IEA Wind Task 36: Practical Application Examples from the Recommended Practices for Forecast Solution Selection

John W Zack
UL Services Group LLC
Albany, NY, USA
john.zack@ul.com

Corinna Möhrlen
WEPROG
Assesns, Denmark
com@weprog.com

Abstract—**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. However, there is considerable evidence that the full potential value of the wind and solar forecasts in many applications is often not realized. This is typically related to three 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 for the user's application and (3) the use of non-optimal evaluation metrics to assess the performance of candidates or existing forecast solutions.**

This issue was addressed by a group of experts within the scope of the International Wind Energy Agency (IEA) Wind Task 36. This group prepared an IEA Recommended Practice on Forecast Solution Selection (RP-FSS), which provides guidance on the selection of a new, alternative or additional forecast solution, the execution of a forecasting trial or benchmark, and the evaluation metrics and methods used to assess forecast quality. The RP is composed of three documents. The first part, "Forecast Solution Selection Process", deals with the selection and background information necessary to collect and evaluate when developing or renewing a forecasting solution. The second part of the series, "Benchmarks and Trials", offers recommendation on how to best conduct benchmarks and trials in order to evaluate the relative performance and the "fit-for-purpose" of forecasting solutions. The third part, "Forecast Evaluation", provides information and guidelines for the effective evaluation of the performance of forecasts and forecast solutions. The RP documents are intended to provide guidance to stakeholders who are seeking to initiate or optimize a forecasting solution that will maximize the benefit for their specific applications.

Initial feedback from users of the RP-FSS document series indicated that there was a strong desire for examples that illustrated the key points of the RP-FSS, their practical implementation and their impact on the identification of an optimal forecast solution. A number of examples have been constructed now, some in collaboration with an emerging online and open source tool for forecast evaluation called the Solar Forecast Arbiter (SFA)

The paper provides (1) insight into the reasons why the full potential value from existing forecast solutions is often not realized and (2) a brief overview of the contents of the three IEA RP documents and where to obtain but focuses on (3) practical examples of key points from the RP-FSS.

Keywords—*forecast solution selection, forecast benchmarks and trials, optimization of forecast value*

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., [1], [2]). 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 often related to three factors.

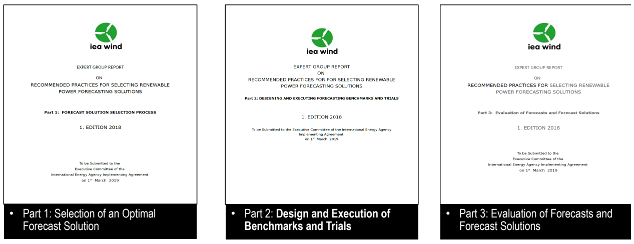
The first is the specification of the wrong forecast performance objectives in the forecast solution selection process. For example, a user may implicitly or explicitly state that the objective is to minimize the typical or average error of the forecast. However, the user's application may be more sensitive to large errors or errors associated with specific types of events. While it would be ideal to have a system that produces perfect forecasts in all situations, the reality is that the error characteristics of forecasts are linked to the way in which they are optimized. For example, a forecast system that is optimized to minimize the average error will generally not produce the best forecasts of anomalous events.

A second key issue is the use of poorly designed benchmarks or trials to select a forecast solution for the user's application. 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 trail 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.

The result of these and other flaws in the forecast solution selection process is that the value of renewable energy forecast information is reduced below its full potential for both the specific users and implicitly for a broad range of stakeholders in the energy community since it results in higher integration costs for wind and solar electricity generation and also inhibits a higher penetration level for these generation resources on grid systems

In order to address this issue, 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") to develop a set of three recommended practices documents to provide guidance on forecast solution selection. The IEA is an independent international entity that is composed of 30 member countries and 8 associate countries. Information about the IEA may be found at https://www.iea.org/about/. IEA Wind Task 36 is a focused activity that facilitates the interaction of international experts to address issues associated with short-term wind power forecasting. The first phase of the Task 36 activities extended from 2016 through 2018. The second phase began at the start of 2019 and will extend through the end of 2021. Information about the past, current and future activities of Task 36 can be obtained from the task's web portal at ieawindforecasting.dk. The interim progress of this activity was summarized in papers and presentations at the 2017 [3], 2018 [4] and 2019 [5] Wind Integration Workshops (WIW). An initial version of a set of three Recommended Practices for Forecast Solution Selection (RP-FSS) documents was completed at the end of the first phase of Task 36 and presented in a paper at WIW19 [6]. The title pages of these three documents are shown in Fig. 1. This set of documents provides guidance on almost all aspects of the selection of a renewable power forecast solution. The first document, "Forecast Solution Selection Process", deals with the selection and background information necessary to collect and evaluate when developing or renewing a forecasting solution. The second document, "Benchmarks and Trials", of the series offers recommendation on how to best conduct benchmarks and trials in order to evaluate the relative performance and the "fit-for-purpose" of alternative forecasting solutions. The third part, "Forecast Evaluation", provides information and guidelines for the effective evaluation of the performance of forecasts and forecast solutions. The next three sections of this paper provide an overview of the contents and key points addressed in each of these documents.


iea wind

EXPERT GROUP REPORT
ON
RECOMMENDED PRACTICES FOR SELECTING RENEWABLE
POWER FORECASTING SOLUTIONS

Part 1: FORECAST SOLUTION SELECTION PROCESS

1. EDITION 2018

To be Submitted to the
Executive Committee of the
International Energy Agency Implementing Agreement
on 1st March 2019

- Part 1: Selection of an Optimal Forecast Solution

iea wind

EXPERT GROUP REPORT
ON
RECOMMENDED PRACTICES FOR FOR SELECTING RENEWABLE
POWER FORECASTING SOLUTIONS

Part 2: DESIGNING AND EXECUTING FORECASTING BENCHMARKS AND TRIALS

1. EDITION 2018

To be Submitted to the Executive Committee of the International Energy Agency
Implementing Agreement
on 1st March 2019

- Part 2: **Design and Execution of Benchmarks and Trials**

iea wind

EXPERT GROUP REPORT
ON
RECOMMENDED PRACTICES FOR SELECTING RENEWABLE
POWER FORECASTING SOLUTIONS

Part 3: Evaluation of Forecasts and Forecast Solutions

1. EDITION 2018

To be Submitted to the
Executive Committee of the
International Energy Agency Implementing Agreement
on 1st March 2019

- Part 3: Evaluation of Forecasts and Forecast Solutions

Fig. 1. Title pages of the three IEA Wind Recommended Practices for forecast solution selection documents.

II. PART 1: FORECAST SOLUTION SELECTION PROCESS

The first of the three RP-FSS documents 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 probability 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 these core components of RP-FSS 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; mean error scores are not that important. Large errors 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 are least correlated with other forecasts in the market.

When choosing a forecast solution, understanding the underlying requirements is key. It is not sufficient 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 the solution selection process and the implementation of the forecast solution?
- What criteria will be used to determine which is the best solution for the target application?
- What forecast services are available from solution providers?
- What level of customization is available?
- 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 three 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 lot of time and resources can be wasted by conducting trails or benchmarks (to determine the best performing solution for the user's application) that are not aligned with the user's requirements and also planned and conducted by personnel who are not experienced with these issues. 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 RP-FSS 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, especially, when the tests are based on a short time period. In many cases they do not answer the questions an end-user needs to answer. This is because such tests are usually simplified 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 also addressed in Part 2.

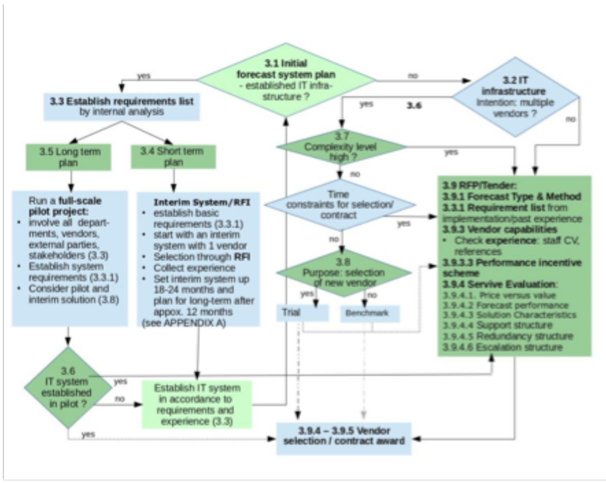


Fig. 2. A decision support tool for the planning and design of a variable generation forecast solution. The numerical citations in the flow chart objects refer to the sections in RP-FSS Part 2 document in which associated with the referenced topic are discussed

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 shows a decision support tool aimed to high-level decisions of managers and nontechnical staff when establishing a business case for a forecasting solution. The high-level thought construct shown in Fig. 2 is targeted to assist in considering the required resources and staff involvement in the 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.

III. PART 2: DESIGN AND EXECUTION OF BENCHMARKS AND TRIALS

The second of the three RP-FSS documents provides guidance for the design and execution of benchmarks or trials (B/T). For the purposes of the RP-FSS 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 trail 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 these cases the fundamental objective of the trial is to determine which solution is 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 an application, the use of a B/T as part of the solution selection process is not always the best option and has a number of limitations as well as benefits. The trade-off between the limitations and benefits of a B/T should be carefully considered before a decision is made to conduct a B/T. The Part 2 document addresses the benefits and limitations of a B/T in a range of scenarios. The structure of Part 2 is based on the three phases of a B/T: (1) preparation, (2) execution and (3) evaluation and decision-making. Some key issues are summarized in the following 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 had 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. RP-FSS Part 1 summarizes the key attributes of a trial that have an impact on both the cost and quality of the information produced by a B/T.

B. *Execution*

The execution phase is the period during which forecasts are produced and submitted by the participating solution providers. In a 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 dataset 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).

In a well-designed B/T, most of the communication between the trial operator and the solution providers should be during the preparation period. However, issues often arise during a trial (especially a live trial). 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 on the part 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 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 share the anonymized forecast results from each solution provider at the time-interval frequency that forecasts were being made (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 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 trail 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 some of the significant issues that have frequently been encountered is presented in RP-FSS Part 2.

IV. 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 but an evaluation is also an important component of an ongoing performance assessment program.

The recommendations for evaluation in RP-FSS Part 3 are based on the following set of principles:

- Evaluation is subjective, i.e. it is important to understand the limitations of a chosen metric
- Evaluation has an inherent uncertainty due to its dependence on the evaluation dataset
- Evaluation should contain a set of metrics in order to measure a range of forecast performance attributes
- Evaluation should reflect a “cost function”, i.e. the metric combinations should provide an estimate of the value of the solution for the specific target applications(s)

The formulation and use of an application-specific cost function is perhaps the most critical component of the evaluation process since it determines how well the evaluation is aligned with the needs of the application. In response to this statement many users ask, “What is a cost-function and how can it be implemented?”

A cost function can be defined as the minimum cost of producing a minimum level of outcome from a specific set of input. In the case of evaluating forecasts this means that the evaluation in fact is an optimization problem, as the objective is to find the “minimum level of outcome” for a specific input. If we turn around the objective in this way, we take account for the subjective part of the evaluation: we try to find the metrics that provide the best way of optimizing our forecasts!

This can be done by constructing an evaluation framework that links the needs of the application to a set of performance metrics that assess different but application-relevant attributes of forecast performance. Such a framework is an effective way to also mitigate the so-called "relevance" issues associated with the tuning (optimization) of forecasts to target metrics that may or may not be optimal indicators of value for an end user's application.

Errors in forecasting are inevitable. The primary objective is, of course, to minimize the magnitude of the error. However, a second objective should always be to shape the error distribution in ways that are beneficial to a specific application.

There are three fundamental components of a comprehensive evaluation procedure. These are described in the following subsections.

A. *Deep Analysis of the Prediction Errors*

An analysis of the prediction errors provides considerable insight into the characteristics of forecast performance as well as information that can allow users to differentiate situations in which forecasts are likely to be trustworthy from those that are likely to produce large errors.

B. *Definition of the Optimization Target*

A detailed specification of the forecast optimization target is a crucial component of the evaluation framework

There are three primary components to this specification: (1) definition of the types of errors are most "costly" in terms of economics, security, penalties, (2) identification of which time frames are most important or subject to penalties, (3) the mapping of the optimization target to the appropriate evaluation metric

C. *Setup an Evaluation Framework*

The evaluation framework should contain a set of metrics, time frames and targets that are appropriate for an application. Table III maps some typical forecast applications to appropriate optimization targets and associated metrics.

Details on evaluation metrics can be found in [10] and [9], explanation of specific metrics can be found in [7], [8], and [11] for deterministic forecasts inclusive of solar forecasting and for probabilistic forecast metrics in [12]. Information about significance tests can be found in [13]

In a best practice approach, the evaluation framework should reflect (1) the importance of forecasts in their role in business processes and (2) provide incentives for the forecast service provider to generate forecasts that fit the outlined (and evaluated) purpose. As a minimum requirement when establishing such an evaluation framework the following four factors should be considered.

1) *Definition of the forecast framework*

The first consideration should be a detailed specification of what is desired from the forecast. This should include the specification of: (1) the forecast application (2) the key forecast time frames, and (3) a ranking of the relative importance of the forecast performance attributes

2) *Evaluate a Clearly Defined Set of Forecasts*

The evaluation should be conducted on a well-defined set of representative forecasts and should have a baseline set of "typical error" metrics in order to monitor an overall performance level. The "typical error" metrics should include nMAE, nRMSE, BIAS. The RP-FSS documents discuss the many factors that should be considered in order to have a representative set of forecasts. The sample size is one of the key factors: 1-year is ideal but it should have a minimum of 3 months.

3) *Quality Control of Evaluation Sample*

The detection of missing or erroneous data and a clear strategy how to deal with such missing data needs to be made at the outset of any evaluation period to ensure that verification and forecasting is fair and transparent.

4) *Use a Set of Error Evaluation Approaches*

A set of evaluation approaches should be employed in order to assess a range of forecast performance attributes. Ideally, the approaches should be customized for the application so that the most important forecast performance attributes are evaluated. However, some of the approaches that should be considered: (1) visual inspection, (2) use of more specific metrics: SDE, SDBIAS, StDev, VAR, CORR, (3) use of histogram or box plot for evaluation of outliers, (4) use of contingency tables for specific event analysis, (5) use of improvement scores relative to a relevant reference forecast

Establishing an evaluation framework or matrix is complex, but can be straightforward if the principles of forecast uncertainty and choice of appropriate metrics are incorporated into the evaluation strategy. The best practice for the establishment is to go through the steps outlined above to select the components for the evaluation framework. The core concept is to use this framework to define a formal structure and then add multiplication factors to weight each of the selected individual metrics according to their relative importance.

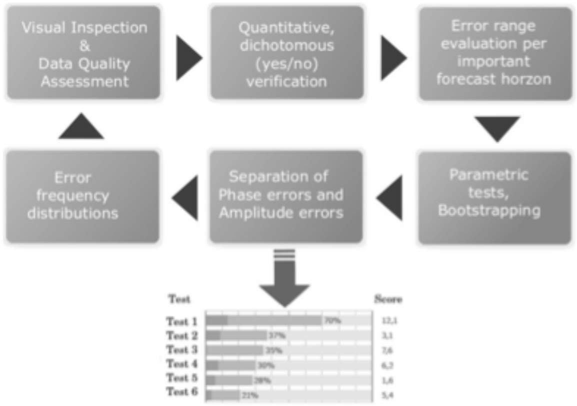


Fig. 3. Example of an evaluation matrix that verifies forecasts against six test metrics and displays the scores for a holistic overview of the forecast performance.

TABLE I. EXMPLES OF APPLICATION-SPECIFIC FACTORS TO BE CONSIDERED IN THE FORMULATION OF A FORECAST EVALUATION FRAMEWORK

Forecast Application	Target	Metric
Forecast to system operator for grid operation	Penalty for squared errors in the time frame 12-24h >±20% of installed capacity	RMSE on time frame 12-24h
	Size of forecast error is important due to reserve restrictions (1) 0-5% MAE/RMSE (2) 6-10% MAE/RMSE (3) 11-15% MAE/RMSE (4) 16-25% MAE/RMSE (5) 25-50% MAE/RMSE (6) 51-100% MAE/RMSE	MAE/RMSE 1 (5% penalty) MAE/RMSE 2 (5% penalty) MAE/RMSE 3 (10% penalty) MAE/RMSE 4 (20% penalty) MAE/RMSE 5 (20% penalty) MAE/RMSE 6 (40% penalty)
	Size of ramping due to reserve restrictions	Dichotomous verification with contingency table, Critical Success Index (CSI) Split up amplitude and phase of: RMSE SDBIAS BIAS StDEV CORRELATION
	High-speed shutdown prediction error due to costs associated with curtailment	Dichotomous verification with contingency table with additional scores, for example: Bias score Probability of detection (POD) Probability of false detection (POFD) Perfect score Success ratio Relative value curve
Forecast to power exchange day-ahead market (bid)	Absolute error of forecast versus production in the look-ahead time frame of 24-48h	MAE
	Distinguishing errors at different look-ahead times due to pricing of balancing power (1) 6-10h (2) 12-14h (3) 17-19h	MAE 1 (45% penalty) MAE 2 (20% penalty) MAE 3 (35% penalty)
Forecast to power exchange intra-day market (bid)	Absolute error for 1-3h forecasts	BIAS, MAE at 1-3h horizon

The matrix can be setup in a spreadsheet environment with macros or within a database environment, where all data is available and metrics may even be directly computed though the database software. The key point of the matrix is that the forecast performance results can be collected, multiplied with an "importance factor", normalized and transferred into a summary table to visualize the scores. For example the scores can be visualized with a bar chart that indicates the performance in a scale from e.g. 0 to 1 or 0 to 100 as shown in Fig. 3.

V. PLANS FOR RP-FSS VERSION 2

As noted earlier, the first version of the three RP-FSS documents was the culmination of a 3-year effort under Phase 1 of IEA Wind Task 36 and that this series of documents has been released as a formal IEA report in September 2019.

The work on the RP-FSS is continuing under the second phase of Task 36. The second phase of work on the RP-FSS is planned to have two major components: (1) a campaign to obtain feedback from forecast users and other power system stakeholders to identify areas in which the RP-FSS can be improved and (2) the preparation of a second version of the RP-FSS that addresses the issues identified in the feedback from stakeholders and also expands and refines the scope of the documents. All parties interested in the RP-FSS are strongly encouraged to become part of this international collaboration to improve the value of wind power forecasting.

The campaign to obtain stakeholder feedback has several components. One of these components is the organization and execution of "feedback" workshops at several opportunistic and geographically diverse venues and via presentations at stakeholder gatherings such as the Wind Integration Workshop. All material from these feedback workshops can be found at the IEA Wind task 36 homepage in the section "Publications → Workshops & Special Session". In addition, a feedback capability that will be established on the IEA Wind Task 36 web portal.

Stakeholders have already provided some valuable feedback on the initial version of the RP-FSS. One key issue that has been raised is that the first version of the RP-FSS is heavily focused on the evaluation and use of deterministic forecast solutions and that very little information is provided about the evaluation and selection of probabilistic forecast solutions, even though probabilistic solution are often better choices for many applications.

Some stakeholders have also identified a need for background information about the sources of uncertainty in wind power forecasts and the relative magnitude of those sources. This would allow forecast users to better understand the role of the different parts of a forecast system and the limitations on forecast performance. A separate publication that addresses this topic is in preparation that is expected to be ready in mid 2021.

A third issue in stakeholder feedback is that it would be valuable to provide recommendations for the use of experienced third parties to design and execute benchmarks or trials. This can be a more effective approach in cases in which the forecast user does not have sufficient knowledge, experience to conduct a satisfactory B/T. However, if this path it chosen, it raises the question of how to identify a qualified third party.

ACKNOWLEDGMENTS

Significant contributions to some or all of the three parts of the three RP-FSS documents were made by: Jeff Lerner, Jakob Messner, Jethro Browell, Craig Collier, Aidan Tuohy, Justin Sharp and Gegor Giebel. In addition, Gregor Giebel, the operating agent for IEA Wind Task 36 provided abundant support and guidance in the preparation of the documents and the navigation of the IEA report publication process.

REFERENCES.

[1] 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.

[2] 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

[3] Giebel, G., J. Cline, H. Frank, W. Shaw, B.-M. Hodge, C. Draxl, P. Pinson,J. Messner, G. Kariniotakis, C. Möhrlen, 2017: IEA Wind Task 36 Forecasting: Proceedings of the 16th Interntiaonal Workshop on Large-Scale Integration of Wind Power into Power Systems (Berlin, Germany). Energynautics GmbH, Darmstadt Germany, ISBN: 978-3-9816549-6-7.

[4] 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.

[5] Giebel, G., W. Shaw, H. Frank, C. Draxl, J. Zack, P. Pinson, C. Möhrlen, G. Kariniotakis and R. Bessa, 2019: IEA Wind Task 36: The New Phase for the Wind Power Forecasting Task. Proceedings of the 18th International Workshop on Large-Scale Integration of Wind Power into Power Systems (Dublin, Ireland). Energynautics GmbH, Darmstadt Germany.

[6] Zack, J., C. Möhrlen, and C. Draxl: 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.

[7] Zhang, J., B.-M. Hodge, A. Florita, S. Lu and H. F. Hamann, V. Banunarayanan, 2013: Metrics for Evaluating the Accuracy of Solar Power Forecasting. Proceedings of the 3rd International Workshop on Integration of Solar Power into Power Systems (London, England). Energynautics, Darmstadt Germany.

[8] Fowler T. Brown B. Lazo J. Haupt S.E. Jensen T.Metrics for evaluation of solarenergy forecasts. Tech. rep. NCAR, 2006.url:http://opensky.ucar.edu/islandora/object/technotes:538.

[9] Messner, JW, Pinson, P, Browell, J, Bjerregård, MB, Schicker, I. Evaluation of wind power forecasts—An up-to-date view. Wind Energy. 2020; 23: 1461– 1481. https://doi.org/10.1002/we.2497

[10] Kariniotakis G. Nielsen HA Nielsen TS. Madsen H. Pinson P. "Standardizingthe Performance Evaluation of Short-Term Wind Power Prediction Models".In:Wind Engineering29.6 (2005), 475 ˘489.doi:10.1260/030952405776234599

[11] Towards the definition of a standardised evaluation protocol for probabilistic windpower forecasts. Tech. rep. Anemos.Plus Project, 2012.url:http : / / www .anemos-plus.eu/images/pubs/deliverables/aplus.deliverable_d1.3-protocol_v1.5.pdf.

[12] Braun A. Koch J. Jost D. Dobschinski R.J. Vogt S. "Benchmark of Spatio-temporal Shortest-Term Wind Power Forecast Models". In: Stockhlm, Swe-den, 2018.

[13] Zhang. "A suite of metrics for assessing the performance of solar powerforecasting". In:Solar Energy111 (2015), 157.