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

PUBLIC REVIEW VERSION

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

Prepared as part of the IEA Wind Task 36, WP 2.1.
Version:3.2 - **PUBLIC REVIEW VERSION**
Date: 07. December 2018

Edited by: Jeff Lerner and Corinna Möhrlen

With Contributions from: Jethro Browell, Jakob Messner, John Zack, Craig Collier, Aidan Tuohy, Justin Sharp and Gregor Giebel

Table of Contents

1 INTRODUCTION TO BENCHMARKS AND TRIALS

1.1 BEFORE YOU START

This is the second part of a series of three recommended practice documents that deal with the development and operation of forecasting solutions. This document "Execution of Benchmaks and Trials" deals with the configuration and steps for carrying out a benchmark or trial of different forecasting solutions against each other prior to selection.

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 third part "Forecast Evaluation", provides information and guidelines regarding effective evaluation of forecasts, forecast solutions and benchmarks and trials.

If your main interest is in selecting a forecasting solution or verifying the quality of your forecast solution, please move on to part 1 or part 3 of this recommended practices guideline, respectively.

1.2 BACKGROUND

The effectiveness of forecasts in reducing the variability management costs of power generation from wind and solar plants is dependent upon both the accuracy of the forecasts and the ability to effectively use the forecast information in the grid management decision-making process or trading purposes. Therefore, there is considerable motivation for stakeholders to try to obtain the most effective forecast information as input to their respective decision tools.

This document is intended to provide guidance to stakeholders on a primary mechanism that has been used extensively in the past years to assess the accuracy of potential forecasting solutions: benchmarks and trials.

Note, forecasting trials and benchmarks will throughout this document be abbreviated with "t/b" for simplicity.

This guideline focuses on the key elements to carry out a successful trial or benchmark and on typical pitfalls. It will also provide recommendations as to when it is beneficial or too risky or expensive in terms of resources to carry out a trial or benchmark.

1.3 OBJECTIVES

The guidelines and best practices recommendations are based on years of industry experience and intended to achieve maximum benefit and efficiency for all parties involved in such benchmark or trial exercises. The entity conducting the trial will have the following benefits:

1. *Being able to evaluate, which of a set of forecast solutions and forecast service providers (FSP) fits best the need, specific situation and operational setup*
2. *Short term internal cost savings, by running an efficient trial*
3. *Long term cost savings of forecast services, by following the trial standards and thereby help reduce the costs for all involved parties*

2 INITIAL CONSIDERATIONS

This section is targeted to the task of engaging a forecast service provider (FSP) and how to navigate through the vast amount of information.

2.1 TACKLING THE TASK OF ENGAGING A FORECASTER

The most important considerations when starting out to plan a forecasting trial or benchmark (t/b) is to be clear about the desired outcome.

The following tables provide information about the benefits and drawbacks of selecting to carry out a t/b. Before starting a t/b it is recommended to go through these tables and determine if the effort is warranted.

2.1.1 Benefits of Trials and Benchmarks

Table 1: Decision support table for situations, where trials/benchmarks are beneficial

Situation	Benefit
Real-time trial for a selected number of sites	Relatively low costs, gaining some experience, getting an idea about the error level, verifying which solution or FSP fits best to current or planned operational setup
Real-time trial for an entire portfolio	Higher costs, experience gain is accordingly higher and more realistic
Retrospective Benchmark with historic data for a specific time period different to the training set	Very cost efficient In multi-vendor systems, the error level is secondary, while the correlation determines whether the new vendor improves the overall error post the mixing procedure

2.1.2 Limitations with Trials and Benchmarks

Table 2: Decision support table for situations, where trials/benchmarks are NOT beneficial

Situation	Limitation	Recommendation
Finding best service provider for large Portfoilo (> 1000MW) distributed over larger area	Test of entire portfolio is expensive for client and service provider in terms of time and resources. Simplifying test limits reliability of result for entire portfolio.	RFI and RFP, where service provider's methods are evaluated and incentive scheme on the contract terms provides more security on performance.
Medium sized Portfolio (500MW< X < 1000MW) over limited area	est of entire portfolio is expensive for client and service provider in terms of time and resources. Simplifying test limits reliability of result for entire portfolio.	RFP, where service provider's methods are evaluated. Built of system that enables change of service provider and incentive scheme more reliable than a test.
Finding best service provider for small sized portfolio (< 500MW)	Test of portfolio requires significant staff resources a budget; usually requires 6 months	Test is possible, but expensive. Difficult to achieve significance on target variable in comparison to required costs and expenses – trial costs makes solution more expensive. Cheaper to setup incentive and a system where the suppliers can be exchanged relatively easily.
Micro portfolio (<	Cost of a trial with	Evaluate methodologies

100MW) or single plants	many parties can easily be higher than the cost of 1 year of forecasting. Time for a trial can delay real-time experience by up to 1 year!	and set up internal system with an incentive scheme, and ease of adding or switching service providers
Power marketing	Best score difficult to define, as sale is often also dependent on market conditions and a statistical score like RMSE or MAE cannot reflect the best marketing strategy	More efficient and timely to perform back test of historical forecasts combined with historical prices, or make a strategic choice with an performance incentive.
Market share of service provider is high	Monopolies in the power market mean that forecast errors are correlated and hence increase balancing costs.	Ask about the market share of a provider and do not choose one with a share > 30% as the only provider!
Blind forecasting, i.e. no measurement data available for the park or portfolio	Without measurements testing is very limited due to the significant improvement from training forecasts. Evaluation can only be carried out for day-ahead or long-term forecasts.	If you have a portfolio > 500MW, a blind test against a running contract can provide an inexpensive way to test the potential of a new provider. For single sites, the benefits of training are so large (>50% of error reduction at times) that blind forecasting is not recommended. It wastes resources for everybody without providing useful results.

2.2 TIME LINES IN A TRIAL OR BENCHMARK

Time lines need to be set strictly in any trial or benchmark in order to fulfill the attributes of a fair, transparent and representative exercise.

The following time lines should be considered:

(1) start and stop of the t/b must be fixed

(2) start and stop must be the same for all FSP

(3) delivery times of forecasts must be set and monitored

(4) missing forecasts from one FSP must be black marked for all

2.3 1-PAGE "CHEAT SHEET" CHECKLIST

The following checklist is provided to help trial organizers save time, apply best practices, and avoid common pitfalls when designing and executing forecast trials. It has been compiled by leading forecast vendors and researchers with many years experience.

Forecast Trial Checklist

--Preparation--

- ☐ Determine outcomes / objectives
- ☐ Consult expert with experience
- ☐ Establish timeline and winning criteria
- ☐ Decide on live or retrospective trial
- ☐ If live trial with datafeed, begin datafeed setup
- ☐ Gather metadata (use IEA checklist spreadsheet)
- ☐ Determine if adequately resourced to carry out
- ☐ Obtain historical data
- ☐ Invite forecast service providers
- ☐ Distribute historical and meta-data
- ☐ Finalize datafeed configuration (if applicable)
- ☐ Allow two weeks Q&A prior to start
- ☐ Begin

--During Trial--

- ☐ Develop validation report
- ☐ Check interim results
- ☐ Provide interim results (if no live data being provided)
- ☐ End

--Post Trial--

- ☐ Provide final results
- ☐ Notify winner(s)
- ☐ Contract with winner(s)
- ☐ Start Service

3 PHASES OF BENCHMARK/TRIAL

There are three main phases of a trial or benchmark exercise: preparation ahead of the trial, actions during the trial, and post-trial follow up.

3.1 PHASE 1: PREPARATION

The time required for the pre-trial preparation is significant and should not be underestimated to insure a successful outcome. If you have no experience in renewable energy forecasting or running a t/b, it would be prudent to contact an experienced individual, organization or forecast provider to get feedback on what can reasonably be accomplished given your time line and objectives. Part 1 of this recommended practice contains a decision support path that may be useful for determining the proper course of action if you have no experience with renewable energy forecasting.

3.1.1 Key Considerations in the Preparation Phase

Once your objectives are known (see Sec Background and Objectives), there are some key decisions to be made that will affect the complexity of the trial.

They are:

(1) **Choice of forecast horizon**

Are forecast horizons less than 6 hours operationally important? If the answer is "no", establishing a live data feed may not be necessary. Although there are advantages of running a trial with a live data feed, it is one of the most time consuming aspects of trial preparation.

Are forecast lead times greater than "day-ahead" operationally important? If the answer is no, this will reduce the volumes of data that need to be processed saving time and resources.

If many lead times are of operational importance, consider that the performance of different providers will likely vary across lead times, therefore, different lead times, e.g. hour-ahead, day-ahead and week-ahead, should be evaluated separately.

(2) **Weather conditions for the exercise:**

Will the benchmark take place during periods of more difficult to predict

weather conditions that reflect the organization's difficulties in handling renewable generation, e.g. windy or cloudy periods? The trial operator wants to make sure the answer here is "Yes" to insure the sample size of harder-to-forecast events is sufficient. If the answer here is "No", trial operator should strongly consider doing a retrospective forecast (also known as "backcast"), if possible.

(3) **Historical data/observations for the exercise:**
Most FSPs require at least 6-12 months of on-site historical observations to train their forecast models. If this much data is not available, trial operator might consider another location or conduct a longer trial on the order of 4-6 months to monitor forecast improvements over time.

(4) **Representativeness:**
Is the benchmark location representative from a wind-climatology perspective of what operator will require contractually? That is, the trial operator should select the location that is needed for subsequent forecasting or a location with a similar climatology. Operators should also be aware of the randomness of forecast performance on single locations, if a large area with many sites is the target.

(5) **Metrics:**
Are the metrics that will be used to evaluate the forecasts meaningful to the bottom line to the success of my project? There are many error metrics to choose from that penalize forecast errors differently. For example, root mean squared error penalizes large errors more than small errors. It is important to choose a metric, or metrics, that reflect the value of an improved forecast and can discriminate between different forecast solutions. Please refer to part 3 of this recommended practice for details on metric selection.

3.1.2 Metadata Gathering in the Preparation Phase

Details of the forecast trial, such as location and capacity of the target generator, are required by all FSPs and comprise the trial Metadata. In Appendix A "Metadata Checklist" provides the typically required information you'll need to prepare for the FSPs participating in the trial and is designed to use as a spreadsheet to fill out. This should also include the desired format (filename and content) of the forecasts you'll be comparing. The clearest way to communicate the forecast file format to multiple FSPs is to provide an example file.

3.1.3 Historical Data Gathering in the Preparation Phase

On-site observations of power production or the renewable resource (e.g., irradiance, wind speed at hub height) are critical for helping the FSPs statistically "train" their forecast models and thus reduce error and bias in the forecasts. Most FSPs require at least 6-12 months of good quality data to train their statistical models. By good quality, we mean that the data does not, for example, contain many gaps. Curtailed power data should be accompanied by plant availability or a curtailment flag.

Data time intervals should be regular and clear documentation of the units, how the observations were averaged, time zone of the data, and whether there's a shift in time due to daylight savings time should be documented. Appendix A of this document has a clear list of the necessary historical data attributes required to efficiently start a benchmark or trial.

3.1.4 IT/Data Considerations in the Preparation Phase

Most organizations are constrained on IT help time, so it's best to plan ahead or keep the sending/receiving of data very simple. If you're doing a live trial (most common), but are not posting realtime data, then you'll need a place for each FSP to send forecast files *to*. One of the metrics that you may be using for deciding on a FSP is the timeliness of the forecasts. In this case, it is important that you can verify the time of delivery.

If you're providing realtime data, you'll most likely need to create a common password protected file server directory that FSPs pick up data from (e.g., Secure Shell File Transfer Protocol (SFTP)). Less common is to push files to each FSP which will require more effort.

Historical data often requires a SCADA engineer or expert on third party software to extract the common data that you'll be providing all FSPs for training their models.

3.1.5 Communication in the Preparation Phase

Anonymizing the FSPs for all external communication is considered a best practice as it promotes competition and entry from smaller FSPs trying to become more established in the industry. Communication via email therefore should always be blind carbon copied to all FSPs and consistent.

Consistent in this context means always sending and sharing emails with the same group of FSP users. Common information sharing engenders trust and the perception of fairness in the benchmark or trial process.

In the preparation phase, it is not uncommon that the FSPs will have questions that could affect how the trial is conducted. For this reason, it is recommended to have a 2-week question and answer period *before* the official start date to allow FSP participants to ask questions that then can be answered in a living document that contains all questions and answers up to the present time. All participants should be notified whenever this document is updated.

The importance of frequent and clear communication cannot be overstated when conducting a trial or benchmark. Not only will the trial operator receive the most accurate forecasts, it will make it much easier the next time a trial is carried out to gage the state-of-the-art in forecasting technologies and features.

3.1.6 One-week test run in the Preparation Phase

It is recommended to allow FSPs a one-week test period before the official start date of the trial or benchmark to identify and remove any technical issues that could invalidate forecast results. This helps to improve the likelihood that all results can be included in the final validation calculations without the need for omitting the first part of the trial.

3.2 PHASE 2: DURING BENCHMARK/TRIAL

Often the most successful forecast provider is one that can show steady improvement over time. Providing an interim validation report will not only prepare the trial operator for the final validation report but will give important feedback to the FSPs.

3.2.1 Communication during the T/B

By design, there should be less reason for back-and-forth communication between the trial operator and FSPs during the trial compared to pre-trial. However, issues do arise especially for a live trial with a real-time data feed. It is recommended that if any decisions are changed during the live part of the trial or benchmark, they should be communicated to all participants immediately as they might require action on the FSP's part. Examples might include: changing the forecast validation metric, if there are unreported outages that should be omitted for future model trainings, or if the location of the data feed or forecast file destination has changed.

Again, all communication should be made with all FSPs equally. Additional communication with individual FSPs (including forecast incumbents) can be construed as showing partiality.

3.2.2 Forecast Validation and Reporting during the T/B

Forecast validation reports are often developed before or during the trial or benchmark. With forecast data coming in at regular intervals, the trial operator has real data to feed into the validation report. If the trial or benchmark is lasting several months (i.e., >3 months), it is recommended to provide at least one interim report to FSPs that include anonymized results from all FSPs. This benefits the trial operator as errors in the report generation can be flagged earlier and ways to make the report generation more efficient can be realized. The interim report benefits the FSPs as course-corrections can be made during the trial to improve the forecasts.

If there are several FSPs participating, efficiencies can be realized by automating part or most of the validation metrics especially as the forecast file format should be the same from all FSPs.

3.3 PHASE 3: POST TRIAL OR BENCHMARK

The post trial phase is important aspect of the trial because FSP selection will likely occur based on the criteria set out from the beginning (see recommended practices part 1 on "evaluation of services and decision support").

3.3.1 Communication at the end of the T/B

If the trial operator hasn't already done so, an email should go out within a week before the end date alerting FSPs to the end of the trial and the time line for sharing results and re-iterating the selection process for the winning FSP.

3.3.2 Forecast Validation and Reporting at the end of the T/B

If an interim report was sent out during the trial, 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, the trial operator may wish to share the anonymized forecast results from each FSP at the time-interval frequency that forecasts were being made at (e.g., hourly). This will help FSPs discover where forecasts are similar or different from the competition which may spawn improved methodologies.

4 BEST PRACTICES

Although there are many different ways that trials and benchmarks may be conducted, there are some common elements of successful trials that provide the trial operator with the best forecast solution and the participants with useful knowledge of where their forecast ranks among the competition. Here are a few:

(a) A clear purpose for the trial/benchmark exercise
(b) Pre-defined, explicit accuracy metrics and winning criteria
(c) A clear time line (start/end dates, winner announcement, contract award)

(d) Anonymized forecast results. Ask FSP's approval to share results. This helps FSPs find ways to improve their forecast accuracy and see their shortcomings.

(e) Question & answer period before benchmark period begins (~ 1-2 weeks)

(f) Sufficient time allocated for testing the transfer of data between participant(s) and operator

(g) Prompt communication to participants regarding any changes or answers to questions that arise

(h) Consistent forecast file format requested of all - example file sent to all

(i) Providing the same historical and project metadata to all participants

(j) Allocation of sufficient resources to furnish data and perform validation

(k) Consistent data formats (both observations and forecast files) ideally as close to (if not identical to) what the trial operator needs, once contract is executed.

5 PITFALLS TO AVOID

Here are a few common mistakes made and how to avoid them in the design, setup and execution of a forecast benchmark or trial.

The consequences of errors and omissions in trials are often under-estimated. However, if results are not representative, the efforts that have gone into a trial are lost. The significance of some common pitfalls can be expensive to the operator as now, they have to make a decision without having objective information to base it on.

1. **Poor Communication**
   All FSPs should receive the same information. Answers to questions should be shared with all FSPs. Fairness, and perception of fairness, are important when running and evaluating the results of trials.

2. **Unreliable Validation Results**

   Don't compare forecasts from two different power plants or from different time periods. Forecast performance will vary depending on location and specific time periods. Only forecasts for the same period and location/power plant/portfolio should be compared.

3. ***Examples of Bad Design***
   (a) Avoid carrying out trials with 1 month length during a low-wind month
   (b) No on-site observations shared with forecast providers.
   (c) Hour ahead forecasts initiated from once a day data update.
   (d) Avoid data only being processed in batches or at the end of the trial in a real-time trial – this is an invitation for cheating to the FSPs. In most cases, there will be some that use the opportunity to do so

4. ***Details missing or not communicated***
   Examples include: daylight savings time changes, whether data time stamp represents interval beginning or ending, plant capacity of historical data differs from present, curtailment and maintenance outages omitted

5. **Remove Possibility of Cheating**

   Forecast trials shouldn't be carried out for a period of time that FSPs are given data for. Also, if there's an incumbent forecaster with a longer history of data, ask for, in writing, that they will not use additional data during the trial that they have exclusive access to. Missing forecasts should be appropriately penalized as missing data may bias "average" forecast metrics, potentially resulting in incorrect conclusions being drawn. If downloading forecasts from a FSP as part of a live trail, files should ideally be downloaded in accordance with the operational process being simulated, and certainly before the time period being forecast. If forecasts are not downloaded or removed shortly after delivery, there is potential for FSPs to cheat by reloading forecasts with newer information, making the results unrealistic and useless. Such an omission should not be underestimated.

6 REFERENCE MATERIAL

6.1 REFERENCES

[1] J. Kehler and D. McCrank, Integration of wind power into Albertas electric system and market operation, Proc. of IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century, Pittsburgh, PA, pp. 1-6. doi: 10.1109/PES.2008.4596824, 2008.

[2] E. Lannoye, A. Tuohy, J. Sharp, V. Von Schamm, W. Callender, L.Aguirre, Solar Power Forecasting Trials and Trial Design: Experience from Texas, Proc. of 5th International Workshop on the Integration of Solar Power into Power Systems,, Brussels, Belgium, ISBN: 978-3-9816549-2-9, 2016.

[3] E. Lannoye, A Tuohy, J Sharp, and W Hobbs, Anonymous Solar Forecasting Trial Outcomes, Lessons learned and trial recommendations, Proc. of 7th International Workshop on the Integration of Solar Power into Power Systems, Paper SIW-126, Berlin, Germany, 2017.

[4] C. Möhrlen, C. Collier , J. Zack , J. Lerner , Can Benchmarks and Trials Help Develop new Operational Tools for Balancing Wind Power?, Proc. of 7th International Workshop on the Integration of Solar Power into Power Systems, Paper SIW-126, Berlin, Germany, 2017. Online access: http://download.weprog.com/WIW2017-292_moehrlen_et-al_v1.pdf.

[5] Bessa, R.J.; 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, 1402. Online access: http://www.mdpi.com/1996-1073/10/9/1402

2.1 CONFERENCE PAPERS

Corinna Möhrlen, Recommended Practices for the Implementation of Wind Power Forecasting Solutions Part 1: Forecast Solution Selection Process, Proc. 17th International Workshop on Large-Scale Integration of Wind Power into Power Systems as well as on Transmission Networks for Offshore Wind Power Plant, Stockholm, Sweden, October 17.-19, 2018

Online Access: http://www.ieawindpowerforecasting.dk/publications

Corinna Möhrlen, John Zack, Jeff Lerner, Aidan Tuohy, Jethro Browell, Jakob W. Messner, Craig Collier, Gregor Giebel
Part 2&3: DESIGNING AND EXECUTING FORECASTING BENCHMARKS AND TRIALS AND EVALUATION OF FORECAST SOLUTIONS, Proc. 17th International Workshop on Large-Scale

Integration of Wind Power into Power Systems as well as on Transmission Networks for Offshore Wind Power Plant, Stockholm, Sweden, October 17.-19, 2018

Online Access: http://www.ieawindpowerforecasting.dk/publications

C. Möhrlen, R. Bessa, Understanding Uncertainty: the difficult move from a deterministic to a probabilistic world, Proc.
17th International Workshop on Large-Scale Integration of Wind Power into Power Systems as well as on Transmission Networks for Offshore Wind Power Plant, Stockholm, Sweden, October 17.-19, 2018

Online Access: http://www.ieawindpowerforecasting.dk/publications

C. Möhrlen, R. Bessa, G. Giebel, J. Jørgensen,G. Giebel, Uncertainty Forecasting Practices for the Next Generation Power System, Proc. 16th Int. Workshop on Large-Scale Integration of Wind Power into Power Systems as well as on Transmission Networks for Offshore Wind Power Plant**, Berlin (DE), 26-29 June 2017.**
Online Access: http://www.ieawindpowerforecasting.dk/publications

C. Möhrlen (WEPROG, Denmark), C. Collier (DNV GL, USA), J. Zack (AWS Truepower, USA), J. Lerner (Vaisala, USA)
Can Benchmarks and Trials Help Develop new Operational Tools for Balancing Wind Power?, Prog. 16th Workshop on Large-Scale Integration of Wind Power into Power Systems as well as on Transmission Networks for Offshore Wind Power Plant **, Berlin (DE), 26-29 June 2017.**
Online Access: http://www.ieawindpowerforecasting.dk/publications

2.2 PRESENTATIONS

C. Möhrlen, C. Collier , J. Zack , J. Lerner , Can Benchmarks and Trials Help Develop new Operational Tools for Balancing Wind Power?, Proc. of 7th International Workshop on the Integration of Solar Power into Power Systems, Paper SIW-126, Berlin, Germany, 2017.
Online access: http://download.weprog.com/WIW17-292_MOEHRLEN-ET-AL_PRESENTATION_20171028.pdf

J. W. Zack (2017)**,** Wind and solar forecasting trials experience: do's and don'ts, Part 2 UVIG 2017 Forecasting Workshop, Atlanta (US), 21-22 June 2017
Online access:
http://www.ieawindforecasting.dk/-/media/Sites/IEA_task_36/Publications/forecast_trials_session_4_uvig2017_jzack.ashx?la=da

C. Collier (2017), Why Do Forecast Trials Often Fail to Answer the Questions for which End-Users Need Answers: A Forecaster's Point of View UVIG Forecasting Workshop, Atlanta (US), 21-22 June 2017. Online access:

http://www.ieawindforecasting.dk/-/media/Sites/IEA_task_36/Publications/forecast_trials_session_4_uvig2017_ccollier.ashx?la=da
T. Maupin (2017), Wind and Solar Forecasting Trials: Do's and Don'ts, Part 1 Best practices. UVIG 2017 Forecasting Workshop, Atlanta (US), 21-22 June 2017. Online access:http://www.ieawindforecasting.dk/-/media/Sites/IEA_task_36/Publications/forecast_trials_session_4_uvig2017_ccollier.ashx?la=da

2.3 GLOSSARY

T/B: *Trial and Benchmark*

FSP: *Forecast Service Provider*

Forecast Creation Time: *The time at which a forecast is created. This is useful when determining skill at different lead times though usually deliver time will be used instead.*

Forecast Delivery Time: *Similar to creation time, only this is the time the forecast was actually received by the end user. This is then used to define what lead time should be ascribed.*

Forecast Lead Time: *The time between the delivery (or creation) time and the beginning of the first interval being forecasted. For example, a forecast delivered at 8:30, where the first entry is for 5-minute period ending 9:05 has a 30 minute lead time.*

Forecast Horizon Time: *The time of the last forecast interval relative to the delivery time. For instance, a day head forecast with hourly intervals from midnight to midnight the following day has a horizon time of midnight on date+2*

Forecast Interval: *The length of time between the forecast start time and the forecast end time.*

Forecast Valid Time: *The time interval for which a forecast is valid. The last valid time is the same forecast horizon.*

Appendix A: Metadata Checklist

The following checklist (Table A.1), when filled out, will greatly aid FSPs in configuring forecasts efficiently. Many of the essential questions relevant to benchmark and trial forecast model configuration are provided here.

Note that the following table is an example and may not contain all necessary information required for the FSP to setup a solution for your purpose. The table is meant to serve as a guideline and can be copied, but should be carefully adopted to the specific exercises before sending out to FSP with questions filled in. If this is done with care, it will expedite forecast configuration and save back and forth communication time.

Table A.1: Example of a Metadata Checklist

Wind Power Forecast Trial Checklist	
Metadata	
Name of site(s) as it should appear in datafile	
Latitude and longitude coordinates of sites	
Nameplate capacity of each site	
Will a graphical web tool be needed?	
Turbine make/model/rating	
Number of turbines	
Hub height of turbines	
Please attach suitable plant power curve	
Forecast output information	
Forecast output time intervals (e.g., 15-min, 1-hourly)	
Length of forecast required	
Timezone of forecast datafile	
Will local daylight savings time be needed?	
Forecast update frequency (e.g., once a day, every hour)	

Value of Forecast	
Which variables will be forecasted and validated?	
Which forecast horizons are being validated?	
Which metrics are being used to gage forecast performance?	
List criteria for determining winning forecast provider	
Will results be shared as a report? Will results be anonymized?	
On what frequency will results be shared with forecast provider?	
Historical Data Checklist	
Is the data in UTC or local time?	
Is the data interval *beginning* or *ending* or *instantaneous?*	
What are the units of the data?	
If met tower histories being provided, indicate height of measurements.	
Realtime Data Checklist (if applicable)	
Is the data in UTC or local time?	
Is the data interval *beginning* or *ending* or *instantaneous?*	
What are the units of the data?	
Email and Telephone number of technical point of contact (POC)	
Email and Telephone of datafeed POC	
Name and email of users that need website access	
Person name and email that filled out this checklist:	

Appendix B: Sample downloadable forecast file

Back and forth communication can sometimes delay the start of a trial or benchmark. One of these delays is getting the forecast file output format just right for the beginning of the trial. Standardization of the format will make the trial operators life much easier when time comes to validating forecasts. A best practice here is for the trial operator to use a format that is already in use or a format that has already proven to work in operations.

Table B.1 below shows the first few fields of a forecast file template.

Plant Output	Acme Wind Farm	1.11.2017 4:00	1.11.2017 5:00	1.11.2017 6:00	1.11.2017 7:00
Power	MW	41.43	41.43	41.43	40.89
Windspeed	m/s	11	10	10	10
Time zone: Central European Summer Time (CEST)					
Intervals: hour ending					
Date time format: dd.mm.yyyy hh:mm (e.g., 06.08.1969 08:30)					

Table B.2 shows typical XSDs for forecasts and SCADA data in a trial with WebService

```
<?xml version="1.0" encoding="utf-8"?>
<xs:schema attributeFormDefault="unqualified" elementFormDefault="qualified"
xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="WindForecast">
    <xs:complexType>
      <xs:attribute name="VendorCode" type="xs:string" use="required" />
      <xs:attribute name="ImportTime" type="xs:dateTime" use="required" />
      <xs:sequence>
        <xs:element name="CUSTOMER">
          <xs:complexType>
            <xs:attribute name="name" type="xs:string" use="required" />
            <xs:sequence>
              <xs:element name="Forecast">
                <xs:complexType>
                  <xs:attribute name="MWaggregated" type="xs:double" use="required" />
                  <xs:attribute name="time" type="xs:dateTime" use="required" />
                  <xs:sequence>
                    <xs:element name="Probability">
                      <xs:complexType>
                        <xs:attribute name="P95" type="xs:double" use="required" />
                        <xs:attribute name="P50" type="xs:double" use="required" />
                        <xs:attribute name="P05" type="xs:double" use="required" />
                        <xs:attribute name="max" type="xs:double" use="required" />
                        <xs:attribute name="min" type="xs:double" use="required" />
                      </xs:complexType>
                    </xs:element>
                    <xs:element name="WindFarms">
                      <xs:complexType>
                        <xs:sequence>
                          <xs:element name="WindPark1">
                            <xs:complexType>
                              <xs:attribute name="id" type="xs:string" use="required" />
                              <xs:attribute name="mw" type="xs:double" use="required" />
                            </xs:complexType>
                          </xs:element>
                        </xs:sequence>
                      </xs:complexType>
                    </xs:element>
                  </xs:sequence>
                </xs:complexType>
              </xs:element>
            </xs:sequence>
          </xs:complexType>
        </xs:element>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
</xs:schema>
```

- *SCADA XSD for exchange of real-time measurements*

```
<?xml version="1.0" encoding="utf-8"?>
<xs:schema attributeFormDefault="unqualified" elementFormDefault="qualified"
xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="WindSCADA">
    <xs:complexType>
      <xs:sequence>
        <xs:element maxOccurs="unbounded" name="WindPark">
          <xs:complexType>
            <xs:attribute name="ID" type="xs:string" use="required" />
            <xs:attribute name="Time" type="xs:dateTime" use="required" />
            <xs:attribute name="Mw" type="xs:decimal" use="required" />
            <xs:attribute name="Availabilty" type="xs:decimal" use=" optional" />
            <xs:attribute name="CurrentActivePower" type="xs:decimal" use=" optional"/>
            <xs:attribute name="Curtailment" type="xs:string" use="optional" />
            <xs:attribute name="WindSpeed" type="xs:decimal" use="optional" />
            <xs:attribute name="WindDirection" type="xs:decimal" use="optional" />
            <xs:attribute name="AirTemperature" type="xs:decimal" use="optional" />
            <xs:attribute name="AirPressure" type="xs:decimal" use="optional" />
            <xs:attribute name="Outage" type="xs:decimal" use="optional" />
          </xs:complexType>
        </xs:element>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
</xs:schema>
```