead to a difficult decision process, especially if the generation is growing while testing.

A second key issue is the use of poorly designed benchmarks or trials to select a forecast solution for the user's application. Benchmarks or trials are costly for both parties. However, poorly designed benchmarks and trials will frequently provide invalid and misleading information to the solution selection process and can result in the selection of a solution that does not provide the best solution for the user's application even though the user thinks it is the best solution based on the data compiled from the benchmark or trail.

A third factor is the use of non-optimal evaluation metrics. A user may correctly specify the performance objective and then conduct a well-designed and executed benchmark or trial but ultimately evaluate the forecasts with metrics that do not measure the performance attributes that are most important to the user's application. This can result in the selection of a solution that is ideal for some other user's application but not for the application of the user conducting the solution selection process.

A fourth issue is the quality, representativeness and timeliness of the data from the forecast target

All Danish contributors thank EUDP for funding from project 64018-0515.

facilities that is provided as input to the forecast production process. However, an important issue that was not addressed in the first version of the RP-FSI document series was the timely availability of high quality meteorological and power generation data from renewable generation facilities to the forecast process. This data is critical (1) for the monitoring of the current meteorological situation and its agreement with or departure from recent forecasts, (2) input into very short-term (minutes to a few hours ahead) forecasting procedures, (3) the training of statistical component of renewable generation prediction systems and (4) evaluation and refinement of forecast methods. It has been the experience of forecast users, researchers and providers that high quality data is often not available to the forecast process in a timely manner. This has been the result of a number of issues including (1) the choice of the number and types of meteorological instruments that are deployed, (2) the selection of the locations at which the meteorological sensors are deployed, (3) the setup, calibration and maintenance of the meteorological sensors, (4) the real-time quality control procedures used to determine if valid data is being produced and (5) the robustness of the real-time data communications infrastructure to gather data from the meteorological and power generation from the sensors and transmit it for use in the forecast product process. Shortcomings or failures in any of these areas can result in a comprehensive set of quality data not being available to the forecast production process with the result that forecast performance will be below the ultimate potential for a state-of-the-art forecast system. In many cases these shortcomings will have their greatest impact at times when forecast accuracy is most critical

In order to address these issues, an international group of experts has worked under the structure of Task 36 of the International Energy Agency's (IEA) Wind Technology Collaboration Program (known as "IEA Wind") [4] to develop a set of recommended practice documents to provide guidance on forecast solution selection. The first phase of the Task 36 activities extended from 2016 through 2018 and produced a first version of the Recommended Practices for Forecast Solution Selection (RP-FSI_v1) documents, which consisted of three parts. The RP-FSI_v1 was released as an IEA Wind Recommended Practice in 2019 [1]. The second phase of IEA Wind Task 36 extended from 2019 to the end of 2021 and resulted in the update of RP-FSI_v1 that was labeled as version 2 (RP-FSI_v2). The major change in version 2 is the addition of a fourth part to the RP-FSI series which will be entitled, "Meteorological and Power Data Requirements for Real-time Forecasting Applications". This part will be included in a second version of the RP-FSI, which will also include revisions and additions to the existing three parts and be available during the first half of 2022. The progress of work on first version 1 and then version 2 of the RP-FSI was summarized in papers and presentations at the 2017 [5], 2018 [6], 2019 [7] and 2020 [8] Wind Integration Workshops (WIW). The title pages of the four parts of RP-FSI_v2 are shown in Fig. 1. This updated set of documents provides guidance on many aspects of the selection of a renewable power forecast solution.

This paper provides an overview of the content of the four parts of RP-FSI_v2 with a focus on the new material that has been added in version 2. The RP-FSI_v2 documents can be downloaded at https://iea-wind.org/task-36/task-36-publications/recommended-practice/.


iea wind

RECOMMENDED PRACTICE FOR THE IMPLEMENTATION OF RENEWABLE ENERGY FORECASTING SOLUTIONS

- Part 1: FORECAST SOLUTION SELECTION PROCESS -

2. EDITION

Draft for Review by the Executive Committee of the International Energy Agency Implementing Agreement

Prepared by IEA Wind Task 36

iea wind

RECOMMENDED PRACTICE FOR THE IMPLEMENTATION OF RENEWABLE ENERGY FORECASTING SOLUTIONS

- Part 2: DESIGNING AND EXECUTING FORECASTING BENCHMARKS AND TRIALS -

2. EDITION

Draft for Review by the Executive Committee of the International Energy Agency Implementing Agreement

Prepared by IEA Wind Task 36

iea wind

RECOMMENDED PRACTICE FOR THE IMPLEMENTATION OF RENEWABLE ENERGY FORECASTING SOLUTIONS

- Part 3: Forecast Solution Evaluation -

2. EDITION

Draft for Review by Executive Committee of the International Energy Agency Implementing Agreement

Prepared by IEA Wind Task 36

iea wind

RECOMMENDED PRACTICE FOR THE IMPLEMENTATION OF RENEWABLE ENERGY FORECASTING SOLUTIONS

- Part 4: Meteorological and Power Data Requirements for real-time forecasting Applications-

1. EDITION

Draft for Review by the Executive Committee of the International Energy Agency Implementing Agreement

Prepared by IEA Wind Task 36

Fig. 1. Title pages of the second version of the IEA Wind Recommended Practices for the Implementation of Renewable Energy Forecasting Solutions.

II. RP-FSI PART 1: FORECAST SOLUTION SELECTION PROCESS

Part 1 of RP-FSI_v2 addresses the process of selecting an optimal wind forecasting solution for a specific set of applications. This is intended to provide guidance for the **design of an economically viable process that will maximize the proba-**

bility of obtaining an optimal forecast solution for a user's applications. The document is divided into two core sections. The first is a discussion of the "big picture" issues that should be considered before starting the design of a selection procedure. The second is the presentation and discussion of a Decision Support Tool (DST) that steps through the issues that should be considered during the design of a forecast solution selection process. The following two subsections summarize some of the key points in Part 1.

A. *Initial Considerations*

The first step in the forecast solution selection process is to define the objectives of the forecasting application. For example, very different forecasting strategies are needed for the balancing of supply and demand on a system with a significant penetration of renewable generation versus the selling of generated electricity in the power market. In the first application, extremes must be considered and risks estimated, where probabilistic forecast solutions need to be considered; mean error scores are not that important and misleading forecast incentives. Large errors and the associated uncertainty of the forecasts in times, where large errors occur, are most significant, as they could potentially lead to lack of available balancing power. In the second case, it is important to know the uncertainty of the forecast and to use a forecast whose errors have the lowest correlation with other forecasts in the market.

When choosing a forecast solution, understanding the underlying requirements is key. It is not enough to ask for a specific forecast type without specifying the target objective. For this reason, defining the objective is most important. Furthermore, if there is no knowledge in the organization regarding the techniques required to reach the objective, it is recommended to start with a "request for information" (RFI) from a set of forecast providers and thereby gain an understanding and overview of the various existing solutions and their capabilities.

Once the applications objectives are clearly and specifically defined the next step is a detailed specification of the desired outcome of the solution selection process. The key questions to be asked are:

- What specific forecast information is needed for the application?
- What infrastructure and resources does the user have to support in the solution selection process and ultimately the implementation of the forecast solution?
- What criteria will be used to determine which is the best solution for the target application?
- Do we need a probabilistic forecast solution?
- What forecast services are available from solution providers?
- What level of customization is available from solution providers?
- What is the cost range of the available forecasts?
- What is the historical performance level of the available solutions?

The answers to these questions should play a major role in defining the scope of the selection process. The answers to the first four questions define the requirements and limitations of the desired forecast solution. The last four questions provide information about what is available in the forecasting marketplace. The degree of alignment between the user's requirements and limiting factors and what is available in the marketplace should be the basis for the formulation of a selection process.

For example, a considerable amount of time and resources can be wasted by conducting trials or benchmarks (to determine the best performing solution for the user's application) that are not aligned with the user's requirements, too much simplified and maybe also planned and conducted by personnel who are not experienced with these issues and have little or no relation to the real-time environments, where the forecasts will later be used. In order to avoid this, it is recommended that the user compile a "requirements list" at the start of the selection process. An example of a requirements list is presented in Part 1.

In some cases, it can be beneficial to test solutions prior to implementation. The difficulty with this approach lies in the quality of the information from the tests, especially, when they are based on a short time period or an over-simplification of the real-world environment, where the forecasts later will be used. In many cases they do not answer the questions an end-user needs to answer. This is because such tests can usually only be carried out with a sub-set of the forecasts in comparison to the real-time application but still require significant resources to conduct. For such cases, this guideline provides other methods for an evaluation of alternative forecast solutions. The pitfalls and challenges with trials and/or benchmarks are addressed in Part 2.

B. *Decision Support Tool*

From an end-user perspective, it is a non-trivial task to decide which path to follow in the selection and implementation of a forecasting solution for a specific application. In most user situations there are multiple stakeholders involved in the decision-making process. A relatively straightforward way to decide on the path is to use a decision support tool. Fig. 2 depicts the decision support tool presented in Part 1 intended to support the high-level decisions of managers and non-technical staff when establishing a business case for a forecasting solution. The high-level thought construct Fig. 2 is targeted to assist in considering the required resources and involvement of staff in the decision process. The decision tool is constructed to begin with initial considerations to establish a "Forecast System Plan". There are cross-references in the decision tool and referrals to alternative decision streams, depending on the answer at each step of the decision flow.

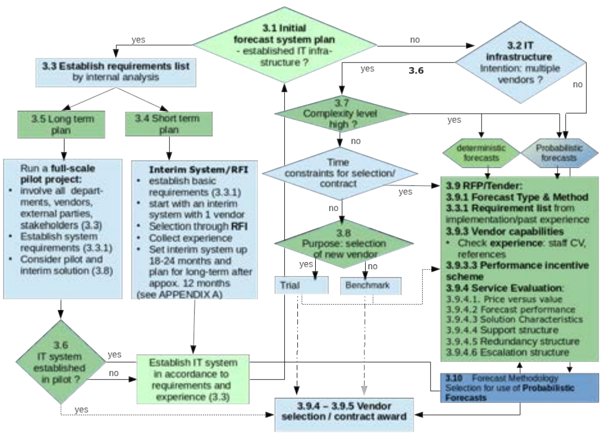

Fig. 2. A decision support tool for the planning and design of a variable generation forecast solution. The numerical citations in the flow chart objects refer to the sections in Part 1 document in which issues associated with the referenced topic are discussed

III. RP-FSI PART 2: DESIGN AND EXECUTION OF BENCHMARKS AND TRIALS

The second part of RP-FSI_v2 provides guidance for the design and execution of benchmarks or trials (B/T). For the purposes of the RP-FSI documents, a **benchmark** is defined as an exercise conducted to determine the features and quality of renewable energy forecast systems or methods such as those used to produce wind or solar power forecasts. The exercise is normally conducted by an institution or their agent and multiple participants including private industry forecast providers or applied research academics. A **trial** is an exercise conducted to test the features and quality of (operational) renewable energy forecast solutions. This may include one or more participants and is normally conducted by a private company for commercial purposes. A trial is a subset of benchmarks. A trial may be part of the process to select an initial, replacement or additional forecast solution providers or part of a periodic evaluation process for an existing forecast solution. In any of these cases the fundamental objective of the trial is to determine which solution represents the best value for a user's application.

While a B/T may intuitively seem to be the best approach to identify the best forecasting solution for a specific application, in practice, the use of a trial as part of the solution selection process is not always the best option and has a number of limitations to go along with its benefits. The trade-off between the limitations and benefits of a trial should be carefully considered before a decision is made to conduct a trial. The Part 2 document addresses the benefits and limitations of a trial for a range of scenarios, including the challenges that come with testing and comparing probabilistic forecast solutions.

The structure of Part 2 is based on the three primary phases of a B/T: (1) preparation, (2) execution and (3) evaluation and decision-making. Some of the key issues in each of these phases are summarized in a cheat-sheet and briefly explained in the following three subsections.

A. *Preparation*

The preparation phase is the period before the start of the forecasting activities during which the structure and protocols of the B/T are formulated by the B/T operator and disseminated to the solution providers that will participate in the B/T. The decisions and actions during this phase often have a very large impact on the ultimate quality and therefore the value of the information obtained from the B/T. The use of information from a poorly designed B/T is often worse than not conducting a B/T since this information is typically viewed as an objective basis for making a selection of a forecast solution.

There are a number of key decisions that will determine the complexity and therefore the level of effort and cost of a trial. It will also play a major role in determining the quality of the information produced by the B/T. Part 2 includes a detailed table that summarizes the key attributes of a trial that have an impact on both the cost of a B/T and the quality of the information produced by the exercise.

B. *Execution*

The execution phase is the period during which forecasts are produced by the participating solution providers and submitted to the B/T operator. In a live (or real-time) trial, the providers should receive near-real-time data for the forecast target facilities from the B/T operator and submit forecast data on a prescribed schedule to IT platforms designated and controlled by the B/T operator. In a retrospective trial, the providers should receive a historical set of data for the target facilities (for statistical model training purposes) and produce forecasts for a specified evaluation period (that does not overlap with the historical data sample period).

In a well-designed B/T, most of the communication between the trial operator and the solution providers should be during the preparation phase. However, issues often arise during a trial (especially in live trials). It may be helpful to all B/T participants to establish an open forum during the first part of the live B/T period to provide a way to effectively and uniformly resolve all issues. However, it is strongly recommended that if any attributes of the B/T are changed at any point during the live part of the B/T, the changes should be communicated to all participants immediately as they might require action of the solution providers.

C. *Evaluation and Decision-making*

Intuitively, one might expect the evaluation and decision-making phase to begin after all the forecast data has been gathered from the solution providers at the end of the live or retrospective B/T periods. However, in a well-designed B/T that should not be the case. The forecast evaluation process should begin soon after the first forecasts have been received from the solution providers. This will enable the B/T operator to assess its evaluation design and results production protocols (e.g. software to calculate error metrics, graphical or tabular displays to view results, etc.) before the end of the B/T execution period and possibly make adjustments to the evaluation or forecast submission process to mitigate issues that may compromise the quality of the information obtained from the B/T.

If an interim report was provided during the B/T, then the final report can either be an updated version of the validation report expressing the bulk metrics or appended month-by-month forecast validation results. For transparency and to promote further forecast improvements, it is recommended that the B/T operator shares the anonymized forecast results from each solution provider at the time-interval frequency that forecasts were being made at (e.g., hourly). This will help solution providers discover where forecasts are similar or different from the competition which may spawn improved methodologies.

D. *Evaluation and Decision-making*

Forecast service providers who have participated in numerous trials over the past decade have indicated that there are a number of design and execution problems that have repeatedly appeared in trials during this period. The consequences of errors and omissions in trials are often underestimated. However, if results are not representative, the efforts that have gone into a trial can effectively be wasted. Some of these common pitfalls can be expensive to the operator, because they result in placing the operator in a position of making a decision without having truly objective and representative information. A list of significant issues that have frequently been encountered is included in this part.

A specific section dealing with the challenges of probabilistic forecast B/T's has also been added. Here, the recommendations are especially done on so-called "event forecasting solutions" such as ramping reserves, high-speed shutdown events, unit commitment, and the development of cost functions for the decision-making with and evaluation of probabilistic forecasts.

IV. RP-FSI PART 3: FORECAST EVALUATION

Part 3 of the document series provides guidance on the evaluation of forecasts. The evaluation process is a large component of the forecast solution selection process, if a benchmark or trial is conducted as part of the process. An evaluation is also an important component of an ongoing performance assessment program.

The Part 3 document is composed of four core sections. These provide (1) a description of the general factors that determine the evaluation uncertainty, (2) an overview of the uncertainty associated with the data from the forecast target site, (3) a discussion of the importance of choosing appropriate metrics to evaluate forecast performance and (4) a compilation of the recommended best practices for forecast evaluation. In the updated version of the document, specific sections have been added that describe metrics and assessment methods for probabilistic forecasts, recommendations on verification and a typical verification use-case for probabilistic forecast solutions.

The following subsections present a summary of the key points in each core components of Part 3.

A. *Overview of Evaluation Uncertainty*

The first component of RP-FSI Part 3 provides an overview of the general factors that differentiate the level of uncertainty among forecast evaluation exercises. The focus is on three key points:

- All evaluations of forecast solutions have a degree of uncertainty, which is associated with the three core attributes of the evaluation process: (1) representativeness, (2) significance and (3) relevance.
- A carefully designed and implemented evaluation process that considers these three attributes can minimize the uncertainty and yield the most meaningful results.
- Disregarding these issues can lead to uncertainty that is so high that the conclusions of the evaluation are meaningless and no valid information is available for decision-making.

Representativeness can be defined as the relationship between the results of a forecast performance evaluation and the performance that is ultimately obtained in the operational use of a forecast solution. It essentially addresses the question of whether or not the results of the evaluation are likely to be a good predictor of the actual forecast performance that will be achieved for an operational application. There are many factors that influence the ability of the evaluation to be a good predictor of future performance. Four of the most crucial factors are: (1) size and composition of the evaluation sample, (2) quality of the data from the forecast target sites, (3) the formulation and enforcement of rules governing the submission of forecasts, (4) availability of a complete and consistent set of evaluation procedure information.

Significance refers to the ability to differentiate between performance differences that are due to noise (quasi-random processes) in the evaluation process and those that are due to meaningful differences in skill among forecast solutions. Performance differences that stem from noise have basically no meaning and will not represent the likely performance in a long-term operational application of a solution. Real performance differences should be stable and should not change, if an evaluation process is repeated, e.g., one year later. A certain degree of noise is inevitable in every evaluation but both noise minimization and awareness of the uncertainty are crucial for reliable decision-making.

Relevance is defined as the degree of alignment between the evaluation metrics used for an evaluation and the true sensitivity of a user's application(s) to forecast error. If these two items are not well aligned then even though an evaluation process is representative and the results show significant differences among solutions, the evaluation results may not be a relevant basis for selecting the best solution for the application.

A. *Measurement Uncertainty*

The second section of Part 3 provides an overview of the factors that contribute to measurement uncertainty, which is a part of the representativeness attribute. The key points are:

- Measurements from the forecast target facilities are crucial for the forecast production and evaluation process and therefore much attention should be given to how data is collected, communicated and quality controlled
- Collection and reporting of measurement data requires strict rules and formats, as well as IT communication standards in order to maximize its value in the forecasting process; standards and methods for collecting and reporting data from multiple sources are noted in RP-FSI Part 3
- An effective quality control process is essential since bad data can seriously degrade forecast performance; standard quality maintenance and control procedures have been documented and some are noted in this section of Part 3.

B. *Targeted Evaluation of Forecast Performance*

The third component of the Part 3 document addresses the importance of employing an appropriate set of metrics in the evaluation process. A number of publications have compiled lists of metrics for the evaluation of a broad range of attributes of wind and solar power generation forecasts and have provided some examples of their application (e.g., [5]). However, there is little guidance available for the selection of the most relevant set of metrics for a specific application.

The relevance of different aspects of forecast performance depends on the user's application(s). For instance, one user may be concerned with the size of typical forecast errors, while another may only be concerned with the size and frequency of particularly large errors. This component of RP-FSI Part 3 provides a description of the key issues in evaluating specific attributes of forecast performance with a focus on: (1) the relationship between forecast performance attributes and widely used error metrics and (2) metric-based forecast optimization (i.e., configuring the forecast system for the best performance for a specific metric).

An example of how one's selection of a performance metric and therefore what forecast performance attribute is measured can determine one's perspective on what is considered the best forecast is presented in Fig. 3. Three forecasts for a wind ramp event are depicted. Despite being the only forecast (the left plot) that correctly predicts the ramp rate and duration, the forecast with a phase error has the largest MAE. Thus, this could be considered the worst (i.e., highest MAE) or best forecast (i.e., lowest ramp rate or duration forecast error) depending on one's perspective.

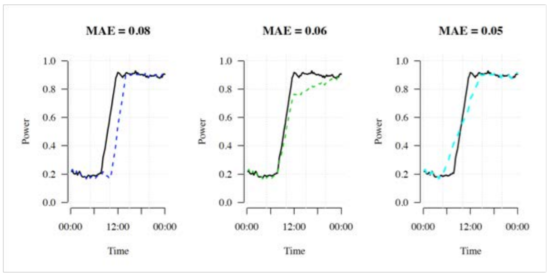

Fig. 3. Examples of 3 types of ramp forecast error. Actual power is shown as solid black lines, forecasts as colored dashed lines. Three types of ramp event forecast errors are depicted: (1) phase or timing error, (2) amplitude error and (3) ramp rate error. The associated MAE (fraction of capacity) is shown at the top of each chart.

C. *Best Practices for Forecast Evaluation*

The fourth component of the RP-FSI Part 3 document provides an overview of the recommended practices for forecast evaluation. The key points of this component are:

• Forecast verification is subjective; it is important to understand its limitations.

• Verification has an inherent uncertainty due to its dependence on the evaluation data set

• Evaluation should contain a set of metrics to measure a range of forecast performance attributes

• Evaluation should include a "cost function" i.e. the metric set should assess the value of the solution to the application.

V. RP-FSI PART 4: METEOROLOGICAL AND POWER DATA REQUIREMENTS FOR REAL-TIME FORECASTING APPLICATIONS

The new Part 4 of the RP-FSI is designed to provide background information on meteorological instrumentation, their recommended setup, maintenance, quality control and use in the real-time environment of wind and solar power generation plants. The focus is on the impact that measurement-data-related decisions that affect the characteristics (e.g. availability, quality, representativeness, timeliness, etc.) of the data available from a wind or solar generation facility ultimately have on wind or solar generation forecast performance

A. *Unique Requirements of Short-term Forecasting*

Part 4 begins with a look at the meteorological and generation data requirements of short-term generation forecasting and how they are different than those for other wind or solar applications that also require data gathering. This is a key issue because it is often not satisfactory to simply use data from a resource assessment campaign for project development or a hardware performance assessment project for short-term forecasting purposes.

For example, the near real-time collection and dissemination of data from a generation facility is often critical for some short-term forecasting applications but is not typically an issue for non-real-time applications such as resource assessment. Another example is that measurements during extreme weather events (high winds, heavy precipitation etc.) are typically crucial for forecasting since these events can have significant impacts on grid systems. However, these infrequent events typically have little impact on a resource assessment. A practical issue associated with this factor is that some sensors are often unable to reliably provide data under these types of conditions but otherwise perform well and may even have advantages for other types of applications such as resource assessment.

B. *Meteorological Instrumentation*

The second section of Part 4 provides an overview of the meteorological sensor options and characteristics of each sensor that can have an impact on generation forecast performance. The measurement of all variables that have an impact of short-term forecasting performance is considered.

The wind component of this section looks at the salient characteristics of sensors related to wind generation forecasting on (1) meteorological masts, (2) passive and active remote sensing devices and (3) nacelle-mounted platforms. As noted previously the focus is on the attributes of each type of sensor that can impact forecast performance.

The analogous solar component of this section examines the characteristics of ground-based sensors that measure incoming solar radiation as well as other meteorological parameters (e.g. temperature, wind) that can impact solar generation. However, it also considers alternative remote sensing devices for the measurement/estimation of solar radiation such sky-imagers and satellite-based sensors. These have the advantage of providing not only an estimate of the solar radiation at a point but also information about the spatial pattern of solar radiation variations due to clouds that can be valuable for the forecasting of solar generation on a minutes to a couple of hours time frame.

C. *Power Measurements*

Although it is not produced by a meteorological sensor, power measurements typically play in an important role in short-term generation forecasting since it is the target variable for the forecasts. Therefore, knowledge on recent trends and possibilities in handling power generation measurements can be extremely valuable for very short term forecasting and historical generation data alike, which is usually employed for the training of statistical/machine learning forecast models. This section addresses a number of potential issues with power generation data that can impact forecast performance and how to manage or mitigate these issues.

A key issue is that power production is affected by non-weather effects, such as the control actions of system operators or plant operators. Thus, it is important to have additional data that specifies the operating status (outages, curtailments, available active power, etc.) of the facility for each data interval so that the generation data can be properly interpreted in the forecast process. Another less significant but important issue is the point at which the generation is measured, which may result in differences between the measurements (e.g. due to line losses or internal facility power use) and what is designated as the "target" for the forecasts.

D. *Measurement Setup and Calibration*

The fourth section of Part 4 addresses the selection, setup, calibration and maintenance of meteorological sensors. The discussion of the selection of instrumentation builds upon the sensor characteristics described in Section 2 and provides guidance on the benefits or drawbacks of different sensor options for short-term forecasting.

The "setup" section addresses how to select the vertical and horizontal location of a sensor to best represent the conditions experienced by the entire generation facility. The objective is to obtain sensor data that has the highest correlation with the weather-based variations in generation.

The last component of this section considers the calibration, inspection and maintenance of the sensors. This is a frequently overlooked factor that can result in subtly bad data that is difficult for data quality control algorithms to identify.

E. *Assessment of Instrumentation Performance*

The fifth section addresses the assessment of sensor performance and the associated issue of quality control of the data that is produced by the sensors and used by short-term forecasting applications. While the impact of low quality data can be offset with longer collection periods for non-real-time applications such as plant assessment and configuration, resource assessment, etc., this is not possible in real-time environments, where the data has to be quality checked at the time of retrieval and either used in the application or discarded.

From a real-time forecasting perspective, bad data is often worse for a forecast model than no data. State-of-the-art forecasting systems usually have automatic algorithms that are very effective at discarding data that is out of range or obviously unrepresentative in other ways. However, it is often difficult for such an algorithm to identify "bad data" that looks realistic. This type of "bad data" can have a significant negative impact on short-term forecast performance. The section of Part 4 describes the challenges and mitigation strategies for the quality assessment and control of measurement data.

F. *Measurement Best Practice Recommendations*

The final section of Part 4 provides a concise summary of the recommended best practices that can facilitate the availability of high quality meteorological and power data for short-term forecasting. The summary of recommendations provides an easily accessed reference of key practices for forecast users and is built upon the material presented in the previous sections.

VI. Summary

A second version of IEA Wind's Recommended Practice for the Implementation of Forecasting Solutions (RP-FSI_v2) has been released in early 2022. It includes a new Part 4 that addresses the recommended practice for meteorological and power measurements for short-term forecasting and also several substantial additions, for example the inclusion of solar forecast evaluation, probabilistic forecast solution select, benchmarking and evaluation, and revisions to the original three parts of version 1. The work on the RP-FSI will continue under the new IEA Wind Task 51 entitled “Forecasting for the Weather Driven Energy System”. The new task will address a number of a broader range of electricity-supply-and-demand-related forecasting issues.

ACKNOWLEDGMENT

The participants of IEA Wind Task 36 have made significant contributions to the content of the RP-FSI parts 1 through 4. The Danish authors were funded by the Danish EUDP project IEA Wind Task 36 Forecasting Danish Consortium, Grant Number 64018-0515. Task 36 was led by “co-operating agents” Gregor Giebel of Technical University of Demark (DTU) and Will Shaw of Pacific Northwest National Laboratory (PNNL) of the United States Department of Energy.

REFERENCES

[1] IEA Wind Task 36, 2019: Recommended Practice on Renewable Energy Forecast Solution Selection Part 1-3, Online: https://iea-wind.org/task-36/task-36-publications/recommended-practice/

[2] Hodge, B-M, A. Florita, A., J. Sharp, M. Margulis, and D. Mcreavy, 2015: The Value of Improved Short-Term Wind Power Forecasting. Technical Report NREL/TP-5D00-63175. National Renewable Energy Laboratory, Golden, Colorado, USA. Available from http://www.osti.gov/scitech.

[3] Orwig, K., B.-M. Hodge, G. Brinkman, E. Ela, M. Milligan, V. Banunarayanan, S. Nasir, and J. Freedman, 2012: “Economic evaluation of short-term wind power forecasts in ERCOT: Preliminary results,” Proceedings of the 11th International Workshop on Large-Scale Integration of Wind Power into Power Systems. Energynautics GmbH, Darmstadt Germany

[4] Giebel, G., W. Shaw, H. Frank, C. Draxl, J. Zack, P. Pinson, C. Möhrlen, G. Kariniotakis, R. Bessa, 2021: IEA Wind Task 36 Forecasting: Proceedings of the 16th Interntiaonal Workshop on Large-Scale Integration of Wind Power into Power Systems (Berlin, Germany).

[5] Giebel, G., J. Cline, H. Frank, W. Shaw, B.-M. Hodge, C. Draxl, P. Pinson,J. Messner, G. Kariniotakis, C. Möhrlen, 2017: Games and other news from IEA Wind Task 36 Forecasting for Wind Energy: Proceedings of the 20th Interntiaonal Workshop on Large-Scale Integration of Wind Power into Power Systems (Berlin, Germany)..

[6] Möhrlen, C., J. Lerner , J. W. Messner , J. Browell, A. Tuohy, J. Zack, C. Collier, and G. Giebel, 2018: IEA Wind Recommended Practices for the Implementation of Wind Power Forecasting Solutions – Part 2 & 3: Designing and Executing Forecasting Benchmarks and Evaluation of Forecast Solutions. Proceedings of the 17th International Workshop on Large-Scale Integration of Wind Power into Power Systems (Vienna, Austria). Energynautics GmbH, Darmstadt Germany.

[7] Zack, J., C. Möhrlen, and C. Draxl, 2019: IEA Wind Task 36 Session Topic 4: Request for Feedback on Version 1 of the Recommended Practices for Forecast Solution Selection. Proceedings of the 18th International Workshop on Large-Scale Integration of Wind Power into Power Systems (Dublin, Ireland). Energynautics GmbH, Darmstadt Germany.

[8] Zack, J., and C. Möhrlen, 2020: IEA Wind Task 36: Practical Application Examples from the Recommended Practices for Forecast Solution Selection. Proceedings of the 20th International Workshop on Large-Scale Integration of Wind Power into Power Systems (virtual). Energynautics GmbH, Darmstadt Germany.