Insight on Human Decision-making from Probabilistic Forecast Games and Experience: an IEA Wind Task 36 initiative

Corinna Möhrlen
WEPROG
Assens, Denmark
Email: com@weprog.com

Nadine Fleischhut
Max Planck Institute for Human Development
Berlin, Germany
Email:fleischhut@mpib-berlin.mpg.de

Ricardo J. Bessa
INESTEC
Porto, Portugal
Email:ricardo.j.bessa@inesctec.pt

***Abstract*—As the penetration levels of RES increase and climatic changes lead to increasingly more extreme weather conditions, the uncertainty of the weather forecasts and power production forecasts can no longer be ignored for the grid operation. In order to support the power industry in the adaptation of uncertainty forecasts into their business practices, the IEA Wind Task 36 has started an initiative in collaboration with the Max Planck Institute for Human Development to investigate the existing barriers in the industry to the adoption of such forecasts into decision processes. In the first part of the initiative, a forecast game was designed as a demonstration of a typical decision making task in the power industry. The game was introduced in an IEA Wind Task 36 workshop and thereafter released to the public. When closed for a final evaluation, it had been played by 105 participants. We will discuss our first experience with the experiment and introduce some of the new features of the second and third generation of experiments we are currently working on as a continuation of the initiative. We will also discuss specific questions that emerged when we started the experiments. With these future experiments, we want to help the industry to learn about the benefits of using uncertainty forecasts by providing training tools and realistic simulation decision scenarios.**

I. Introduction

The challenges of climate change require unprecedented investments in renewable energy sources (RES) and a fundamental transformation of existing infrastructures such as the power grid. A key challenge is that RES like wind and solar power are characterised by an intermittent and inherent variability that is unknown to fossil fuels and other weather-independent energy sources. To deal with the inherent variability of RES and seamlessly integrate them into the power grid requires a combination of weather forecasts and weather-to-power conversion methods. However, current approaches are based on deterministic weather forecasts, which provide no means to explicitly represent the varying uncertainty of forecasts. As the penetration levels of RES increase and extreme weather conditions rise with climatic change, current methods have reached their limit because they neither model nor convey forecast uncertainties. Instead, new methods for power production forecasts are called for that explicitly model uncertainty in order to make robust predictions and allow to communicate forecast uncertainty to market participants and system operators. Integrating uncertainty into the decision process is important, because due to the high flexibility, RES can be used at much lower costs than traditional energy sources- although only one-sided for down-regulation and with limited capacity for balancing power at the grid level.

II. Explaining the Need for the Initiative

The need to take into account and explicitly model forecast uncertainty is at the heart of many scientific and applied enterprises. For instance, the ever-increasing accuracy of weather forecasts has been driven by the development of *ensemble forecasts*, where a large number of forecasts are generated either by generating forecasts from different models or by repeatedly perturbing the initial conditions of a single forecast model. Importantly, this approach provides robust estimates of forecast uncertainty, which supports human judgment and decision-making (e.g. [**?**]). Although weather forecasts and their uncertainty are also crucial for the weather-to-power conversion for RES forecasting, the industry has been reluctant to adopt ensemble methods and other new technologies that can help managing highly variable and uncertain power feed-ins under different weather conditions. For instance, energy traders can use uncertainty forecasting to optimize the amount of power generation that they bid into the market, and transmission system operators can define the required reserve to account for the uncertainty in generation or be prepared for grid congestion much further in advance than with the currently used deterministic forecast. As a direct consequence of adopting these state-of-the art methods, the renewable energy sector forgoes critical potentials to reduce its vulnerabilities, build more robust prediction models, and improve judgment and decision making of the involved parties.

There are two intertwined challenges responsible for the limited adoption of uncertainty forecasts in the renewable energy sector. First, current energy management systems simply cannot operate with probabilistic forecasts.

The second challenge is the human factor: New methods and approaches are bound to fail if transmission system operators and other relevant agents do not know how to harness new forecast methods and systematically integrate uncertainty in their decision-making processes.

Our aim is to address these challenges in a unified and trans-disciplinary approach, bringing together hitherto separate fields (meteorology, behavioral insights, cognitive science, energy sector) and competencies. The overarching goal is to demonstrate the value of using ensemble forecasts in the RES sector.

III. Objective of the Exploratory Phase

We want to answer the most pressing questions that prevent the development by using behavioural decision experiments to simulate real-time problems for specific user groups using ensemble data and the corresponding power production forecasts.

The results and feedback from these experiments can then be used to formulate strategies for further research and how to overcome the barriers that prevent the use of such forecast information and the harvest of the associated benefits from the additional information.

A. First Experiment – background and setup

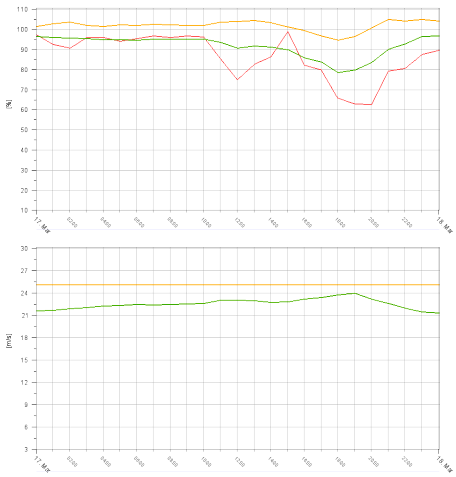
The goal of the first experiment was to simulate a decision making task in the power market, where there is high uncertainty of the production forecasts and where wrong decisions directly could be related to costs. Uncertainty in weather and also power production forecasts is not a new topic (e.g. [2], [3], [4], [5]) and has been discussed in many workshops of industry groups such as the Energy Systems Ingetration Group (ESIG) in the USA or the German Forecasting Platform (IFP) run by the Fraunhofer Institute of Energy economy and energy systems technology (IEE) and the German Weather Service (DWD). Today, such information is readily available and there seem to be obvious advantages of using such information. However, it is also known that a number of factors influence human responses and decisions under uncertainty that go beyond a "rational" calculation of costs and probabilities. In order to investigate the factors that may underlie the hesitant reaction to probabilistic forecast observed in industry, we designed the first experiment in such a way that the participants had a direct comparison of deterministic and probabilistic information as input to their decision making task.

For the first experiment, we chose a trading situation that provided different types of information for the decision making:

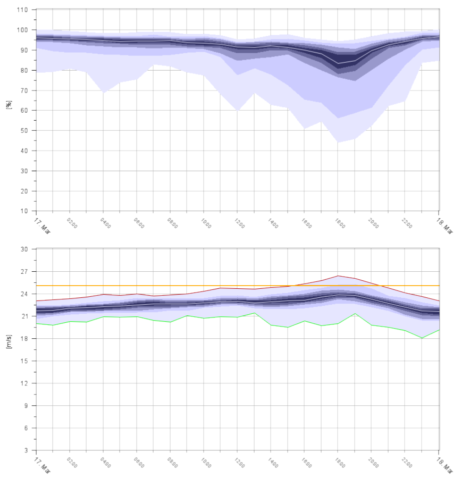
1) deterministic forecasts of wind speed and wind power
2) probabilistic ensemble forecasts of wind speed and wind power

In the experiment we assumed that the most traders would agree that over time the success of trading wind and solar power in a power market situation is related to the costs of the balancing power that is required to level out forecast errors. The income is relatively strongly correlated with these balancing costs and that few events with high forecast error would be the driver for the costs and reduction of the income. We assumed that approximately 5% of the time, where the forecasts are off track account for 95% of the costs over a longer period of time, e.g. some months or a year. The assumption here is that it is more beneficial to reduce the balancing costs for large errors than seeking for a general forecast improvement of 1-2%.

Fig. 1 shows one situation that was presented to the participants in the first example, where the deterministic forecast does not indicate or at least not clearly indicates the risk for a high-speed shut down, whereas the probabilistic forecasts upper percentiles P80-P90 and the maximum forecast in wind speed clearly indicate the risk, confirmed by the power forecast that shows the effect on the power generation with the percentiles P10-P40.



((a)) Deterministic forecasts



((b)) Probabilistic forecasts

Fig. 1. The deterministic forecasts (a) and the probabilistic forecast (b) as example of one of the chosen situations. For each type of forecast, the upper figure shows the wind power forecasts and the lower figure shows the wind speed forecast. The orange line in the wind speed forecasts marks the threshold (25 m/s) around which a high-speed shutdown can occur.

As shown in the example plot, we simplified the task by letting the users decide whether the generating power of an offshore wind park should be traded fully or partially given the possibility of a high-speed shutdown, where the wind park stops generating due to excessive wind conditions. Such so-called high-speed shutdown events occur in wind ranges between 21-27m/s and are often referred to as cut-off wind threshold at 25 m/s. Because the wind turbines are not only calibrated to react on average wind speeds over a certain period of time, but also short-lasting wind gusts there is considerable uncertainty on a wind farm to enter into high-speed shutdown (**HSSD**). All situations were real world forecasts and reflect a decision that traders or operators are faced with in their daily work.

B. *Second Experiment - taking a different perspective*

In the first experiment, we used an experimental design where participants were shown the deterministic forecast first before they could make a second decison based on the probabilistic forecast for the second situations. One disadvantage of this updating design was that it does not allow to quantify whether or not participants will make better decisions by using probabilistic uncertainty forecasts alone. Instead, the participants essentially decided whether to hold on to the previously made decision or change their mind. From a psychological perspective this is a critical aspect to consider when drawing conclusions from an experiment. For instance, often people do not update their beliefs sufficiently to change their minds. Thus, people may decided differently if they would have seen the probabilistic forecast independent of the deterministic forecast. For the second experiment we will instead use a design suited to evaluate the benefit of probabilistic forecasts independently.

The experiment will be setup in two different ways:

1)
2) Every participant makes all decisions based on deterministic as well as on probabilistic forecasts
3) For each forecast type, the situations are presented in blocks randomized among each other
4) For every participant, the order of blocks is randomly chosen at the beginning.

In both runs, the forecast situations will be randomized so that participants are not getting the same sequence of forecasts to prevent that they remember the situations. The goal is to be able to identify the benefits, if any, by using the probabilistic forecasts and deterministic forecasts on their own and explore problems in the use of the probabilistic forecasts in more detail.

IV. Summary of the Lessons Learned from the First Experiment

The first experiment showed us a number of interesting aspects both from the participants reaction and the results of the experiment:

- A significant amount of the participants would like to have both probabilistic wind speeds and wind power generation forecasts or a deterministic "best guess" inside the uncertainty bands
- The improvement in decisions with additional probabilistic information was not equally strong in all cases, which points to the importance to understand the indicators in the forecast and the strategies that people use to make their decisions
- A situation with low visible uncertainty caused participants to expect no HSSD and to take a false risky decision, which raises the questions how threshold values are intuitively or rationally evaluated for uncertainty bands

V. Lessons learned and Next Steps

The first experiment in the series has revealed a number of interesting aspects regarding decision making with and without uncertainty information from probabilistic ensemble forecasts. Although we can conclude that there is a potential benefit, and most participants also confirmed in the follow-up questionnaire that they clearly preferred to have probabilistic forecasts at hand when making decisions on extreme events, there are still many open questions. Some of these will be answered in the second experiment, for instance, about the impact of either probabilistic or deterministic forecasts on their own.

Additionally, this first experiment in the planned series has been focusing solely on decision making with the help of graphical time series forecasts for both, the deterministic and the probabilistic forecast types. However, we have not been investigating, as described e.g. in [6], whether the type of representation of the uncertainty also could have an impact on decisions, i.e. whether some or all end-users may be more likely to make better decisions, if they would get the probabilistic information in a text format. Or, if for example a graph or a series of graphs of the large-scale weather situation in a horizontal plot as shown in Fig. 2 could be useful for the decision making process. The example shows the wind speed (colors), wind direction (arrows) and isobars, joining together places of equal atmospheric pressure, for the time of highest probability of a shut-down from the example in Fig. 1 as mean (large left graph) of 75 forecasts and the minimum (upper right) and maximum (lower right) thereof, respectively.

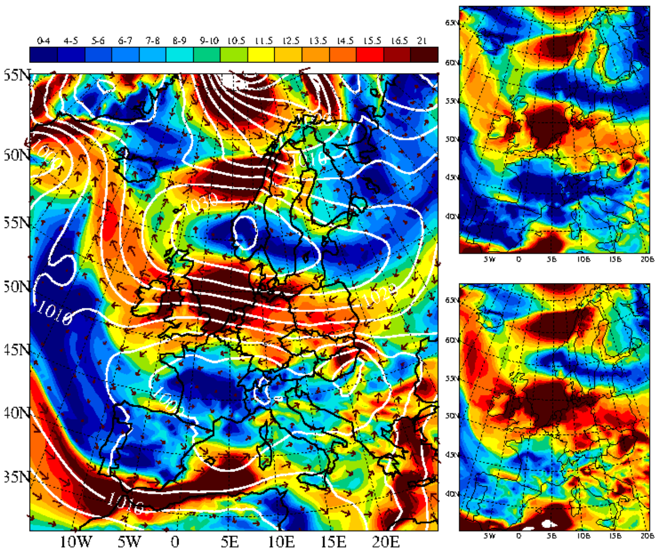


Fig. 2. Weather maps of wind speed (color scale), wind direction (arrows) and isobars. The large graph is the EPS Mean of 75 members, the upper right graph is the minimum and the lower right graph the maximum wind speed of the 75 members in each grid point.

In our future work, we will consider the following research directions (focused on low probability, but high impact scenarios):

- Scenarios with wind speeds above cut-off value for multiple hours, that uncertainty forecasts generated with a statistical model are not able to capture
- Scenarios with extreme balancing power prices (this sometimes occurs in the Nord Pool or EPEX market), where forecast errors in one direction can be highly penalized

- Presenting percentiles versus all ensembles members to the decision-maker to investigate how decisions may change when only a subset of ensembles members are presented.
- Investigate whether decision-makers are more risk averse or prone given probabilistic forecasts and whether the amount of uncertainty makes a difference

Examples for case 1 are scenarios, where the (1) event was never observed in historical data or (2) the temporal dependency structure is not modelled by the statistical approach used. An experiment would allow to investigate how well different modeling approaches are understood and applied correctly when introduced to the industry. Since the science behind the probabilistic approaches are rather complicated, the IEA Wind Task 36 has started to establish documentation about different approaches broken down to a level that adressses the typical educated power engineers. However, it is unclear whether this type of scenario and its hidden uncertainty play a role in the currently observed barriers of implementing probabilistic applications.

An example for case 3 could be to use the probabilistic game from utility theory to study the choice between a feed-in tariff y (guaranteed income) or direct participation in the electricity market with probability p of earning less than y or $1 - p$ for winning more than y.

ACKNOWLEDGMENT

The authors would like to thank all the participants of the experiment and the IEA Wind Task 36 group for valuable input in the preparation of the experiment. Corinna M ohrlenwas supported by the Danish EUDP project IEA Wind Task36 Forecasting Danish Consortium under the contract no.64018-0515. Nadine Fleischhut was supported by the Hans-Ertel-Centre for Weather Research of the German National Weather Service.

REFERENCES

[1] Joslyn S, LeClerc J. Decisions With Uncertainty: The Glass Half Full. Current Directions in Psychological Science. 2013;22(4):308-315. doi:10.1177/0963721413481473

[2] V. J. Fundel, N. Fleischhut, S. M. Herzog, M. G ober, and R. Hagedorn,"Promoting the use of probabilistic weather forecasts through adialogue between scientists, developers and end-users,"QuarterlyJournal of the Royal Meteorological Society, vol. 145, no. S1, pp.210–231, 2019.

[3] A. H. Murphy and Q. Ye, "Optimal decision making and the valueof information in a time-dependent version of the cost-loss ratiosituation,"Monthly Weather Review, vol. 118, no. 4, pp. 939–949,1990.

[4] R. J. Bessa, C. M ohrlen, V. Fundel, M. Siefert, J. Browell, S. H. E.Gaidi, B.-M. Hodge, U. Cali, and G. Kariniotakis, "Towards improvedunderstanding of the applicability of uncertainty forecasts in theelectric power industry,"Energies, vol. 10, no. 9, p. 1402, Sep 2017

[5] C. M ohrlen, M. Pahlow, and J. U. Jørgensen, "Untersuchungverschiedener handelsstrategien f ur wind- und solarenergie unterber ucksichtigung der eeg 2012 novellierung,"Zeitschrift f ur Energiewirtschaft, vol. 36, no. 1, pp. 9–25, March 2012

[6] LeClerc, J. and Joslyn, S. (2015), The Cry Wolf Effect and Weather-Related Decision Making. Risk Analysis, 35: 385-395. doi:10.1111/risa.12336.