

RESEARCH ARTICLE

A decision-making experiment under wind power forecast uncertainty

Corinna Möhrle¹  | Ricardo J. Bessa²  | Nadine Fleischhut³ 

¹Department of Research and Development, WEPROG ApS, Assens, Denmark

²INESC Technology and Science (INESC TEC), Porto, Portugal

³Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin, Germany

Correspondence

Nadine Fleischhut, Center for Adaptive Rationality, Max Planck Institute for Human Development, 14195 Berlin, Germany.

Email: nfleischhut@gmail.com

Present address

Nadine Fleischhut, Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin, Germany

Funding information

Danish EUDP project IEA Wind Task 36 Forecasting Danish Consortium, Grant/Award Number: 64018-0515; Hans-Ertel-Centre for Weather Research of the German National Weather Service, Offenbach, Germany, Grant/Award Number: 4818DWDP3A; Portuguese National Fund FCT - Fundação para a Ciência e a Tecnologia, Grant/Award Number: UIDB/50014/2020

Abstract

As the penetration levels of renewable energy sources increase and climatic changes produce more and more extreme weather conditions, the uncertainty of weather and power production forecasts can no longer be ignored for grid operation and electricity market bidding. In order to support the energy industry in the integration of uncertainty forecasts into their business practices, this work describes an experiment conducted with 105 participants from the energy industry. In the framework of an IEA Wind Task 36 workshop, the experiment aimed to investigate existing psychological barriers in the industry to adopt probabilistic forecasts and to better understand human decision processes. We designed and ran a ‘decision game’ to demonstrate the potential benefits of uncertainty forecasts in a realistic—although simplified—problem, where an energy trader had to decide whether to trade 100% or 50% of the energy of an offshore wind park on a given day based on deterministic and probabilistic uncertainty day-ahead forecasts. The focus thus was on a decision-making process dealing with extremes that can cause high costs in the form of security issues in the electric grid for system operators, or high monetary losses for traders, who have bid a power production into the market that failed to be produced due to high-speed shutdown of the wind turbines. This paper presents the obtained results, extracts behavioural conclusions and identifies how to overcome psychological barriers to the adoption of uncertainty forecasts in the energy industry.

KEYWORDS

ensemble forecasting, experiment, forecast value, *human decision-making*, uncertainty, wind power

1 | INTRODUCTION

The key strategy to fight climate change worldwide is to invest in renewable energy sources (RES) and increase their integration into power systems. In recent years,

however, we observed how extreme weather conditions, together with growing penetration levels of RES, are increasingly affecting the power system operation and planning, as well as electricity markets (Chandramowli et al., 2014). The inherent uncertainty of such events and

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Meteorological Applications* published by John Wiley & Sons Ltd on behalf of Royal Meteorological Society.

the associated uncertainty in the power generation from RES can no longer be ignored by the energy industry.

While the development of RES forecast models based on so-called deterministic weather forecasts has already started in the 1990s (Giebel et al., 2011; Hong et al., 2020; Mengelkamp, 1988) and is today a mature technology, the integration of RES into power systems requires the design and application of new methods for power forecasts that explicitly model uncertainty in order to achieve robust predictions, communicate forecast uncertainty to stakeholders and policymakers, and integrate forecast uncertainties into the decisions process. However, as the penetration levels of RES increase and climatic changes produce more and more extreme weather conditions, current deterministic methods have reached their limit due to the inherent inability to model and convey forecast uncertainties. Examples of these extreme events are wind speeds above the cut-out of wind turbines (e.g., 25 m/s) (Lin et al., 2012) that are the focus of this work; extreme high energy shortfall events at low wind speeds can however also become more frequent in the future (van der Wiel et al., 2019).

These developments require the design and application of new methods for RES power forecasts that explicitly model uncertainty to achieve robust predictions, communicate forecast uncertainty to stakeholders and policymakers, and integrate forecast uncertainties into the decision-aid systems. Probabilistic information and forecasts have been shown to improve decision-making in weather-related processes (Joslyn & LeClerc, 2013; LeClerc & Joslyn, 2015). Moreover, probabilistic forecasts allow to better evaluate whether the decision strategy should be adjusted in the face of negative outcomes.

In meteorology, the uncertainty of weather development is simulated by so-called ensemble prediction systems (EPS), where a sufficient large number of forecasts are generated by either combining many different numerical weather prediction (NWP) models or by perturbing the initial conditions of a single NWP model (e.g., Houtekamer et al., 1996; Molteni et al., 1996; Toth & Kalnay, 1997). The methodology was implemented operationally at the large met centres in the late 1990s, at about the same time as research on power production models began (Mengelkamp, 1988). Ensemble forecasts, when calibrated on short-term horizons (Möhrlen, 2004; Pinson, 2012), can also be used to quantify the uncertainty of the expected wind or solar power generation. They enable market traders, market operators, transmission system operators or parties responsible for power balancing to act far in advance on the uncertain part of the expected generation (Möhrlen et al., 2012; Pahlow et al., 2009). For instance, some traders use this kind of forecast to optimally adjust the amount of power

generation that is bid into the market (Grimit, 2017). Or transmission system operators use it to define a reserve required to account for the uncertainty of the power generation and to prepare for grid congestion much further in advance than based on deterministic forecasts (Haupt et al., 2019). Power system and grid resilience to adverse weather events is also becoming an important case for the application of probabilistic weather forecasts (Moreno et al., 2020). Statistical learning methods can also be used to generate renewable energy uncertainty forecasts and statistically-based ensembles (Hong et al., 2020). Although there is a fair number of applications today that make use of ensemble forecasts, large parts of the industry still have difficulties adopting these types of forecasts into their operation (Bessa et al., 2017; Haupt et al., 2019; Würth et al., 2019). As a result, the renewable energy sector forgoes a huge potential to reduce its vulnerabilities.

1.1 | Simulated experience in games to overcome psychological barriers

Reluctance to use new methods, especially if they seem to be too complicated to grasp, is a well-known phenomenon—not only in the energy industry. Not surprisingly, reluctance to use probabilities in decision-making has been reported before in numerous sectors, from medical decision-making to emergency management (e.g., Fundel et al., 2019; Kalke et al., 2021; Mackintosh & Armstrong, 2020; Renooij & Witteman, 1999).

Often, such reluctance may simply result from misunderstandings due to a lack of transparency or poor explanations and representations of the method(s) used to generate the probabilistic forecast. For instance, studies have shown that presenting verbal probabilities (e.g., “likely”) in combination with numerical probabilities leads to more consistent interpretations across persons and reduces reluctance to use probabilities (Budescu et al., 2009, 2014; Wintle et al., 2019). In fact, over the years, research dealing with weather forecasts, flood forecasting, disaster warnings or climate change has accumulated evidence that uncertainty can be understood and improve decisions, if it is transparently communicated (e.g., Joslyn & LeClerc, 2013).

Reluctance to use probabilities may also result from missing experience how to translate probabilistic forecasts into a binary decision (Fundel et al., 2019). Often, users are presented with a description of a probabilistic forecast—without a chance to experience their use in context. Basic research on “decisions from experience” (e.g., Hertwig & Erev, 2009; Hertwig & Wulff, 2021), however, shows that people may adjust decisions over trials, if given feedback about the outcomes of their

decisions—and can develop a less biased interpretation of the uncertainty than when making decisions based on a description alone. This suggests that “simulated experience” with such decisions could help users to overcome their reluctance to use probabilistic wind power forecasts, develop their own decisions thresholds but also a better understanding of the conveyed forecast uncertainty.

Experimental games are a second research strand that incorporates the use of feedback and experience, by allowing people a hands-on experience with a decision scenario and the outcomes of their decisions. In the weather domain, for instance, Roulston et al. (Roulston et al., 2006) ran experiments set up as a game in which the participants had to manage a road maintenance company responsible for salting the roads, with probabilistic forecasts improving their decisions (see also Joslyn & LeClerc, 2012, 2013; LeClerc & Joslyn, 2015).

‘Gamification’ provides an opportunity for ‘hands-on’ decision experience and training in a clearly defined context and a relaxed atmosphere, without the responsibility and potentially serious consequences in real application (Ramos et al., 2018). For instance, the international HEPEX¹ initiative on hydrological ensemble prediction has been fostering the development and application of so far six publicly available role-play games, described in Ramos et al. (2018). They describe ‘...these games as simplified representations of reality, that do not intend to reproduce the full context of operational environments. Nevertheless, they have been successfully used as support material during teaching and training activities’. In the HEPEX context, it was also found (Ramos et al., 2018) that ‘games are an excellent way to introduce complex forecasting concepts and create a relaxed atmosphere for discussion during trainings or workshops’.

Gamification approaches have so far been developed for flood forecasting or water management, for example, HEPEX (HEP, 2021; Ramos et al., 2018; SMH, 2021), firefighter management (WEXICOM, 2019 ‘Feuerwache Game’), rainfall and temperature probabilistic forecasts by the MetOffice Weather Game (Stephens et al., 2019), which used a hypothetical ‘ice-cream seller’ scenario to ‘...test the decision-making ability of the participants using different methods of representing uncertainty and to enable participants to experience being “lucky” or “unlucky,” when the most likely forecast scenario did not occur’.

The main goal of the present paper is to empirically investigate how to overcome the psychological barrier ‘reluctance’ to the adoption of probabilistic forecasts in the wind energy area, by enabling stakeholders to explore their benefit and use.

For this purpose, we used a simplified energy trading scenario to allow participants to experience hands-on decisions based on probabilistic forecasts in an experimental game—and to allow us to explore the effect of probabilistic compared with deterministic forecasts on individuals’ decisions.

For the decisions, we presented a series of deterministic forecasts, each followed by their probabilistic counterparts, to human decision-makers, who are promising candidates to benefit from probabilistic forecasts, and ensemble forecasts in particular. The experimental design focused on a simplified, but realistic decision problem that is important for a diverse group of decision-makers—and that at the same time allows to gain insights into the benefit and use of probabilistic forecast by analysing how people update their decision based on deterministic forecasts after receiving probabilistic forecasts of the same situation.

To the best of our knowledge, this is the first work to investigate wind power human decision-making with uncertain forecasts, in order to: (a) investigate how people respond to and decide based on additional probabilistic forecasts compared with the deterministic forecasts that are currently in use, and (b) enable participants to explore how to use probabilistic forecasts, stimulate discussions and collect ideas on the use of such forecasts.

The current study thus extends complementary previous work on the communication and use of forecast uncertainty outside the wind energy industry, such as: (i) the EPS training course created by the Meteorological Service of Canada (CMC, 2016) intended to introduce participants to ensemble forecasting and promote a paradigm shift to probabilistic approaches in decision-making (EPS); (ii) studies about (partly biased) perceptions of the uncertainty of deterministic weather forecasts and factors that influence the perceived uncertainty (e.g., time horizon) of the public and professional user groups, such as emergency managers (e.g., Bessa et al., 2017; Fleischhut et al., 2020; Joslyn & Savelli, 2010; Kox et al., 2015); (iii) research that evaluates the impact of different representations for communicating forecast uncertainty and decision-making (e.g., Grounds & Joslyn, 2018; Joslyn et al., 2013; Joslyn & LeClerc, 2012; LeClerc & Joslyn, 2015; Marimo et al., 2015; Ramos et al., 2013; Stephens et al., 2019) (iv) statistical analysis of cases in which uncertainty forecasts have been used with a certain degree of success (e.g., aboriginal whaling quotas, weather forecasting, HIV/AIDS epidemic, population projections, number of funded graduate students to admit) (e.g., Raftery, 2016); (v) approaches and challenges to communicate ensemble weather forecasts to different professional user groups, including transmission system operators, emergency managers (e.g., fire brigade control) and road workers

¹Hydrologic ensemble prediction experiment.

(e.g., Demeritt et al., 2010; Frick & Hegg, 2011; Fundel et al., 2019).

The remaining of this paper is organized as follows: Section 2 describes the experiment, which used 12 randomized situations per participant; Section 3 presents the analyses and results for the main research questions; Section 4 discusses the main limitations and avenues for future research; concluding remarks are presented in Section 5.

All participant data, the code for reproducing the analyses as well as the forecasts used in the experiment can be found at a public repository at OSF (All participant data, the code for reproducing the analyses as well as the forecasts used in the experiment can be found in a public repository at the Open Science Framework at https://osf.io/t8q29/?view_only=5cc6e406ab9243229da88d983ead8cb1).

2 | DECISION-MAKING EXPERIMENT

2.1 | A high-stakes decision scenario under wind power uncertainty

In the decision scenario, participants assumed the position of an energy trader for an offshore wind park. The scenario reflects a realistic, but simplified daily decision in the electricity market that is strongly affected by uncertain weather development. There were two main reasons for this choice: First, most users of weather forecasts in the energy industry are familiar with ‘uncertainty in weather forecasts’ and there is some understanding about the benefits or potential improvements for decision-making tasks (e.g., Bessa et al., 2017; Fleischhut et al., 2020; Fundel et al., 2019; Joslyn & Savelli, 2010; Mylne, 2002). Our experiment thus provides a familiar starting point for decision-makers. Second, wind power trading is highly sensitive to forecast errors and uncertainty. Success strongly depends on trading no more power than will be available. While trading too little power is inefficient, trading too much often results in high costs from buying ‘balancing’ power at a high market price. This asymmetry in cost and income is reflected in the simplified ‘classical single-stage’ cost-loss function, where the loss from trading too much is higher than trading less (Mazzi & Pinson, 2017; Murphy, 1985). In fact, power traders claim that approximately 5% of the situations are responsible for 95% of the costs in a month or a year, namely those situations with large forecast errors. Reducing the costs from large forecast errors is hence more important than improving the general forecast by 1%–2%. Probabilistic forecasts explicitly quantify how likely and

how large forecast errors might be, and thus promise to be a key tool to improve power trading decisions.

The focus of this experiment thus is on a decision-making process dealing with extremes that can cause the above described ‘additional’ high costs in the form of security issues in the electric grid for system operators, or high monetary losses for traders who bid a power production into the market that failed to be produced due to a high-speed shutdown. As uncertainty in forecasts is unavoidable, decisions need to consider it within a cost-loss evaluation.

It may be a part of human nature to thrive for the ability to produce or receive a perfect forecast. Thus, maybe it is not surprising that people tend to trust in deterministic forecasts and rely on them today to a large extent. However, a deterministic forecast is nothing else than a ‘best guess’ of one possible outcome. Trusting in a deterministic forecast therefore simply means to ignore that uncertainty exists (World-Meteorological-Organisation, 2012). The ensemble forecasts in this experiment were chosen to provide exactly the type of uncertainty forecast that corresponds to the spread of which the deterministic forecasts are just best guesses. In this way, participants had the chance to learn how the ensemble forecasts relate to one or a few deterministic forecasts. For the deterministic forecasts, we chose three wind power forecasts (with different wind speed forecasts as inputs) and one wind speed forecast (a mean least-square error optimized forecasts). The reason why we chose only one wind speed forecast was to reflect the information typically requested and used by most end-users—and thus to demonstrate the lack of information about the underlying uncertainty contained in a single deterministic forecast, when translated into wind power. Figures 1 and 2 show examples of these forecast types.

2.2 | Participants and experimental design

One hundred and five participants, mostly experts in the energy industry or from the meteorological community with relation to the energy industry, decided whether to trade 100% or 50% of the energy of an offshore wind park on a given day, based on (past) real-world deterministic and probabilistic day-ahead forecasts from an operational ensemble prediction system (see details in Section 2.4) in 12 randomized situations (see Figures 1 and 2). For each situation, everyone first made their decisions based on deterministic forecasts before they could update their decisions based on a probabilistic forecast for the same situation and before they received feedback. At the end of the experiment, participants were asked which forecasts they would prefer for the decisions they made.

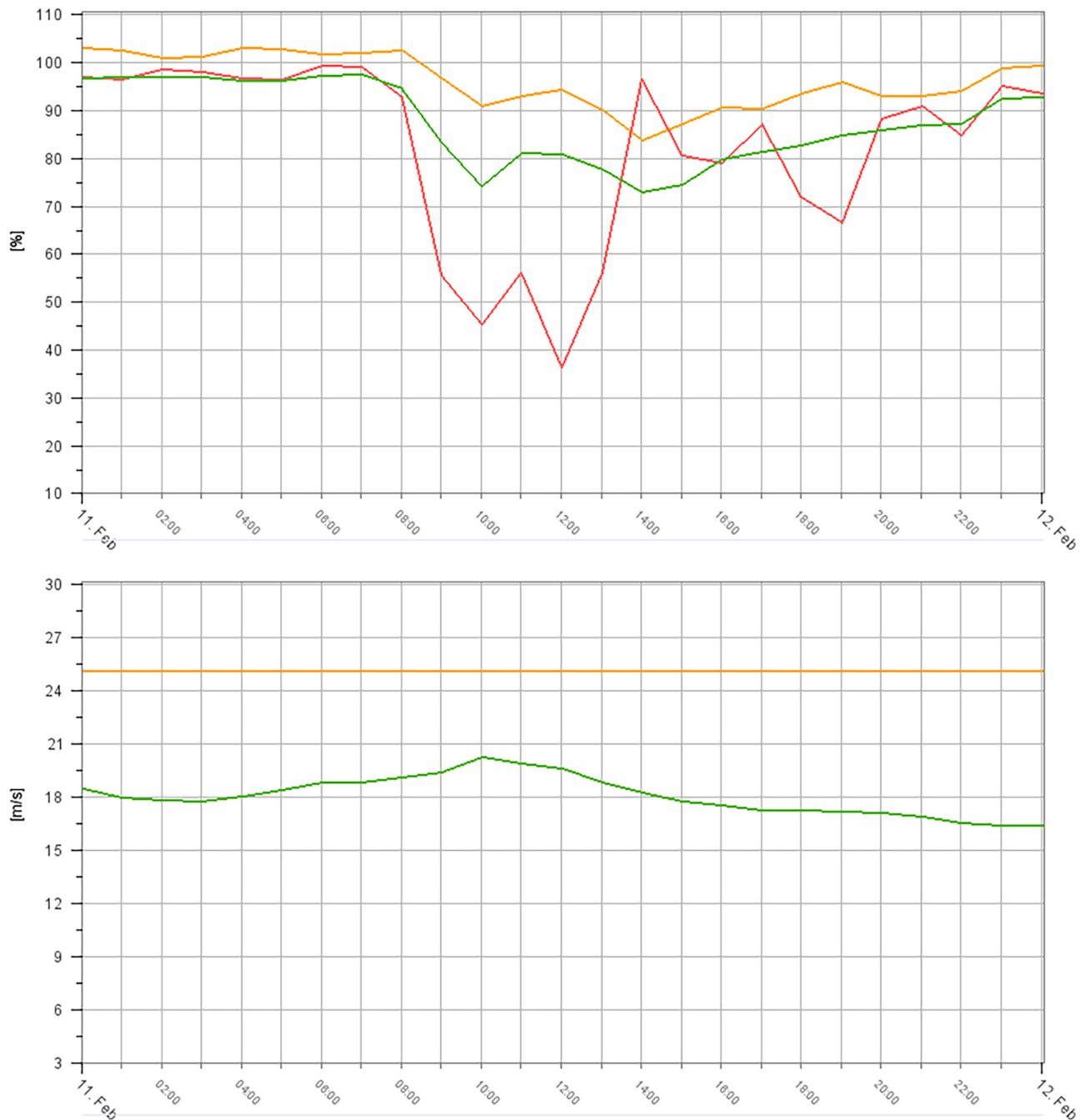


FIGURE 1 Deterministic forecast of Situation 1 as an example: Three independent and different wind power forecasts, selected from NWP models of the 75 member multi-scheme ensemble prediction system (MSEPS) (upper plot), one wind speed forecast randomly chosen from one of the 75 MESPS members (bottom plot); the yellow line is the threshold (25 *m/s*) around which high-speed shutdowns are likely to occur.

This design allowed participants to reconsider the decision they made based on the deterministic forecasts after receiving an additional probabilistic forecast. The main advantage of this design is that it allowed us to quantify how often and in which situation each participant changed their mind. It can thus help to understand in which situations probabilistic forecasts may provide additional information perceived as useful by professional users. A second advantage was that participants went from more familiar

deterministic forecasts to their probabilistic representation. The direct sequential comparison should help them to understand the forecasts, to see the information that the deterministic forecasts hide, and stimulate discussion for the workshop afterwards.

Participants were recruited before a workshop of the IEA Wind Task 36 in January, 2020, where the experiment and type of probabilistic forecasts available for decision-making was explained in more detail. The study was

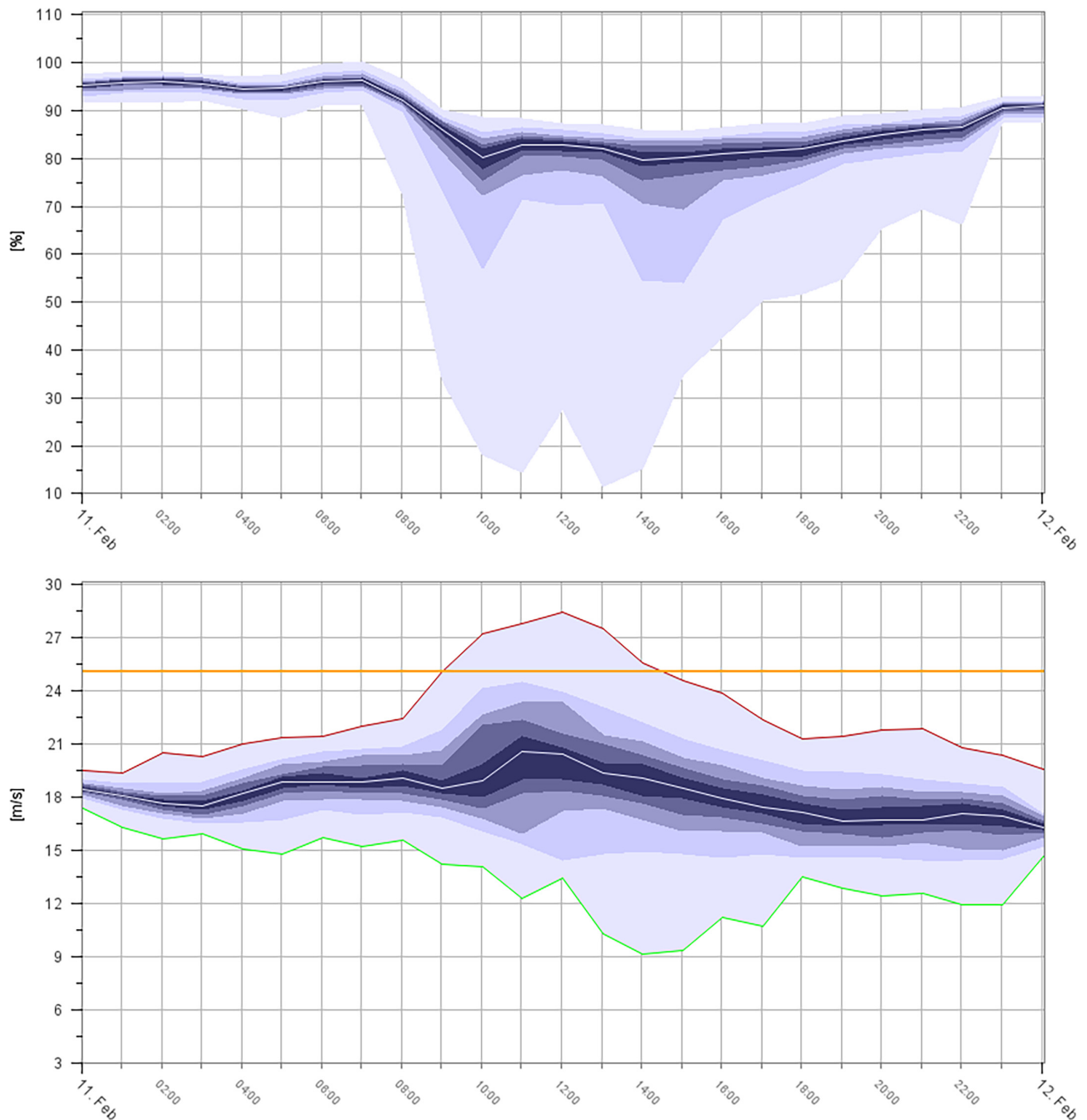


FIGURE 2 Probabilistic forecast of Situation 1 as an example: Wind power (upper plot), wind speed forecast (bottom plot), both in the form of 10 percentiles computed from the 75 member multi-scheme ensemble prediction system (MSEPS); the yellow line is the threshold (25 m/s) around which high-speed shutdown is likely to occur, the red line is the maximum and the green line the minimum of the 75 ensemble forecasts.

approved by the ethics board of the Max Planck Institute of Human Development, and all participants gave informed consent at the beginning.

2.3 | Experimental task and procedure

Based on a given forecast, participants had to decide whether to trade 100% of the generating power of an

offshore wind park or to trade only 50%, given the possibility of a high-speed shutdown (HSSD), where the wind park stops generating due to excessive wind conditions. Participants were informed that high-speed shutdowns typically occur in the wind range above 21–27 m/s, mostly known as the cut-off wind threshold of 25 m/s; and that wind turbines use both the mean wind and wind gusts to determine whether they turn into a high-speed shutdown.

TABLE 1 Cost–loss table for the ‘Offshore Wind Power Trading’ experiment.

	HSSD	No HSSD
Trading 100%	–5000	5000
Trading 50%	0	2500

The instructions also explained the potential payoffs of each decision, depending on whether a high-speed shutdown would occur. To reflect the costs of large and small errors, we used a simplified cost function: Trading 100% generated an income of 5000 EUR, if there was no high-speed shutdown. In case of a high-speed shutdown, a cost of 5000 EUR occurred. Trading 50% generated an income of 2500 EUR, if there was no high-speed shutdown. In case of a high-speed shutdown, income and costs balanced each other, so the payoff was zero (see Table 1). This cost function follows the classical single-stage cost–loss model from the weather forecast literature (Murphy, 1985) and has been used in laboratory experiments to study different ways to communicate forecast information (Bolton & Katok, 2017). Real-world electricity market bidding problems have of course more complex mathematical formulations, such as the ones discussed in (Botterud et al., 2012; Dai & Qiao, 2015). Despite the simplified problem description, the chosen cost function nevertheless reflects a typical asymmetry in price for forecast errors of different sign.

For scoring, we only considered whether there was a high-speed shutdown (HSSD) any time during the forecast period, where the actual generation was either full load production (100%) or a shutdown scenario ($\geq 50\%$ reduced production) for this particular wind farm. No other costs were considered. The wind farm has a capacity of 100 MW and the spot market price is 50 Eur/MW. The cost function assumed that balance costs are equivalent to spot market prices and participants were informed about these details.

Thus, trading 100% was a risky choice that could lead to a large income or a large loss. The more conservative 50% trading strategy eliminates the risk, because no balance cost arise. Given this cost–loss ratio, the expected value of both options, trading 100% and trading 50%, can be calculated using

$$\mathbb{E}[X] = \sum_{i=1}^k x_i p_i \quad (1)$$

where $\mathbb{E}[X]$ is the expected value of a decision option (e.g., trade 100% or trade 50%), x values are the potential outcomes depending on the events (HSSD or no HSSD)

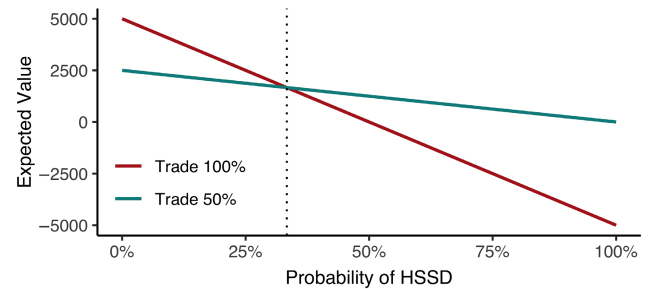


FIGURE 3 Expected payoff of decisions as a function of different probabilities of an HSSD, given the cost–loss table shown in Table 1. The decision threshold is 30%, that is, if the probability of HSSD is higher than 30%, trading 50% has higher expected payoff, and if the probability is lower than 30%, trading 100% has a higher expected payoff.

and p is the probability of each event. If the probability of an HSSD is lower than 30%, the expected value of trading 100% is higher than trading 50%; and it is lower if the probability of an HSSD is higher than 30% (see Figure 3). According to an expected value decision strategy, it would therefore be optimal to trade 100% only if the probability of an HSSD is lower than 1/3, following the red line in Figure 3, and otherwise choose the safe option to trade only 50%, following the green line.

Whereas it is impossible to infer the probability based on the deterministic forecasts, the percentiles in the probabilistic forecasts give an indication of the uncertainty of the forecasts and the probability of exceeding the threshold of 25 m/s in the form of the 9 percentiles (P10...P90). The percentiles provide an indication of when to follow the green line in Figure 3 in the range above 30% probability (P70 wind speed > wind threshold and P30 wind power < full power load) or the red line in Figure 3 below 30% probability (P70 wind speed \leq wind threshold and P30 wind power = full power load). Still, it is worth noting that it is far from trivial to follow an expected value strategy based on the probabilistic forecasts. On the one hand, the representations of uncertainty as a fan chart do not directly convey the probability of exceeding a threshold but require to implicitly infer it from the quantiles as described. On the other hand, empirical studies in psychology and behavioural economics have shown repeatedly that people most of the time do not implement cost–loss or expected value strategies but instead often follow simpler heuristics (Gigerenzer et al., 2011; Gigerenzer & Brighton, 2009; Kahneman, Slovic, & Tversky, 1982).

This shows that for the design of real-world applications, careful thought needs to be given to how best to present forecast data in order to support a specific decision-making task of the user. For example, the probability of exceeding a specific threshold may be further

enhanced by turning ‘red’ when that probability exceeds the level at which the relevant cost function would give a negative expected outcome.

2.4 | Forecast used in the experiment

Participants made decisions based on 12 real-world deterministic and probabilistic day-ahead power production and wind speed forecasts, respectively. The forecasts were selected over a period from October 2018 to April 2019 from an offshore wind park in the North Sea, whose power generation is traded directly in the electricity market. For all forecasts, it was known whether a high-speed shutdown had occurred or not.

Both the deterministic and probabilistic forecasts were taken from an operational multi-scheme ensemble prediction system (MSEPS) (e.g., Möhrlen et al., 2012) that is suitable and well tested for such extreme events in wind energy.² The multi-scheme approach is especially useful for extremes that may happen at any time throughout the forecast horizon (e.g., Bessa et al., 2017; Cali, 2010; Haupt et al., 2019). The MSEPS system, run by the weather service provider WEPROG, is dedicated to the variations of the fast physical and dynamic processes in the lower atmosphere and used for wind energy applications since over 15 years (Möhrlen, 2004).

For the presentation of the probabilistic forecasts, the 75 ensemble members of the MSEPS system were summed up into 9 percentiles P10–P90, represented by a fan chart, with P50 being the median line. As counterpart to the ensemble forecasts, we presented participants with three deterministic forecasts, a typical amount of information provided in real-world applications, for example, from different modelling centres. The three power forecasts were chosen from the MSEPS as independent forecasts with a very different set-up of the NWP models (e.g., see Figures 1 and 2). Although three deterministic forecasts may at times provide a certain hint of the forecast uncertainty, such a small number of forecasts cannot convey a representative uncertainty.

Because wind turbines are calibrated to take both the average wind speed and the short-term gusts into account, the uncertainty in wind speed is also one important indicator to judge the probability of a shutdown of

the wind turbines. Therefore, the power forecasts were always presented together with a wind speed forecast. In the deterministic case, one wind speed forecast, randomly chosen from one ensemble model, was presented. In the probabilistic case, the wind speed forecasts showed 9 percentile bands to indicate the uncertainty underlying any single/deterministic wind speed forecast (see footnote 1).

A horizontal line at 25 m/s indicated the typical average wind speed threshold at which high-speed shutdown may occur (Figure 1), while the variable effect of gusts adds further uncertainty to any decision based on the wind forecast, and can occur at mean speeds from about 21 m/s. Reflecting a real-world issue with extreme events such as an HSSD, in some cases the selected wind speed and power forecasts can also appear inconsistent.

Overall, we selected the 12 forecast situations to systematically cover four categories depending on whether a high-speed shutdown (HSSD) occurred and whether we hypothesized that there were visible indicators of an HSSD in the respective forecasts (see Table 2). Situations in which an HSSD occurred (category one and two) are most important for the prevention of high losses, and in particular Category 1, because it is in these situations where a probabilistic forecast may actually provide a better indication of an HSSD than the deterministic forecasts.

3 | RESULTS

Overall, the experiment provided us with several insights about the usefulness of probabilistic forecasts in a power trading application. First, we found that 70% of the participants had an equal or better final outcome with the additional information from the probabilistic forecasts and that participants made more correct decisions and took less risk when this was appropriate, that is, when the uncertainty forecasts of both wind power and wind speed indicated no or very little probability of a high-speed event to occur. Overall, participants changed their mind after seeing the probabilistic forecasts in 18% of all decisions, and 90% of participants changed their mind at least once.

The results indicate that probabilistic forecasts have the potential to improve human decision-making, although the increased income based on probabilistic forecasts was not significant given the current selection of decision situations. However, the current selection provided important first insights that indicators people may use to correctly predict a high-speed shutdown based on probabilistic forecasts. That participants have made less risky decisions is two-fold: first, it shows that not having the probabilistic information may lead to risky decisions due to a lack of information, which is generally unwanted. Second, it demonstrates that

²In this so-called multi-scheme approach, all ensemble members contain the same NWP kernel, whereas various parameterisation schemes are exchanged. In the MSEPS, the parameterisations dealing with the fast physical processes in the lower atmosphere are solved with five different condensation schemes, five different vertical diffusion schemes and three different dynamic schemes. This provides $5 \times 5 \times 3 = 75$ ensemble members that are well-defined, their difference is known and does not change.

TABLE 2 Categories of the forecasts in the 12 situations presented in the experiment.

Category	HSSD event	Visible in		Situation IDs used in the results
		Probabilistic forecast	Deterministic forecast	
1	Yes	Yes	No	1, 2, 3, 4
2	Yes	Yes	Yes	5, 6
3	No	Yes	No	7, 8, 9, 10
4	No	Yes	Yes	11, 12

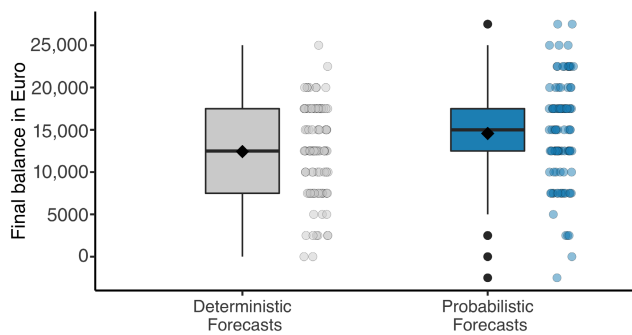


FIGURE 4 Boxplot and raw data of participants' final balance based on deterministic compared with probabilistic forecasts. The horizontal line in the boxplot represents the median, the diamond shows the mean. The distance between the lower and upper limit of the box shows the inter-quartile range (IQR) of the distribution; each dot represents the final balance of a participant. The dots were jittered for better visibility.

the additional probabilistic information can generate more risk-averse behaviour when this is appropriate.

In that sense, the experiment revealed a number of interesting aspects for the decision-making of extreme events and shows that probabilistic forecasts can add value to the decision-making process. In the following sections, we will present the outcome of the various parts of the experiment in more detail.

3.1 | How did probabilistic forecasts affect decisions?

One way to answer this question is to compare participants' final outcome based on the deterministic forecasts alone to their final outcome with the additional probabilistic forecasts. For each participant, we thus aggregated the payoffs over all decisions, separately for the decisions based on deterministic and based on the probabilistic forecasts. The distribution of final outcomes for both conditions is displayed in Figure 4. Based on the probabilistic forecasts, participants had a higher final (mean = 14,571, SD = 5969, median = 15,000, interquartile range [IQR] = 12,500–17,500) than based on the deterministic forecast (mean = 12,429, SD = 5229, median = 12,500,

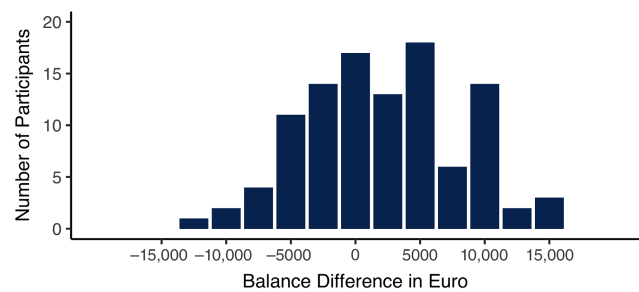


FIGURE 5 Histogram of the difference between participants' final balance based on probabilistic and deterministic forecasts. Values above zero mean that probabilistic forecasts benefited the decisions of a participant, values below zero mean that it harmed decisions.

IQR = 7500–17,500; $V = 2779$, $p < 0.005$, Wilcoxon signed-ranks test on paired sample).

If we look at the central 50% of the participants (IQR), it is worth noting that the spread in outcomes is only half the spread compared with the spread based on the deterministic forecasts. Thus, probabilistic forecast not only improved the overall performance, but also resulted in a more uniform performance of the middle 50% of participants. However, there are also participants in the lower tail of the distribution who performed considerably worse than others.

To understand whether some participants had worse outcomes with the probabilistic forecasts in particular, we calculated how many participants benefited from the probabilistic forecast. We subtracted for each participant their final outcome based on deterministic forecasts from their outcome based on probabilistic forecasts (see Figure 5). Overall, the results are encouraging: 70% of the participants either had better or identical outcomes based on probabilistic forecasts than on deterministic forecasts. Exactly 53% in fact improved their outcomes; however, 30% also did worse than with the deterministic forecasts.

3.2 | Where did the benefits come from?

A better final balance can only be achieved by making more correct decisions, that is, by more often predicting

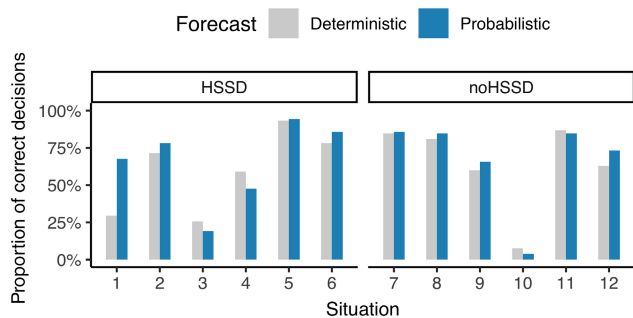


FIGURE 6 Proportion of participants taking the correct decision for each of the 12 decisions situations based on deterministic or additional probabilistic forecasts.

correctly when an HSSD occurs and when it does not. Aggregated across the 12 decisions and 105 participants, the proportion of correct decisions based on probabilistic forecast (66% of $105 \cdot 12 = 1260$ decisions) was slightly higher than based on deterministic forecasts alone (62% of 1260 decisions). If we look at the individual decision situations, the probabilistic forecasts led to more correct decisions in 8 out of 12 situations (Figure 6). Among these, it is Situation 1 that is responsible for the overall better decision outcomes based on probabilistic information in the experiment (for the corresponding forecast see Figure 2).

3.3 | Risk assessment

An important research question is to understand how knowledge of the uncertainty of forecast changes the risk assessment of the situation. Aggregated over all situations and participants, the risky option, that is, trading 100%, was chosen slightly less often based on the additional probabilistic forecast (51% out of 1260) compared with the decisions based on the deterministic forecasts alone (52% out of 1260).

Figure 7 compares the proportion of risky decisions for each of the 12 decision situations. Reflecting the previous results, decisions are quite similar, except for Situation 1. In this case, only 32% choose the risky option based on probabilistic forecasts, compared with 70% based on the deterministic forecasts.

Thus, Situation 1 is a prototypical example of a situation, where income was generated from a decision based on the additional probabilistic forecasts, whereas decisions based on the deterministic forecast alone generated a cost due to an incorrect risky decision. This is in fact what we should expect: in cases with high uncertainty, probabilistic forecasts are more likely to hold decision-makers back and trade less in order to reduce the possible loss. If a deterministic forecast is wrong (e.g., due to

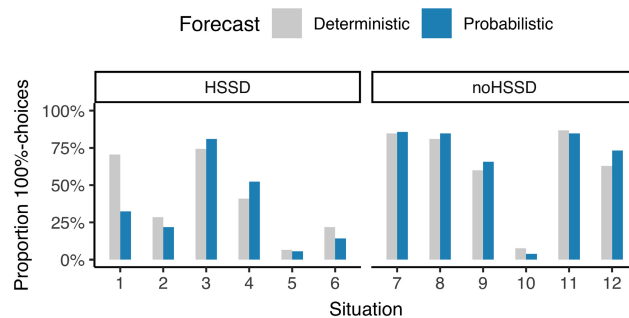


FIGURE 7 Proportion of participants taking the risky option (trading 100%) for each of the 12 decision situations based on deterministic or probabilistic forecasts.

phase error or amplitude error), the decision-maker has no possibility to know why and may thus take a risky decision (trade 100%). The probabilistic forecast, in contrast, warns about a potential risk and may suggest to be careful. It is situations like these that traders fear as the most economic dangerous decision situations, and it is these situations that are often responsible for the large individual losses that generate the bulk of the overall costs. The experiment can thus be considered successful in producing a realistic example application for this type of problem.

For all results, it is important to keep in mind that in the current experimental design, participants always received the probabilistic forecast right after their decision based on the deterministic forecast. When presented with the uncertainty of the forecast, participants therefore had already thought about the situation. Their risk assessment was thus never based on the probabilistic forecast alone.

Instead, the experimental design allows to observe, whether participants change their mind once they are presented with the uncertainty of the forecasts. Overall, participants changed their mind after seeing the probabilistic forecasts in 18% of all decisions, and 90% of participants changed their mind at least once.

Figure 8 shows that in 9 out of 12 situations, more than 10% of the participants changed their mind when presented with the additional information. In three cases, 30%–23% changed their mind (Situations 4, 3 and 9) and in one case—again Situation 1—48% did. The overall low proportion of changed decisions on the one hand reflects the small differences between decisions.

In a real application, it would now be interesting to examine, whether and when the additional information contained in probabilistic forecasts overall improves the risk assessment capabilities of decision-makers, that is, reduces costs and enhances income, and whether it reduces the costs of very expensive cases sustainably. This evaluation would require a higher number and a more

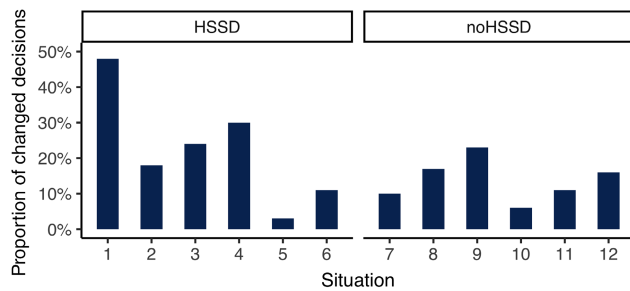


FIGURE 8 Distribution of changed decisions after being presented with the probabilistic forecasts for each of the 12 cases.

representative selection of critical cases as we used in this first experiment. However, the experiment provides a good example of how real applications could be evaluated for a customer of such a product, promising avenues for future research, and a good starting point to explore how to select critical cases for future studies.

3.4 | How did the different situations affect decision-making?

The strong improvement in Situation 1 compared with other situations reflects that probabilistic forecasts did not add relevant information in all situations where we had expected this (category one and three; Table 2).

Here, the current set of situations can help to develop the first hypotheses, which indicators in the forecasts participants used to make their decision, and thus how to select critical cases for future research. For illustration, we focus on situations in Category 1, which are particularly important: in these situations, we expected the probabilistic forecast to be a better indicator for a HSSD than the deterministic forecasts, and hence to reduce the impact of missed events with high losses.

For Situation 1 (Figure 1), the deterministic forecasts barely indicate the risk of a high-speed shutdown. Two out of the three power forecasts indicate no drop in power generation. This is further confirmed by the wind speed forecast that is clearly below the critical threshold of 25 m/s and does not change throughout the forecast period. In the probabilistic wind power forecast (Figure 2), in contrast, it becomes obvious that more than one member of the ensemble predicts a drop in power generation—in line with the wind speed forecast that reveals a considerable uncertainty during the same period.

In Situation 2, 71% and 78% of the participants correctly expected an HSSD and thus traded 50% based on either deterministic and probabilistic forecasts, respectively—compared with Situation 3, where only 26% and 19% correctly expected an HSSD (see Figure 6). If we look at the deterministic forecasts, the wind speed gets much closer to

the critical threshold in Situation 2 than in Situation 3. Moreover, in Situation 2, the deterministic power forecast shows a drop in generation at the same time when the wind speed forecast gets close to the threshold, whereas in Situation 3, the wind speed prediction is constantly away from the threshold. If we look at the probabilistic forecast, the picture does hardly change despite the additional percentiles shown. For instance, in Situation 3, none of the percentiles comes close to the wind speed threshold.

Especially interesting is Situation 4, as slightly more participants correctly expected an HSSD based on the deterministic forecasts (59%) than based on the probabilistic forecasts (48%). Here, the deterministic wind speed forecast is again close to the critical threshold; at the same time, the deterministic power forecasts show no simultaneous performance drop. The conflicting information may have made this a harder decision, and may explain that only 59% of the participants correctly expected an HSSD based on the deterministic forecasts (compared with about 71% in Situation 2). Importantly, the probabilistic forecasts look quite similar. The predicted uncertainty is rather small, with the worst-case ensemble member barely scratching the wind speed threshold. The low visible uncertainty may thus have reassured some more participants to expect no HSSD and falsely take a risky decision (Figure 7). It is important to note that uncertainty forecasts with a larger spread can often be helpful in catching low-probability-high-impact events, but at the same time lead to expensive decisions due to high uncertainty. Narrow forecast intervals can on the other hand lead a decision-maker to over-confidence in a decision—see, for example, Roussos et al. (2021) for a detailed discussion and modelling framework.

In summary, the pattern suggests that a decision-maker needs to understand the cost–loss function of the problem, in this case knowledge about the 30% probability as described in Section 2.3 and shown in Figure 3, or at least use some kind of heuristics such as a simple sequential decision tree: Consider whether the wind speed gets close to the threshold of 25 m/s or not. If it does not get close, do not expect an HSSD. If it gets close and there is a simultaneous drop in power generation, expect an HSSD—otherwise expect an HSSD with some probability. Such a decision strategy would produce similar decisions based on the deterministic and probabilistic forecasts whenever the predicted uncertainty is low.

3.5 | Which forecast information did participants prefer?

In the last part of the experiment, participants indicated which forecasts they would prefer for the decisions they

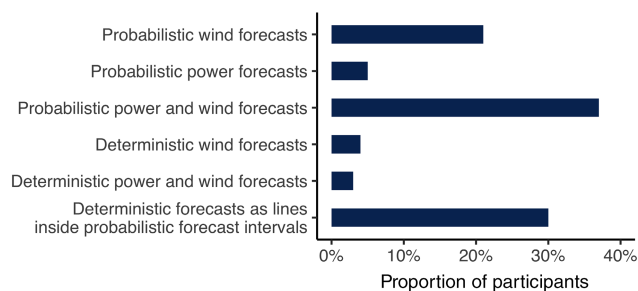


FIGURE 9 Preferred forecast information for the kind of decision made.

made. The questions can be seen in Figure 9. The participants were able to select multiple options. The aim was to investigate whether the participants were satisfied with the provided information and whether they would prefer more information, and if so, what kind of information they would like to receive.

The most striking result was that 93% preferred some type of probabilistic forecast. Among these participants, 37% preferred having probabilistic wind speed and wind power forecasts. Whereas 21% would prefer probabilistic wind forecasts alone, only 5% would base their decision on a power forecast alone. 30% of the participants would like to have a deterministic ‘best guess’ inside the probabilistic forecasts. Such a ‘best guess’ forecast may be trained or tuned with information about the wind farm that are not weather related and can provide additional information relevant to the decision, for example, turbine control mechanisms, when blades are pitched to reduce power output, or related to wake effects in and around the wind farm, direction dependency, etc.

Overall, the results show that decision-makers prefer probabilistic forecasts, but also indicate that the combination of probabilistic wind and power forecasts is still perceived as more helpful than the power forecasts alone. Considering the few users of ensemble forecasts in the industry, the result may also be interpreted as a lack of available information on the concrete implementation possibilities of ensemble forecasts into decision support tools in particular and probabilistic tools in general. Once decision-makers get inspired by the use of such tools through practical applications in the form of games and experiments, some of the identified barriers may disappear.

4 | DISCUSSION

4.1 | Experiment’s limitations

Whereas deterministic forecasts show no information about the probability of a high-speed shutdown,

probabilistic forecasts explicitly quantify the uncertainty information. In this experiment, we showed a typical 24-h window of a trading day in order to help the decision-maker to get an overview and become familiar by working visually with such uncertainty forecasts.

In a real-life application, the decision would have to be made for each hour separately, or, alternatively with time spans and thresholds for such hours, where the probability exceeds these thresholds. Those shutdown periods with lower or no production at all will also be traded or handled over a number of hours before and after the event due to known phase errors in forecasts. In that sense, the current experiment was not too far away from a real-life application.

For the current experiment, it is also important to keep in mind that individuals often do not update their beliefs sufficiently in the face of new information. In this case, seeing the deterministic forecast first is likely to anchor participants so that they likely do not decide based on the probabilistic forecasts as they would without seeing the deterministic forecast first. Moreover, their thoughts about the situation will also be influenced by considering the deterministic forecasts first. The next step will thus be to use an experimental design that tests the benefit of deterministic forecasts and probabilistic forecasts on their own by presenting both independently from each other.

Finally, the main benefit of probabilistic information was observed in Situation 1 but not in all others, where we had it initially expected. Here, it is important to develop a better understanding of which situations probabilistic forecasts contain information that can improve decisions.

4.2 | Avenues for future research

There are a number of promising avenues to further develop this research.

First, to better understand and quantify the benefit of probabilistic forecast, it is key to select a representative set of critical decision situations. On the one hand, this might require the use of complementary information (or indices) that characterize the ‘degree of uncertainty’ (see Bessa et al., 2017 for some examples) and study their correlation with past events (and context), where the use of probabilistic forecasts resulted in a significant positive benefit. On the other hand, it requires to investigate what cues and features human decisions-makers focus in a given forecast representation in order to predict a high-speed shutdown. Knowing the cues people rely on would allow to predict in which cases different decisions should be expected.

From a risk communication perspective, it is also key to investigate whether and how different representations affect decisions (e.g., Joslyn & LeClerc, 2013), and how to best communicate uncertainty information transparently (Spiegelhalter et al., 2011). For instance, would decisions differ, if forecasts were given in text or table format; or by showing all 75 members of the probabilistic forecasts as a ‘spaghetti plot’? More information is not always better: ‘Spaghetti plots’ do not generally add useful information, especially, when the uncertainty of the weather situation is high and ensemble members deviate strongly from each other (Bessa et al., 2017). They may highlight the uncertainty of certain, specific situations, for example, strong, short-lasting ramps, more clearly compared with percentile bands. However, in the HSSD case, spaghetti plots would not add more or better information to the task.

In the current experiment, 30% of the participant performed worse with the additional probabilistic forecasts than with deterministic forecasts alone. There are a number of possible reasons for this: they may have been unfamiliar with probabilistic forecast and information in general, they may have misunderstood the forecast representation, or they may have used a less appropriate decision strategy. One main focus in future studies thus will be to better understand why some people perform worse than other, in order to develop better representation formats and explanations to boost their decisions competencies.

The first step is to better understand which decision strategies are adaptive in order to assess, if decision-makers are risk averse or risk prone under uncertainty. On the one hand, individuals risk preferences vary. On the other hand, even in simplified decision situations, the expected value of an option is not necessarily equivalent to its expected utility for decision-makers. To study whether risk preferences change appropriately, one can model behaviour in probabilistic games with utility theory, for example, a 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 .

The second important step is to understand how decisions depend on the structure of the decision context. Depending on the context, strategies other than expected utility maximization may be more suited, for example, a more robust strategy that puts a stronger emphasis to avoid large losses, for example, to avoid insolvency (Doherty & Eeckhoudt, 1995). A risk-averse strategy can be unwarranted in one decision context, but be called for in another where optimizing a long-term expected gain is only the second goal after maintaining operative processes.

Third, whereas probabilistic wind forecasts are normally calibrated, wind power forecasts incorporate different sources of uncertainty. Thus, the overall value of

course depends on the quality of the forecast itself. Probabilistic forecasts are one important step to make this transparent. In that respect, it might be interesting to show the difference in methodology and application of different methods to generate probabilistic forecast, focusing on low probability, but high impact scenarios, such as:

- Scenarios with wind speed above cut-out value for multiple hours (e.g., 2–3 h) and where uncertainty forecasts generated with a statistical model are not able to capture this event since: (1) a similar event was never observed in historical data; (2) temporal dependency structure is not captured by the statistical model.
- Hours with extreme regulation of power prices (this sometimes occurs in Nord Pool market) and where forecast errors in one direction can be highly penalized (even if they have small deviations).

Another important step is to investigate where non-psychological barriers to the application of probabilistic forecast in the operational environments exist. Possible barriers may be the costs of implementation, training of staff, lack of knowledge about available applications or lack of proper communication about the possibilities and advantages of using such forecasts. Finally, energy traders and transmission system operators generally handle a portfolio of RES power plants and this aggregation mitigates the impact of forecast uncertainty (see e.g., (Miettinen & Holttinen, 2017) and (Pahlow et al., 2009) section ‘Pooling of energy’). In contrast, distribution system operators and local energy communities are more exposed to RES uncertainty and adverse weather, in particular its impact in electrical grid technical constraints. This motivates the need to show to end-users the added value of uncertainty information for finding remedial actions that solve predicted technical problems. An experiment similar to the one presented in this paper would be highly relevant since the impact of RES uncertainty is an indirect result (i.e., influence in voltage and current in the grid) and therefore not straightforward to be understood by a human operator in a control centre.

5 | CONCLUSIONS

This experiment revealed a promising way of examining how to overcome psychological barriers to the adoption of probabilistic forecasts in the operation of the electric grid and in power trading applications. In total, there were 105 participants, mostly experts in the energy industry or from the meteorological community with relation

to the energy industry. Although the experiment was simplified, it still provided a realistic scenario for many decision-makers in the industry, and for this reason was well received by the participants as an exemplary application for the use and application of probabilistic forecasts from a physical-based ensemble weather prediction system, and considered a useful tool for training purposes. The feedback we received also confirmed us in our approach to build on ‘decision from experience’ and gamification. To create simplified representations of reality, also described in the hydrological science (Ramos et al., 2018), increases the engagement level of people, when learning to use a new technology. The *learning-by-doing* and ‘playing’ a game in a safe environment may not reproduce the entire context of a decision-making problem in a specific operational environment, but instead provides a platform to introduce people with a complex topic, train and teach awareness for the challenges and benefits that come with the advanced technology.

The results hence revealed a number of interesting aspects of decision-making in the energy industry, and also how complex the evaluation of good or bad decision-making is for real-life applications. Nevertheless, the results encourage a further development of this experiment by increasing the number of cases to create a more representative distribution. This will also allow us to evaluate how we could use experimental games like this as training tools that enable decisions-makers to learn from experience how to deal with uncertainty in a structured and state-of-the-art way.

AUTHOR CONTRIBUTIONS

Corinna Möhrle: Conceptualization (equal); data curation (equal); investigation (lead); methodology (equal); project administration (lead); resources (equal); validation (equal); visualization (lead); writing – original draft (lead); writing – review and editing (lead). **Ricardo J. Bessa:** Methodology (supporting); validation (supporting); writing – review and editing (equal). **Nadine Fleischhut:** Conceptualization (equal); data curation (equal); formal analysis (lead); investigation (supporting); methodology (equal); resources (equal); software (lead); supervision (lead); validation (equal); visualization (equal); writing – review and editing (equal).

ACKNOWLEDGMENT

Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

The authors declare no conflict of interest. Aside from research and development (e.g., as member of the board of the IEA Wind Task 36), Corinna Möhrle is Co-

founder and director of WEPROG ApS (GmbH, Ltd.), a company that developed and operates an Ensemble Prediction System (MSEPS) based on a multi-scheme approach and provides solutions and services to users with a focus area in renewable energies. The forecasts used in this study are from WEPROG’s MSEPS and are from an existing wind farm. To keep WEPROG’s customer data confidential, outcomes are only given in wording and no measurements are shown.

ORCID

Corinna Möhrle  <https://orcid.org/0000-0002-8842-1533>

Ricardo J. Bessa  <https://orcid.org/0000-0002-3808-0427>

Nadine Fleischhut  <https://orcid.org/0000-0002-6807-3455>

REFERENCES

- Bessa, R.J., Möhrle, C., Fundel, V., Siefert, M., Browell, J., Gaidi, S.H.E. et al. (2017) Towards improved understanding of the applicability of uncertainty forecasts in the electric power industry. *Energies*, 10, 1402.
- Bolton, G.E. & Katok, E. (2017) Cry wolf or equivocate? Credible forecast guidance in a cost-loss game. *Management Science*, 64, 983–1476.
- Botterud, A., Wang, J., Zhou, Z., Bessa, R.J., Keko, H. & Miranda, J.S.A.V. (2012) Wind power trading under uncertainty in LMP markets. *IEEE Transactions on Power Systems*, 27, 894–903.
- Budescu, D.V., Broomell, S. & Por, H.-H. (2009) Improving communication of uncertainty in the reports of the intergovernmental panel on climate change. *Psychological Science*, 20, 299–308.
- Budescu, D.V., Por, H.-H., Broomell, S.B. & Smithson, M. (2014) The interpretation of ipcc probabilistic statements around the world. *Nature Climate Change*, 4, 508–512.
- Cali, Ü. (2010) *Grid and market integration of large-scale wind farms using advanced wind power Forecasting: technical and energy economic aspects*. PhD thesis, University of Kassel, Kassel, Germany. Available at: <https://nbn-resolving.org/urn:nbn:de:0002-30319>.
- Chandramowli, S.N. & Felder, A. (2014) Impact of climate change on electricity systems and markets – a review of models and forecasts. *Sustainable Energy Technologies and Assessments*, 5, 62–74.
- CMC (2016) Training on Ensemble Prediction System of the Meteorological Service of Canada. Available at: <http://collaboration.cmc.ec.gc.ca/cmc/ensemble/Formation-Training/Read-me.html>.
- Dai, T. & Qiao, W. (2015) Optimal bidding strategy of a strategic wind power producer in the short-term market. *IEEE Transactions on Sustainable Energy*, 6, 707–719.
- Demeritt, D., Nobert, S., Cloke, H. & Pappenberger, F. (2010) Challenges in communicating and using ensembles in operational flood forecasting. *Meteorological Applications*, 17, 209–222.
- Doherty, N.A. & Eeckhoudt, L. (1995) Optimal insurance without expected utility: the dual theory and the linearity of insurance contracts. *Journal of Risk and Uncertainty*, 10, 157–179.

- Fleischhut, N., Herzog, S.M. & Hertwig, R. (2020) Weather literacy in times of climate change. *Weather, Climate, and Society*, 12, 435–452.
- Frick, J. & Hegg, C. (2011) Can end-users' flood management decision making be improved by information about forecast uncertainty? *Atmospheric Research*, 100, 296–303.
- Fundel, V.J., Fleischhut, N., Herzog, S.M., Göber, M. & Hagedorn, R. (2019) Promoting the use of probabilistic weather forecasts through a dialogue between scientists, developers and end-users. *Quarterly Journal of the Royal Meteorological Society*, 145, 210–231.
- Giebel, G., Brownsword, R., Kariniotakis, G., Denhard, M. & Draxl, C. (2011) *The state of the art in short-term prediction of wind power – a literature overview*. 2nd edition. Technical Report, ANEMOS.plus Deliverable D1.2. Available at: <https://orbit.dtu.dk/en/publications/the-state-of-the-art-in-short-term-prediction-of-wind-power-a-lit>.
- Gigerenzer, G. & Brighton, H. (2009) Homo heuristics: why biased minds make better inferences. *Topics in Cognitive Science*, 1, 107–143.
- Gigerenzer, G.E., Hertwig, R.E. & Pachur, T.E. (2011) *Heuristics: the foundations of adaptive behavior*. Oxford: Oxford University Press.
- Grimit, E. (2017) *Providing uncertainty information to end users in the electric sector*. UVIG Forecasting workshop. Tutorial on integration of uncertainty forecasts into power system operations, Atlanta, GA.
- Grounds, M.A. & Joslyn, S. (2018) Communicating weather forecast uncertainty: do individual differences matter? *Journal of Experimental Psychology: Applied*, 24, 18–33.
- Haupt, S.E., Garcia Casado, M., Davidson, M., Dobschinski, J., Du, P., Lange, M. et al. (2019) The use of probabilistic forecasts: applying them in theory and practice. *IEEE Power and Energy Magazine*, 17, 46–57.
- HEPEX (2021) 3 HEPEX “Pathways to running a flood forecasting centre: an adventure game!”. Available from: <http://textadventures.co.uk/games/view/vfl90ntyh0yo2vr0wkeqea/pathways-to-running-a-flood-forecasting-centre-an-adventure-game>.
- Hertwig, R. & Erev, I. (2009) The description–experience gap in risky choice. *Trends in Cognitive Sciences*, 13, 517–523.
- Hertwig, R. & Wulff, D.U. (2021) A description–experience framework of the psychology of risk. *Perspectives on Psychological Science*, 17, 174569162111026896.
- Hong, T., Pinson, P., Wang, Y., Weron, R., Yang, D. & Zareipour, H. (2020) Energy forecasting: a review and outlook. *IEEE Open Access Journal of Power and Energy*, 7, 376–388.
- Houtekamer, P.L., Lefaiivre, L., Derome, J., Ritchie, H. & Mitchell, H.L. (1996) A system simulation approach to ensemble prediction. *Monthly Weather Review*, 124, 1225–1242.
- Joslyn, S.L. & LeClerc, J.E. (2012) Uncertainty forecasts improve weather-related decisions and attenuate the effects of forecast error. *Journal of Experimental Psychology: Applied*, 18, 126.
- Joslyn, S. & LeClerc, J. (2013) Decisions with uncertainty: the glass half full. *Current Directions in Psychological Science*, 22, 308–315.
- Joslyn, S., Nemeč, L. & Savelli, S. (2013) The benefits and challenges of predictive interval forecasts and verification graphics for end users. *Weather, Climate, and Society*, 5, 133–147.
- Joslyn, S. & Savelli, S. (2010) Communicating forecast uncertainty: public perception of weather forecast uncertainty. *Meteorological Applications*, 17, 180–195.
- Kahneman, D., Slovic, S.P. & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge university press.
- Kalke, K., Studd, H. & Scherr, C.L. (2021) The communication of uncertainty in health: a scoping review. *Patient Education and Counseling*, 104, 1945–1961. Available at: <https://www.sciencedirect.com/science/article/pii/S0738399121000616>
- Kox, T., Gerhold, L. & Ulbrich, U. (2015) Perception and use of uncertainty in severe weather warnings by emergency services in Germany. *Atmospheric Research*, 158, 292–301.
- LeClerc, J. & Joslyn, S. (2015) The cry wolf effect and weather-related decision making. *Risk Analysis*, 35, 385–395.
- Lin, J., Zhang Sun, Y., Cheng, L. & Zhong Gao, W. (2012) Assessment of the power reduction of wind farms under extreme wind condition by a high resolution optimization model. *Applied Energy*, 96, 21–32.
- Mackintosh, N. & Armstrong, N. (2020) Understanding and managing uncertainty in health care: revisiting and advancing sociological contributions. *Sociology of Health & Illness*, 42, 1–20. <https://doi.org/10.1111/1467-9566.13160>
- Marimo, P., Kaplan, T.R., Mylne, K. & Sharpe, M. (2015) Communication of uncertainty in temperature forecasts. *Weather and Forecasting*, 30, 5–22.
- Mazzi, N. & Pinson, P. (2017) Wind power in electricity markets and the value of forecasting. In: Kariniotakis, G. (Ed.) *Woodhead Publishing Series in Energy. Renewable Energy Forecasting. From Models to Applications*. Duxford, UK: Woodhead Publishing, pp. 259–278. Available at: <https://www.sciencedirect.com/science/article/pii/B978008100504000010X>
- Mengelkamp, H.-T. (1988) On the energy output estimation of wind turbines. *International Journal of Energy Research*, 12, 113–123.
- Miettinen, J.J. & Holttinen, H. (2017) Characteristics of day-ahead wind power forecast errors in nordic countries and benefits of aggregation. *Wind Energy*, 20, 959–972. <https://doi.org/10.1002/we.2073>
- Möhrlén, C. (2004) *Uncertainty in wind energy forecasting*. PhD thesis, National University of Ireland, Cork, Ireland. Available at: <http://library.ucc.ie/record=b1501384>.
- Möhrlén, C., Pahlow, M. & Jørgensen, J.U. (2012) Untersuchung verschiedener handelsstrategien für wind- und solarenergie unter berücksichtigung der eeg 2012 novellierung. *Zeitschrift für Energiewirtschaft*, 36, 9–25.
- Molteni, F., Buizza, R., Palmer, T.N. & Petroliagis, T. (1996) The ECMWF ensemble prediction system: methodology and validation. *Quarterly Journal of the Royal Meteorological Society*, 122, 73–119. <https://doi.org/10.1002/qj.49712252905>
- Moreno, R., Panteli, M., Mancarella, P., Rudnick, H., Lagos, T., Navarro, A. et al. (2020) From reliability to resilience: planning the grid against the extremes. *IEEE Power and Energy Magazine*, 18, 41–53.
- Murphy, A.H. (1985) Decision making and the value of forecasts in a generalized model of the cost-loss ratio situation. *Monthly Weather Review*, 113, 362–369.
- Mylne, K.R. (2002) Decision-making from probability forecasts based on forecast value. *Meteorological Applications*, 9, 307–315. <https://doi.org/10.1017/S1350482702003043>
- Pahlow, M., Möhrlén, C. & Jørgensen, J. (2009) *Application of cost functions for large-scale integration of wind power using a multi-scheme ensemble prediction technique*. *Optimization advances in*

- electric power systems*. Castronuovo: Nova Science Publishers, pp. 151–180.
- Pinson, P. (2012) Adaptive calibration of (u,v)-wind ensemble forecasts. *Quarterly Journal of the Royal Meteorological Society*, 138, 1273–1284.
- Raftery, A.E. (2016) Use and communication of probabilistic forecasts. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 9, 397–410.
- Ramos, M.H., van Andel, S.J. & Pappenberger, F. (2013) Do probabilistic forecasts lead to better decisions? *Hydrology and Earth System Sciences*, 17, 2219–2232.
- Ramos, M.-H., van Andel, S.J., Pappenberger, F., Crochemore, L., Wood, A., Zappa, M., et al. (2018) Role-play games to advance probabilistic forecasting in hydrology. *EGU General Assembly Conference Abstracts*, 15793.
- Renooij, S. & Witteman, C. (1999) Talking probabilities: communicating probabilistic information with words and numbers. *International Journal of Approximate Reasoning*, 22, 169–194. Available at: <https://www.sciencedirect.com/science/article/pii/S0888613X99000274>
- Roulston, M.S., Bolton, G.E., Kleit, A.N. & Sears-Collins, A.L. (2006) A laboratory study of the benefits of including uncertainty information in weather forecasts. *Weather and Forecasting*, 1, 116–122.
- Roussos, J., Bradley, R. & Frigg, R. (2021) Making confident decisions with model ensembles. *Philosophy of Science*, 88, 439–460.
- SMH (2021) Call for water game. Available from: <https://hypeweb.smhi.se/showcases/desktop-call-for-water-game/>.
- Spiegelhalter, D., Pearson, M. & Short, I. (2011) Visualizing uncertainty about the future. *Science*, 333, 1393–1400.
- Stephens, E.M., Spiegelhalter, D.J., Mylne, K. & Harrison, M. (2019) The met office weather game: investigating how different methods for presenting probabilistic weather forecasts influence decision-making. *Geoscience Communication*, 2, 101–116.
- Toth, Z. & Kalnay, E. (1997) Ensemble forecasting at NCEP and the breeding method. *Monthly Weather Review*, 125, 3297–3319.
- van der Wiel, K., Stoop, L.P., van Zuijlen, B., Blackport, R. & van den Broek, M. (2019) Meteorological conditions leading to extreme low variable renewable energy production and extreme high energy shortfall. *Renewable and Sustainable Energy Reviews*, 111, 261–275.
- WEXICOM (2019) “Feuerwache”: Wie gut verstehen Sie die Unsicherheit von Wettervorhersagen? Probabilistic forecast game presented at the 11th Industry and Research Platform Prognosis (IFP) meeting, Kassel, Nov. 2019
- Wintle, B.C., Fraser, H., Wills, B.C., Nicholson, A.E. & Fidler, F. (2019) Verbal probabilities: very likely to be somewhat more confusing than numbers. *PLoS One*, 14, 1–18. <https://doi.org/10.1371/journal.pone.0213522>
- World-Meteorological-Organisation. (2012) *WMO guidelines on ensemble prediction systems and forecasting*. Technical Report. 1091. Available at: https://library.wmo.int/index.php?lvl=notice_display&id=12962#.Ya3jhLso_JF.
- Würth, I., Valdecabres, L., Simon, E., Möhrlen, C., Uzunoğlu, B., Gilbert, C. et al. (2019) Minute-scale forecasting of wind power—results from the collaborative workshop of IEA wind task 32 and 36. *Energies*, 12, 712.

How to cite this article: Möhrlen, C., Bessa, R. J., & Fleischhut, N. (2022). A decision-making experiment under wind power forecast uncertainty. *Meteorological Applications*, 29(3), e2077. <https://doi.org/10.1002/met.2077>