CrossMark

Rewarding cooperative virtual power plant formation using scoring rules ☆

Valentin Robu [a, *], Georgios Chalkiadakis [b], Ramachandra Kota [c], Alex Rogers [d], Nicholas R. Jennings [e]

[a] Institute of Sensors, Signals and Systems, Heriot-Watt University, Edinburgh, UK
[b] School of Electronic and Computer Engineering, TU Crete, Chania, Greece
[c] IBM Research, Bangalore, India
[d] Department of Computer Science, University of Oxford, Oxford, UK
[e] Central Faculty, Office of the Provost, Imperial College, London, UK

ABSTRACT

Virtual Power Plants (VPPs) are fast emerging as a viable means of integrating small and distributed energy resources (DERs), like wind and solar, into the electricity supply network (Grid). VPPs are formed via the aggregation of a large number of DERs, so that they exhibit the characteristics of a traditional generator in terms of predictability and robustness. In this work, we promote the formation of such "cooperative" VPPs (CVPPs) using techniques from the field of distributed Artificial Intelligence and game theory. In particular, we design a payment mechanism that encourages DERs to join CVPPs with increased size and visibility to the network operator. Our method is based on strictly proper scoring rules and incentivises the provision of accurate predictions of expected electricity generation from member DERs, which aids in the planning of the supply schedule at the Grid. We empirically evaluate our approach using the real-world setting of 16 commercial wind farms in the UK, and we show that it incentivises real DERs to form CVPPs, and outperforms the current state of the art payment mechanism developed for this problem.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

In recent years, a number of strands of Artificial Intelligence (AI) research have taken up the challenge of creating smart and robust electricity supply networks, which can make efficient use of all available energy resources, thereby reducing dependence on carbon-intensive conventional generators. In particular, Ramchurn et al. [3] provide strong arguments for the role that new AI techniques (the "smarts") play in shaping the smart grid research agenda. Kok at al [4]. discuss how intelligent techniques can be used to integrate distributed renewable resources in existing distribution networks, while Pudjianto et al. [5] focus on the role of smart technologies in enabling Virtual Power Plants. Moreover, mechanism design and AI techniques have been used to explore the incentives in smart grid contexts, including in the work of Robu et al. [6] for the coordination of electric vehicle charging, or the work of Ma et al. [7] on incentivising reliability in demand-side response.

One such crucial topic we explore here is the integration of distributed renewable energy resources. While environmental concerns are becoming increasingly important, the overriding concern of national electricity transmission network operators (termed *the Grid* herein, following the standard in the UK) remains the *reliability of supply*. In particular, the Grid is responsible for ensuring that energy demand is met without interruptions, by dispatching power plants to produce and supply energy whenever it is required.

Now, although reliability is easily addressed when energy is produced solely by conventional power plants (which can be dispatched as required, often at short notice), the problem becomes pressing when plants utilizing *renewable energy* sources are

---

involved. In the last decade, *distributed energy resources (DERs)*—essentially small to medium capacity (2kW-2MW) renewable energy generators—have begun to appear in greater numbers in the network. Though their deployment could in principle reduce reliance on conventional power plants significantly [5], their integration into the Grid is problematic since the DERs, given their small size, are largely "invisible" to the Grid. This means they cannot readily be taken into account while planning production schedules, even if their total energy production represents a significant amount. Even if they were visible to the Grid, the uncertainty and intermittency of renewable energy sources inhibits individual DERs from profitably dealing with it directly, or participating in the wholesale electricity market because they are often unable to meet the set generation targets. Nevertheless, as shown in a number of existing works, the need to incorporate renewable energy resources into the existing Grid is a pressing one. Specifically, the importance of incorporating distributed energy resources has been argued by Asmus et al. [8], who compare different approaches to tackle this problem, especially microgrids and virtual power plants, as well as by Ramchurn et al. [3]. Integrating renewable energy resources also plays a crucial role in load balancing, and in this vein, Tarroja et al. [9] discuss different metrics for evaluating the impact of intermittent renewable generation on utility load balancing.

The approach adopted by many countries to incentivise more renewable generation is to encourage small-scale renewable energy producers with payments according to specific *feed-in tariffs (FITs)*. FITs are typically set at significantly higher levels than market prices (an approach adopted in many EU countries, among others). However, with DER numbers projected to be in the range of hundreds of thousands in a single country, the use of feed-in tariffs is increasingly seen as an unsustainable long-term policy [10].

An alternative is to aggregate individual DERs together to form larger energy generating entities. Such entities then have the opportunity to become economically sustainable by overcoming the invisibility and unreliability problems identified above. This has led several researchers to propose the creation of *Virtual Power Plants (VPPs)*, which consist of large numbers of DERs, and thus have the potential to be viewed as the virtual equivalents of conventional power stations. The concept is extensively discussed in the work of Pudjianto et al. [5], who distinguish between Technical Virtual Power Plants and Commercial Virtual Power Plants, and present the advantages of both. Our work fits into the later category, since our focus is on a group of distributed generators selling their energy together as a single entity. Dimeas and Hatziargyriou [11] present a case study application of a practical VPP. Finally, Vasirani et al. [12] consider the use of VPPs to integrate renewable generations from wind with storage from electric vehicles. Unlike this work, however, they do not consider incentives for accurate reporting.

In more detail, in this paper, we focus our attention on the concept of *Cooperative VPPs (CVPPs)*. CVPPs are coalitions of individual DERs that sell their energy jointly to the Grid/DNO as a single coalition (and also provide a single estimate for the production of the whole group, either as a sum of the estimates or as a joint probability distribution). Individual DERs participating in such a cooperative are modeled as game-theoretically "rational" agents, which have the choice of interacting with the Grid by themselves (independently), but choose to pool together resources and predictions with other agents because they have an incentive to do so. Here, the CVPP itself acts as a coordinating agent, whose goal is to incentivise collaboration and truthful/reliable behaviour from the participating DERs. However, while the CVPP agent aims to achieve the best total payoff for its members in dealing with the Grid, it does not aim to make a profit for itself, beyond what is distributed among member DERs.

Recently, Chalkiadakis et al. [1] proposed a pricing mechanism that can be used by the Grid to promote the creation of CVPPs, and act as an alternative payment mechanism from the Grid to feed-in tariffs. A critical limitation of that approach, however, is that a CVPP only presents the Grid with point (mean) estimates of its production. Unfortunately, single point estimates do not provide any indication of how uncertain these estimates may be and taking them at face value runs the risk that predictions can be widely off. An alternative that is more useful to the Grid is that, along with the production estimates, the distribution of potential prediction errors of these estimates is also provided, specifying the *confidence* individual entities place in their estimates. This additional information enables the Grid to optimise the scheduling of all available generators; since it would now be aware of the probability of renewables not meeting the targets, necessitating the dispatch of conventional generators. Depending on the confidence placed on the estimates, the Grid is able to choose the appropriate number of conventional generators needed on standby.

While providing accurate estimates is desirable, it comes with a cost for individual generators (DERs). In particular, they may need to invest in monitoring and smart metering equipment, and in systems which can collect and analyse data from both their own wind turbines, and also other sources such as nearby weather stations. At the moment, since there is no incentive for DERs to provide accurate prediction (they get paid just per unit of generation), very few will actually make this investment. Hence, there is a real need for the right incentive mechanisms that will encourage individual renewable investors (DERs) to make this transition, and contribute not just to additional generation, but also to the reliability of the electricity supply.

Now, scoring rules have long been used to design payment mechanisms that incentivise agents to report private probabilistic predictions truthfully and to the best of their forecasting abilities—via appropriately "rewarding" agents that achieve high scores and "punishing" them otherwise. The concept of scoring rules was first introduced in the seminal work of Savage [13], while Matheson and Winkler [14] first applied it to elicit continuous probability distributions. The work of Gneiting and Rafterry [15] provides a broad overview of scoring rule techniques, and provides a set of mathematical tools which can be used in a variety of domains. Finally, Hersbach [16] shows how the continuous ranked probability score (CRPS), a particular form of scoring rule which we also use, can be used to construct ensemble predictions of weather forecasts. However, his work does not consider the issue of distributed energy generation and virtual power plants.

In more detail, in this work, we would like to design a mechanism which can guarantee that agents will accurately declare privately calculated distributions which reflect their confidence in their own forecast. Without such a mechanism in place, agents may either lie about their estimates to secure higher returns or not bother to provide the most accurate estimates. To counter such trends, *strictly proper* scoring rules are used here to guarantee the so-called *incentive compatibility* (or *truthfulness*) of the agents in reporting their estimates [15]. This means that, for agents participating in such a mechanism, the best strategy is to declare truthfully the distributions reflecting the uncertainty in their predictions. Any other strategy only results in lower returns. Additionally, it incentivises them to provide as accurate estimates as they possibly can.

Taking inspiration from this, we provide the first application of a scoring rules-based mechanism in the renewable energy domain. Specifically, we put forward a payment mechanism that uses a strictly proper scoring rule to incentivise CVPPs—and, in turn, DERs—to provide the Grid with their true expected production and the true estimated probability distribution representing their confidence (or, equivalently, expected error estimates) regarding

their production predictions. The mechanism guarantees that DERs are rewarded for providing estimates that are both accurate and have a high confidence (ensuring that agents are given credit for high probability estimates that are close to the realised ones).

Another key contribution of this paper lies in the experimental analysis of the proposed mechanism. We base our experimental setting on 16 real-world wind farms, distributed around the UK. For these farms, we collect a 10-week dataset of both wind speed predictions and actual wind speeds, for each half hourly settlement period. We used these in conjunction with a model of the characteristics of the wind turbines employed, in order to create both predictions and measurements of the production at these sites. Thus, our experimental conclusions are based on real data.

The rest of this paper is organised as follows. Section 2 presents the formal model of our setting and the role of CVPPs. Section 3 discusses scoring rules and their properties, and presents our payment mechanism. Section 4 details our experimental study and Section 5 concludes.

2. Energy producers and cooperatives

We consider a setting that contains several independent, distributed energy producers (DERs) that can sell their energy directly to the Grid or opt to join a cooperative (CVPP). The CVPP's primary function is to represent its DER members while interacting with the Grid: the CVPP provides the aggregate estimate of the members' production to the Grid, receives the corresponding payment, and distributes it amongst members in some fair manner. The model also assumes that the day is divided into *settlement periods* corresponding to electricity trading intervals (typically 48 half-hour time slots).

Formally, for any settlement period t, each DER i produces a certain amount of energy $prod_{i,t} \in \mathbb{R}_+$ (in kWh). It can also estimate in advance an *expected production* value $\widetilde{prod}_{i,t} \in \mathbb{R}_+$. The method used by a DER to compute this estimate (and its accuracy) depends on its type of generation. For instance, DERs composed of (one or more) wind turbines—such as the ones considered in this paper's analysis—can estimate their $\widetilde{prod}_{i,t}$ based on an hourly wind prediction obtained for their area from the UK meteorological office (as described in detail in Section 4).

Now, each DER can compute a relative prediction error for each of the past settlement periods for which it has historical data over some time horizon T, as follows:

$$e_{i,t} = \frac{prod_{i,t} - \widetilde{prod}_{i,t}}{\widetilde{prod}_{i,t}}, \forall t \in T \tag{1}$$

Note that in Eq. (1), the actual production $prod_{i,t}$ is the random variable to be predicted (unknown in advance), while $\widetilde{prod}_{i,t}$ is the average prediction for this variable.

In practice, there may be wide variances because some DERs may be able to better estimate their future production than others. For example, DERs that use tidal energy tend to be more accurate in estimation than those that use wind. Even among wind-based DERs, there may be substantial differences in the accuracy of predictions, due to the age/characteristics of their turbines, or the geographical area the DER is in. Wind may be easier to predict in some areas than in others, because the meteorological office provides more accurate and timely predictions for some areas (as these might be closer to existing weather monitoring stations).

Formally, $\boldsymbol{E}_i = \langle e_{i,t} \rangle$, $t \in T$ denotes the vector of relative prediction errors that DER i makes over some time horizon T (e.g., if the period of study is 30 days, T contains $30 \times 48 = 1440$ half-hourly time periods). Note however that the actual production $prod_{i,t}$ for

the current time period t is yet unobserved—and hence the error $e_{i,t}$ is not known to the producer either. If DER i is a *perfect predictor*, then the vector $\boldsymbol{E}_i$ would contain only zeros; however, in practice, it will always contain a number of negative values (over-prediction) and/or positive values (under-prediction). If DER i corresponds to an *unbiased predictor* (i.e., one which neither systematically over-predicts or under-predicts), then we could expect the values in $\boldsymbol{E}_i$ to have a mean of zero and follow a normal distribution $\mathcal{N}(\mu = 0, \sigma_i^2)$. In practice, however, one *cannot* assume that the predictions can be always approximated as a normal distribution, or that they are always unbiased. In this paper, we develop a novel method for constructing the distribution empirically, without requiring any normality assumption.

Note that the above definitions can be naturally extended to a CVPP, in which a number of DERs sell their production jointly, and make a joint prediction. Formally, the actual and estimated production of a cooperative C at t are denoted by $prod_{C,t}$ and $\widetilde{prod}_{C,t}$ respectively. If I denotes the set of the participating CVPP members, the total CVPP production at t can be computed as the sum total of its members' production, i.e. $prod_{C,t} = \sum_i prod_{i,t}$. Similarly, the estimate of the CVPPs production at time t is the sum of the estimates of the individual DERs at time t, i.e. $\widetilde{prod}_{C,t} = \sum_i \widetilde{prod}_{i,t}$. The vector $\boldsymbol{E}_C = \langle e_{C,t} \rangle$ is defined in the same way as in the case of a DER.

3. The payment mechanism

In order to reward agents for accurate reports of their uncertainty, we design a payment mechanism which employs *scoring rules*. A *scoring rule* is a real-valued function $S(\widehat{F}, x)$, specifying the reward that a forecaster agent i should receive if it reports a predicted distribution $\widehat{F}$ over the probability of some future event, and the event x occurs (in our case, $\mathbf{x} \in \mathbb{R}$).

Scoring rules with certain properties can be of significant value to a mechanism designer. In particular, *strict propriety* is one such important property [15,13].

Definition 1. *A scoring rule $S(\widehat{F}, x)$ rewarding a prediction with distribution $\widehat{F}$ and realisation x, is called strictly proper if it incentivises a truthful report, meaning reporting the true distribution F is the only report that maximizes expected reward. Formally stated, if F is the true underlying distribution of the random variable x, scoring rule S is strictly proper if $S(F, x) \geq S(\widehat{F}, x)$, with the equality holding if and only if $\widehat{F} = F$.*

While in this paper we use strict propriety, similarly, a scoring rule is said to be proper if $S(F, x) \geq S(\widehat{F}, x)$, but the prediction $\widehat{F} = F$ is not necessarily the only one that maximizes $S(\widehat{F}, x)$. In our case, the use of a strictly proper scoring rule has a specific interpretation: it means that energy suppliers can expect to maximize their payments *if and only if* they accurately report their expectation over the prediction error they can potentially make. Here, the distribution F is constructed from the vector of past prediction errors $\boldsymbol{E}_i$, while observation x is the current observed error $e_{i,t}$. The reason why strict propriety is such a desirable property in our CVPP context is that we assume that member DERs are rational agents, that maximize their own expected utility. Hence, if they would receive a higher utility from reporting different than truthfully, this could potentially break the accuracy of the joint CVPP prediction and payment received by the CVPP from the Grid, and hurt everyone in the coalition.

3.1. Continuous ranked probability score

Many of the proper and strictly proper scoring rules proposed in existing literature [13] do not satisfy the requirements of our application, because they are not *sensitive to distance*—i.e., no credit

is given to agents for predictions assigning high probabilities to values that are close, but not identical, to the realised value. This is a necessary requirement for us, because being far off the predicted production amount is much more detrimental than being only slightly off. However, these desirable characteristics are possessed by the *Continuous Ranked Probability Score* (CRPS) [14], which is a strictly proper scoring rule used for continuous variables. This particular scoring rule has lately attracted renewed interest in the scoring rules literature, and has been used to quantify weather predictions [16] (such as wind speeds, tides or solar intensity) which are essential for renewable generation as well. We have therefore chosen this rule as the basis of our payment mechanism.

Formally, CRPS is defined as (c.f. [15], Eq. (2)):

$$CRPS(F, x) = -\int_{-\infty}^{\infty} (F(y) - 1\{y \geq x\})^2 \mathrm{d}y \qquad (2)$$

where F is the reported underlying distribution of a random variable, and x is the actual observation of this variable, for the instance when the CRPS score is computed. Moreover, here the function $1\{y \geq x\}$ is an *indicator function*, which takes the value of 1 if the expression in the brackets ($y \geq x$ in this case) is true, and 0 otherwise. The following lemma is essentially a result from Gneiting and Rafterry [15] (Section. 4.2, Eq. (2)), which we'll make use in subsequent proofs in this paper.

Lemma 1. (c.f. [15]). *The CRPS rule as defined in* Eq. (2) *is strictly proper.*

Moreover, we note the CRPS score, as defined in Eq. (2) is always a negative number, with a maximum of 0 for a perfect predictor.

3.2. Constructing the expected error probability distribution

Consider the vector of historical observations $incr(\langle e_{i,j} \rangle), j = 1 \ldots N$, where $e_{j,i}$ are defined as in Eq. (1) above, while $incr$ is an operator that orders the vector of values in ascending order. Here, we consider the prediction error of DER i as a random variable drawn from some unknown distribution E_i, for which the e_i values are prior observations.[1] Now, we construct the cumulative distribution function (CDF) of this variable as follows. We consider the sequence of steps between each pair of $e_{i,j}$ and $e_{i,j+1}$ observations ordered in terms of value in an *increasing order* on the X-axis, and an equal distancing on the Y-axis of the respective probability density function (as, in the absence of other prior information, all observations are assigned an equal $\frac{1}{N}$ probability). So, for example, if the error vector contains $N = 3360$ half-hourly observations (as is the case for the real-world dataset used in this paper), the CDF E_i will be formed by 3360 small "steps", where some of these may be vertical (e.g., if the error vector contains many zeros, implying many instances of perfect prediction, those steps will be vertical). Note that the actual time step at which each error observation was made is not relevant here: the observation values are simply placed in the vector in an increasing order fashion, irrespective of the time step they correspond to. An illustration of this process is provided in Fig. 1 for three cases: a theoretical optimal predictor (for which the CDF is a perfect step function, with all probability mass at 0), a good predictor wind farm (i.e. the CDF close to the step function) and a poor predictor, chosen from our dataset.

Note that an alternative to this process would be to assume the distribution E_i must follow a normal distribution, and try to fit a sigmoid shape, symmetric w.r.t. the origin, through the empirical data points. However, the method we are proposing above is much more generic, and does not require any assumptions about the E_i distribution that the errors must follow. In fact, several of the distributions observed empirically (as seen in Fig. 3) are *not* symmetric, and deviate significantly from a normal—but our method can readily handle this fact.

Now, given the expected error E_i distribution empirically constructed as above, the CRPS score, defined in Eq. (2), can be actually computed as follows. Let x be the (yet unseen) error observation. Consider the empirically constructed prediction error vector $incr(\langle e_{i,j} \rangle), j = 1 \ldots N$ earlier introduced, containing N values in an ascending order; and let $pos(x)$ be the position of value x if inserted in that vector (that is, $x = e_{i,pos(x)}$). The value $E_i(x)$ is then the CDF of the observation x, given distribution E_i (i.e. the probability that the random variable will be smaller than x): $E_i(x) = \frac{pos(x)}{N}$, as observations have equal prior probability. Given this, the CRPS from Eq. (2) can be calculated as:

$$
\begin{aligned}
CRPS(E_i, x) = &-\sum_{k=1}^{pos(x)-2} (e_{i,k+1} - e_{i,k})\left(\frac{k}{N}\right)^2 - \Big(x \\
&- e_{i,pos(x)-1}\Big)\left(\frac{pos(x)}{N}\right)^2 - \Big(e_{i,pos(x)} \\
&- x\Big)\left(1 - \frac{pos(x)}{N}\right)^2 - \sum_{k=pos(x)}^{N-1} (e_{i,k+1} \\
&- e_{i,k})\left(1 - \frac{k}{N}\right)^2
\end{aligned}
\qquad (3)
$$

Intuitively, Eq. (3) expresses the integral in Eq. (2) as a discrete sum over the "step" intervals of the CDF function E_i on the X-axis. On the Y-axis, the squared distance to either 0 or 1 is taken, depending on the value of the indicator function following Eq. (2).
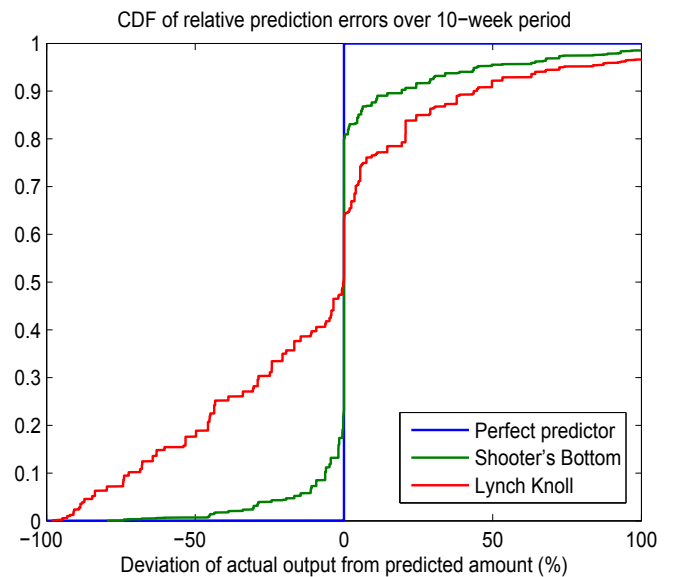
Fig. 1. Illustration of the empirical distribution E_i constructed from real error data for 3 DERs: 1. A theoretical optimal predictor, 2. A good predictor wind farm (Shooter's Bottom), 3. A poor predictor wind farm (Lynch Knoll).

---

[1] To avoid introducing additional notation, we slightly overload the notation E_i to refer to both the vector of past prediction errors, and the empirical distribution constructed directly from it. For the remainder of this paper, E_i refers to the expected distribution.

3.3. Defining the accuracy score using CRPS

The form of the CRPS function in both Eq. (2) and Eq. (3) above is scaled to take negative values, with a maximum of 0 (if the agent is always a perfect predictor). To be useful in the payment functions, we need a score scaled in the interval [0,1], with a maximum of 1 for a perfect predictor. Hence, we normalize the accuracy score as:

$$\mathscr{S}(\boldsymbol{E}_i, e_{i,t}) = \frac{1}{1 - CRPS(\boldsymbol{E}_i, e_{i,t})} \tag{4}$$

Note that, for ease of notation, we will denote the score of DER i at time t as $\mathscr{S}_{i,t} = \mathscr{S}(\boldsymbol{E}_i, e_{i,t})$, where $x = e_{i,t}$ is the new observation at time t predicted by $\boldsymbol{E}_i$. Also, note the same scoring equations can be applied to a CVPP C of DERs which sell energy jointly, by replacing the index i with C.

Lemma 2. *The accuracy score $\mathscr{S}$ defined in* Eq. (4) *above is strictly proper.*

Proof. The proof is by contradiction. Denoting by $\boldsymbol{E}_i^*$ the truthful report, assume there exists another report $\exists \boldsymbol{E}_i, \boldsymbol{E}_i \neq \boldsymbol{E}_i^*$ such that $\mathscr{S}(\boldsymbol{E}_i, e_{i,t}) > \mathscr{S}(\boldsymbol{E}_i^*, e_{i,t})$. This implies: $\exists \boldsymbol{E}_i$ s.t.

$$\frac{1}{1 - CRPS(\boldsymbol{E}_i, e_{i,t})} > \frac{1}{1 - CRPS(\boldsymbol{E}_i^*, e_{i,t})}$$

By definition (Eqs. (2) and (3)), *CRPS* is always a negative value (with a max at 0). Hence we can multiply both sides with $(1 - CRPS(\boldsymbol{E}_i, e_{i,t})(1 - CRPS(\boldsymbol{E}_i^*, e_{i,t}))$, yielding:

$$1 - CRPS(\boldsymbol{E}_i, e_{i,t}) < 1 - CRPS(\boldsymbol{E}_i^*, e_{i,t})$$

This means $\exists \boldsymbol{E}_i$, s.t. $CRPS(\boldsymbol{E}_i, e_{i,t}) > CRPS(\boldsymbol{E}_i^*, e_{i,t})$. This is a direct contradiction with the fact that CRPS is strictly proper (Lemma 1). □

Fig. 2 illustrates the accuracy score in Eq. (3), assuming that the expected error distribution $\boldsymbol{E}_i$ of a DER i follows a perfectly symmetric, normal distribution centered at zero: $\boldsymbol{E}_i = N(\mu_i = 0, \sigma_i)$. In this particular case, as shown in Ref. [15], Eq. (2) reduces to a function that depends only on $e_{i,t}$ and σ_i (again, note that the CRPS form we use in this paper also allows us to work with historical prediction errors that do not follow such normality). However, it is interesting to observe here how this error varies for different values of reported standard deviation σ_i. If DER i is highly confident in its predictions (reporting $\sigma_i = 0$), the maximum reward for accuracy can be achieved, but only if the actual error is also close to 0. However, if the actual relative error is high, then being truthful, by reporting a higher σ_i (i.e. less confidence), provides a better reward. This is the essential property of scoring rules: reporting truthfully the past distribution of prediction errors is always the agent's best available strategy.

3.4. Payment from the grid to individual producers or CVPP

We now present our payment mechanism. We first define the *"Grid-to-CVPP"* pricing function providing payments for the energy supplied by a CVPP to the Grid (or, in fact, any DER i that chooses to sell directly to the Grid). We then present the *"CVPP-to-DER"* redistribution function, used by the CVPP to distribute the received payments among its members internally.

First, we denote the electricity *base price* per kWh produced by π_B. This is a price paid per kWh of electricity produced, regardless
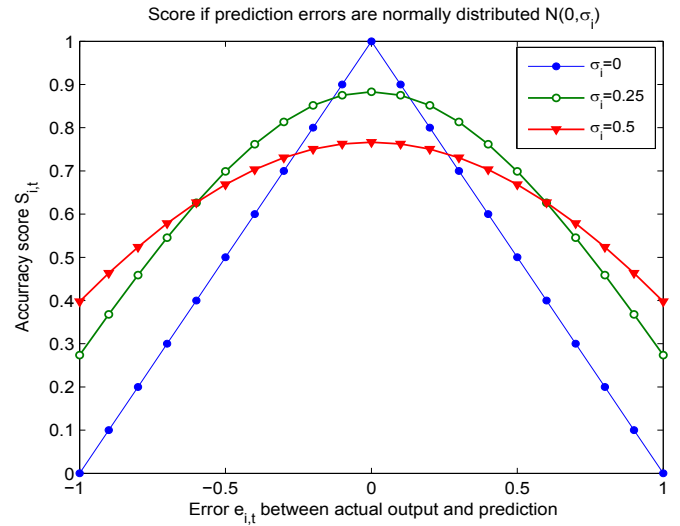
Fig. 2. Illustration of accuracy factor $\mathscr{S}_{i,t}$ assuming past prediction errors of DER i follow a perfect normal distribution $N(0,\sigma_i)$.
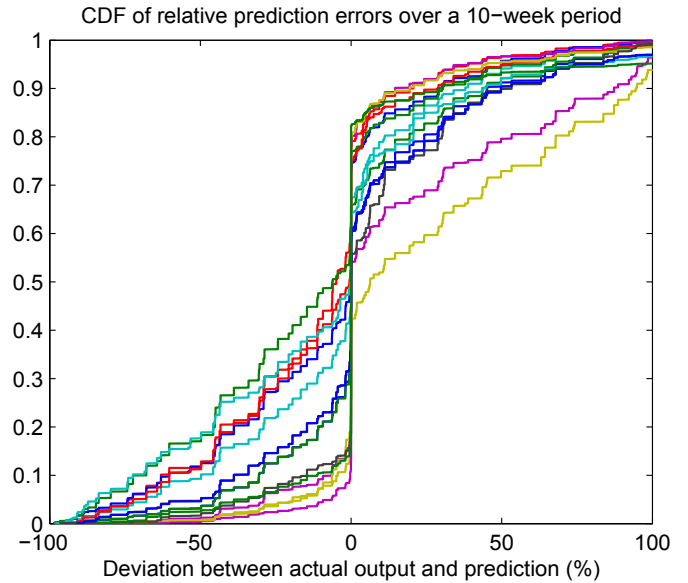

Fig. 3. Cumulative distribution functions of the prediction errors for the 16 wind farms over the 3360 half-hourly time points (10 weeks), when prediction is requested 4 h in advance.

of the predictability of the producing DER (as used by most feed-in tariff schemes[2]). Denoting by $prod_{i,t}$ the production of i at time slot t, and by $\mathscr{S}_{i,t}$ the accuracy score (as defined in Eq. (4) above), the payment from the Grid to DER i is then:

$$\pi_t^{G,i} = prod_{i,t} \times \pi_B \times \mathscr{S}_{i,t} \tag{5}$$

The same payment function applies from the Grid to a CVPP indexed by C (if a group of DERs sell energy together making a joint prediction), by simply replacing index i with C. Essentially, since $\mathscr{S}_{i,t} \in [0, 1]$, the price per unit $\pi_B \times \mathscr{S}_{i,t}$ paid to DER i is a fraction of the baseline price π_B, based on the accuracy. Moreover, this payment rule incentivises individual DERs to submit truthful production estimates, as stated in the following lemma:

Lemma 3. *The payment rule π defined in* Eq. (5) *is strictly proper.*

Proof. The proof is done by contradiction with Lemma 2. Assume $\exists \boldsymbol{E}_i, \boldsymbol{E}_i \neq \boldsymbol{E}_i^*$ s.t. $\pi_t^{G,i}(\boldsymbol{E}_i) > \pi_t^{G,i}(\boldsymbol{E}_i^*)$. This means: $\exists \boldsymbol{E}_i$, s.t. $\mathrm{prod}_{i,t} \cdot \pi_B \cdot \mathscr{S}_{i,t}(\boldsymbol{E}_i, e_{i,t}) > \mathrm{prod}_{i,t} \cdot \pi_B \cdot \mathscr{S}_{i,t}(\boldsymbol{E}_i^*, e_{i,t})$ However, the unit price π_B is a positive constant and $\mathrm{prod}_{i,t}$ is the observed production (not dependent on the report $\boldsymbol{E}_i$) and also positive. Hence, this means: $\exists \boldsymbol{E}_i$, s.t. $\mathscr{S}_{i,t}(\boldsymbol{E}_i, e_{i,t}) > \mathscr{S}_{i,t}(\boldsymbol{E}_i^*, e_{i,t})$, which is a direct contradiction with Lemma 2. $\square$

This proof can also be stated that the payment function π is an affine transformation of the score $\mathscr{S}_{i,t}$, and hence boasts of the same strict propriety properties (c.f. [15]). As shown in the next section, for the redistribution function inside a CVPP this no longer holds and an additional step is needed.

The key intuition of having a reward based on a strictly proper scoring rule is that the reporting DER not only has an incentive to be as accurate as possible (thus report a prediction $\widetilde{\mathrm{prod}}_{C,t}$ as close to $\mathrm{prod}_{C,t}$ as possible), but also to be truthful about the entire past prediction distribution $\boldsymbol{E}_i$. Although, in this way, an imprecise agent that declares less than full confidence (i.e. through the distribution of its prediction error) will not receive the maximum score, and hence reward, even if its current prediction turns out to be completely accurate, on the other hand, if the prediction turns out to be poor, it will get penalised *much less*, because it declared this possibility through the distribution vector of prediction errors.

3.5. Payment mechanism within CVPP

If a set of DERs decide to join together in a virtual power plant C, this CVPP will first aggregate all their reports, past data, and production estimates, as discussed in Section 2; and will then get rewarded by the Grid with payment $\pi_t^{G,C}$ for each period t. This payment is then distributed by the CVPP to each member $i \in I$ (where I is the set of members) as:

$$\pi_t^{C,i} = \frac{\mathrm{prod}_{i,t} \times \mathscr{S}_{i,t}}{\mathrm{prod}_{i,t} + \sum_{j \neq i} \mathrm{prod}_{j,t} \times \mathscr{S}_{j,t}} \, \pi_t^{G,C} \qquad (6)$$

Eq. (6) ensures each member is paid a weighted fraction of the total payment received by the CVPP in period t, with a weight proportional to both its actual energy output $\mathrm{prod}_{i,t}$ and its individual accuracy score. The normalisation is such as to allow us to have the following property:

Lemma 4. *Given a set of reports from other agents $\boldsymbol{E}_j, j \neq i$, the redistribution payment $\pi_t^{C,i}$ in Eq. (6) is maximised when agent i is truthful.*

Proof. For the proof, we first observe that the factor in Eq. (6) that depends on the report $\boldsymbol{E}_i$ of agent i is the score $\mathscr{S}_{i,t}$ (the measured output $\mathrm{prod}_{i,t}$ is not a report). Moreover, note that the *maximum value* that the accuracy score can take (for a perfect prediction) is $\mathscr{S}_{i,t} = 1$ (this can be seen from Section 3.3, as $\mathscr{S}_{i,t}$ takes a value of 1 when $CRPS(\boldsymbol{E}_i, e_{i,t})$ takes the maximum value of 0). Now, a natural choice for the denominator in Eq. (6) would be to use $\sum_{\forall j \in I} \mathrm{prod}_{j,t} \times \mathscr{S}_{j,t}$ as a normalisation factor. However, this would *not* be an affine transformation, because the factor $\mathscr{S}_{i,t}$ (depending on report $\boldsymbol{E}_i$) cannot appear both above and below the fraction. Essentially, what we do here is that in the denominator, for each agent i, we replace $\mathscr{S}_{i,t}$ with its theoretically maximum possible value of 1. This "sum of others plus max" mechanism ensures that $\mathscr{S}_{i,t}$ only appears once, above the line.

The rest of the proof then is, as before, by contradiction, as follows. Assume $\exists \boldsymbol{E}_i, \boldsymbol{E}_i \neq \boldsymbol{E}_i^*$ s.t. $\pi_t^{C,i}(\boldsymbol{E}_i) > \pi_t^{C,i}(\boldsymbol{E}_i^*)$. The equation of $\pi^{C,i}$ consists of two parts: the fraction and $\pi^{G,C}$. Let us focus on $\pi^{G,C}$ first. This factor partially depends on the report $\boldsymbol{E}_i$ of agent i. Given a set

of reports by the rest of the agents, the best strategy for agent i in order to maximize $\pi^{G,C}$ is to also try to be as accurate as possible with its own accuracy report $\boldsymbol{E}_i^*$. Let us assume, however, that i chooses *not* to maximize $\pi^{G,C}$, by reporting $\boldsymbol{E}_i \neq \boldsymbol{E}_i^*$. Then, the only way for i to benefit from this report is if the first $\pi_t^{C,i}$ factor in Eq. (6) is such that $\pi_t^{C,i}(\boldsymbol{E}_i) > \pi_t^{C,i}(\boldsymbol{E}_i^*)$. But $\mathscr{S}_{i,t}$ is the only factor that depends on $\boldsymbol{E}_i$ in this first part. Thus, this necessitates that $\exists \boldsymbol{E}_i$, s.t. $\mathscr{S}_{i,t}(\boldsymbol{E}_i, e_{i,t}) > \mathscr{S}_{i,t}(\boldsymbol{E}_i^*, e_{i,t})$, a contradiction with Lemma 2. $\square$

Note that while the total amount redistribution by the cooperative to the agents in Eq. (6) is guaranteed to be below the total payment $\pi_t^{G,C}$ received by the cooperative from the grid operator, not the entire amount can be redistributed. This is because in the denominator of each agent i: $\mathrm{prod}_{i,t} + \sum_{j \neq i} \mathrm{prod}_{j,t} \cdot \mathscr{S}_{j,t} \geq \sum_{\forall i} \mathrm{prod}_{j,t} \cdot \mathscr{S}_{j,t}$. Hence, there will be a small amount that cannot be distributed, that will be kept by the mechanism design (e.g. by the energy cooperative towards its running costs) - however, in practice for a large number of agents, the non-redistributed amount will be very small (as we show experimentally in Section 4.5).

4. Experimental analysis

We study the performance of the proposed pricing functions in a real-life, renewable electricity generation scenario. Specifically, we consider the setting of Ecotricity, one of the largest renewable generation and distribution companies in the UK[3]. Ecotricity owns 16 wind farms distributed across the UK, with installed nominal capacities ranging from 0.5 MW to 16 MW. These farms differ not only in their nominal capacities, but also by the amount of wind they receive at their geographical locations and, crucially, their ability to use good wind speed predictions in those areas.

The overall question we consider in these experiments is: If these farms were independent producers working with the Grid, would the pricing functions we propose incentivise them to cooperate by forming a CVPP? Moreover, we aim to compare the incentives offered by the proposed scoring rules-based payment function (in which DERs report a full probability distribution) against a benchmark payment function in which the agents report just a single-point estimate of their future production.

4.1. Real-world data collection

Both the actual and predicted electricity generation for each wind farm, for each half hourly settlement period, depends primarily on the wind speeds. For our experiments, we collected half-hourly wind speed data for a 10-week period from 15 February to 30 April 2011. The data was collected from the website uk.weather.com, which essentially records the latest predictions made available by the UK Met Office. Both the actual and predicted wind data for each half hour were collected using the geographical locations of the 16 wind farms of Ecotricity. For each data point, we consider different prediction horizons, ranging from 1 to 24 h.

Given the predicted and actual wind speeds for any given time, the predicted and actual energy produced depends on the so-called *power curve* of each turbine. Power curves follow a sigmoid shape function [17]. At low wind speeds, the power generated is low, then it increases rapidly as wind speed increases and it levels off for high wind speeds. Note that wind turbines also have a safe operating limit for the wind speed they can use, above which the turbine temporarily shuts down to protect itself from damage. However, such high speeds were not recorded in the data set we used, so this does not influence our results. Formally, the energy generated by producer i at period t is:

$$prod_{i,t}\left(w_t^{HH}\right) = \frac{NomCapacity_i}{1 + e^{\alpha*(\beta - w_t^{HH})}} \qquad (7)$$

where $NomCapacity_i$ is the nominal capacity of farm i, w_t^{HH} represents the wind at the hub height at time t and e is Euler's number. The nominal (or installed) capacity is the maximum energy that a wind turbine can produce, under ideal wind conditions. In our case, each of the 16 Ecotricity farms has a different nominal capacity, ranging from 0.5 MW to 16 MW. The hub height w_t^{HH} is a parameter of the wind turbines and, likewise, it differs for each farm (as larger turbines have higher hubs). The method of computing the wind speed at the hub height, given the data (and predictions) of the Met Office is through the standard industry formula [18]:

$$w_t^{HH} = w_t * \frac{10}{36}\left(\frac{height_{hub}}{10}\right)^{0.2}$$

Here, in the original wind data (i.e. the w_t parameter) is expressed in km/h, $height_{hub}$ is taken from the data available for each Ecotricity location, while 0.2 is the so-called shear exponent used in vertical wind potential calculations [18]. In order to set the power curve parameters α and β, we use a technical report from Enercon (the main producer of the wind turbines used by Ecotricity farms) [17], based on which we set the values $\alpha = 0.625$ and $\beta = 9.7$. The above power curve function was used for generating both the *actual* and *predicted* energy production values for each of the 16 wind farms and each of the 70 days × 48 periods = 3360 half hourly periods. For each period, the predicted output was computed and logged for 24 prediction horizons: 1—24 h in advance.[4]

In terms of summary statistics, the installed (nominal) capacity of the 16 wind farms ranged from a minimum of 0.5 MW installed power output to a maximum of 16 MW, with newer wind farms being able to provide considerable more power. The average hub height ranged from 42 m to 85 m, again due to newer wind turbine types operating at considerably higher hub heights. Finally, in terms of prediction accuracy, the standard deviation σ of the estimates over the 3360 datapoints for a prediction time horizon of 4 h (used in the first set of our experiments) ranged from a minimum of $\sigma = 0.3271$ for the best predictor, to a maximum of $\sigma = 2.45$ for the worst predicting plant. While wind forecasting is known to be a very difficult problem, there is clearly quite a lot of variance in how well Ecotricity plants are able to predict generation, both due to the equipment they have installed, but also proximity to weather stations.

For the fixed price parameter π_B in Eq. (5) we assign $\pi_B = 0.8$. This value enables a realistic comparison, since the average amount paid per kWh in this setting matches the range of the feed-in tariffs currently being offered for renewable wind generation by the UK government (see www.fitariffs.co.uk/eligible/levels/). But, unlike our payment functions, feed-in tariffs *do not* reward prediction accuracy, nor do they incentivise formation of CVPPs.

Fig. 3 shows the cumulative distributions (CDFs) of the errors E_i of the 16 DERs over the 3360 half hourly periods, if the prediction occurs 4 h in advance. In this representation, the CDF of a perfect predictor would be a step function (i.e. the CDF for any value <0 is 0, and for any value >0 is 1); while, had errors been following a normal distribution, it would have been a sigmoid function.

As can be seen in Fig. 3, however, the real distributions of the 16

farms considered here do not follow perfect normal distributions (in fact, they are often non-symmetric), with CDFs ranging from those closer to a step function, to one corresponding to a more flat, uniform distribution. Our method can handle these variations, as the CRPS score in Eqs (2) and (3) is computed empirically by integration, and is proportional to the area between this curve and a vertical line corresponding to the actual observation.

4.2. Experimental setup

For our experimental analysis, we compare 4 different generation scenarios. They are as follows.

(1) All the 16 sites (or DERs) interact with the Grid as single, independent producers (i.e. as singletons) and are asked to provide only a single-point production estimate.
(2) All the 16 DERs interact with the Grid as singletons, but provide the Grid with both a production estimate and the expected distribution of their relative prediction errors.
(3) The 16 DERs interact with the Grid grouped together in a CVPP, and are only asked to jointly provide one CVPP-wide single-point production estimate.
(4) The 16 DERs interact with the Grid grouped together in a CVPP as above, but jointly provide the Grid with both a CVPP-wide production estimate, and the expected distribution of their global prediction error. In the two cases (i.e., scenarios 2 and 4) when the DERs, and respectively the CVPP, provide both expected production and distribution of relative prediction error (which is given by their vector of historical prediction errors up to that point), they will be paid according to the function in Eq. (5). In the other two cases (i.e., scenarios 1 and 3) when the DERs (and respectively the CVPP formed by them) only provide single point estimates, they will be paid according to the pricing function proposed in Ref. [1], and which we use as the benchmark for comparison:

$$\pi_t^{G,C} = \frac{1}{1 + \alpha\left|\widetilde{prod}_{C,t} - prod_{C,t}\right|^\beta} \cdot \log\left(prod_{C,t}\right) \cdot \pi_B \cdot prod_{C,t} \qquad (8)$$

Here $prod_C$ and π_B represent the actual production and the base price per kWh respectively, and have the same meaning as in our Eq. (5). However, in the point-based estimate payment function from Eq. (8), agents report *only* a single point estimate $\widetilde{prod}_C$, and not the entire distribution E_i of the relative error they expect to make.

To ensure a fair comparison between the two methods, the α and β parameters of the payment function in Eq. (8) are scaled such that, when the DERs participate in the market as singleton producers, they receive roughly the same payment with both the payment functions (i.e. the ones in Eqs. (6) and (8)). In this way, we have an unbiased benchmark for comparing the effects of these functions towards incentivising the formation of a CVPP.

4.3. Results for a single prediction horizon

First, we considered a setting in which all the DERs are asked to predict their productions 4 h in advance. This prediction horizon is often used in energy markets for short-term wind energy predictions [19], and provides a good benchmark value for our model. Results for this setting are shown in Fig. 4. The 16 DERs are ordered from poor predictors (high standard deviation, i.e., σ) to good predictors (low σ). The prediction error E_i of each farm or DER was computed using all the data from the 3360 half hourly intervals in our 10 week dataset. Results were tested for statistical significance

---

[4] Note that, in real-life, there may be other factors causing a variation in the actual power being produced, besides those captured in Eq. (7)—e.g., losses in transformers and transmission lines, from frequency matching, etc. However, these can be expected to be insignificant, and thus do not alter our conclusions.

using a paired *t*-test and were found significant at the p = 0.05 level.

Now, looking at the results in Fig. 4, two main trends can be observed. First, it is seen that, when DERs are interacting with the Grid as singleton producers and their estimates are reasonably accurate, they receive very similar payments under both payment mechanisms. This is expected because of our choice of parameters for the point estimates-based payment function in Eq. (8). The second observation is that, for both types of payment functions, forming a CVPP is clearly beneficial for all the agents.

The key reason why cooperative action is beneficial is the uncertainty reduction effect of adding together a number of uncertainty distributions from different DERs. When a number of DERs report a single aggregate estimate of joint generation, rather than individual ones, their under- and over-estimate prediction errors often partially compensate each other. This results in a much more accurate joint estimate − and hence in a larger payment than the sum of payments they would have received as singletons. In Figs. 4 and 5, this is reported as a payment per unit of energy generated, in order to level out the fact that wind farms in our setting have different sizes (as discussed before, installed nominal capacities of DERs range from 0.5 MW to 16 MW). Here, the revenue per unit of generation provides an accurate metric for the benefits of grouping DERs into a CVPP to improve accuracy. Moreover, note that while good predictors (those at the right hand side of the axis in Fig. 4) received comparably more benefit of being in the CVPP due to our distribution function, poor predictors also receive more revenue by participating in the CVPP, rather than accessing the market as singletons. Essentially, collective action is incentivised from *all* agents, as the penalty from poor prediction is compensated by the improved joint accuracy from being in the CVPP.

Finally, we note the incentive to form a CVPP is considerably stronger with CRPS payments (where agents report both a predicted mean and an error), than with single-point estimate payments. Intuitively, the reason for this is that, with a scoring rule-based payment function, the CVPP is "punished" less for poor predictions when it also reports the confidence in these predictions (in the form of the expected distribution of the relative prediction errors) than when just reporting a single-point estimate. Thus, our
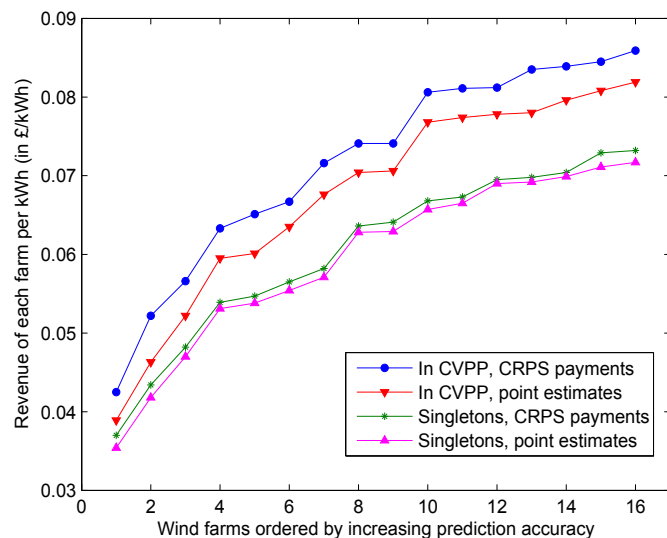
scoring rules payment mechanism results in both gathering more useful information for the Grid (in the form of expected error distributions), and in providing stronger incentives for individual DERs to group together into cooperatives.

4.4. Results for different prediction horizons

Next, we also investigate whether the incentives for forming CVPPs, and the advantages of our scoring rules-based payments mechanism over the point estimate payments hold true over different prediction horizons. To this end, the number of hours in advance that the agents are asked to predict their productions is varied from 1 to 24 h. Moreover, in contrast to the previous section, we look at the aggregate revenue, averaged over all the DERs, rather than individual DER revenues; this allows us to summarize each setting of the prediction-time horizon in a single value. The results and standard error over the different time points are shown in Fig. 5. As in Section 4.3, for the scenarios when DERs interact with the Grid as singletons, they receive a similar per-unit payment (in £/kWh) under the scoring rule-based and point estimates-based mechanisms. Results were found to be statistically significant using paired t-tests.
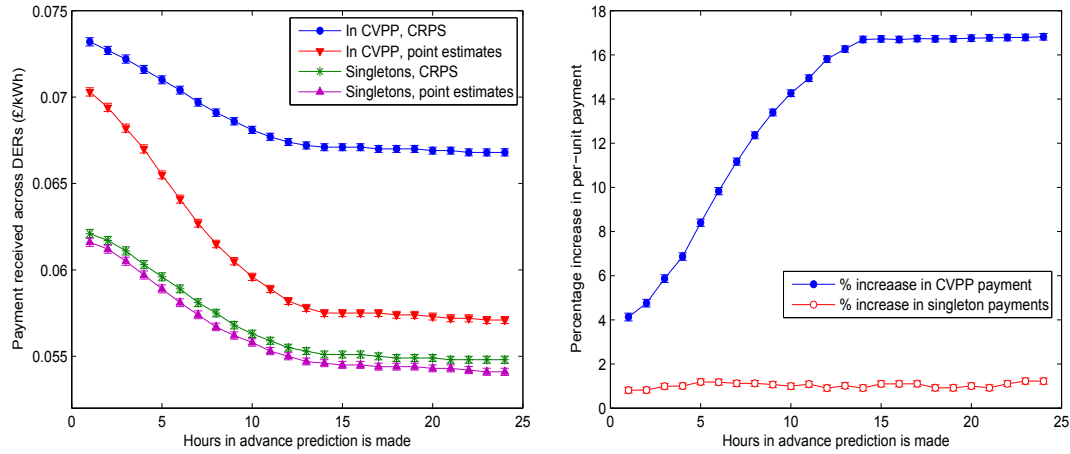
We note that the incentive for collective action (joining the CVPP) remains strong as in Fig. 5(a), the revenue per unit generated in the two scenarios where agents interact in the market through the CVPP clearly dominate the revenue in the singleton settings. Interestingly, the effect becomes stronger as the forecasting horizon (i.e. the number of hours in advance the prediction is requested) increases. This is because the variance of the prediction for all agents (both good and poor agents) decreases with the longer time horizon, hence they get more benefit by forming a CVPP. Moreover, this evaluation also shows that, for all prediction horizons, our mechanism performs much better in incentivising producers to form CVPPs than point-based estimates payments. It is especially noteworthy, however, that the relative advantage offered by the scoring rule-based mechanisms over the single-estimate mechanisms actually increases considerably as agents are required to predict within a longer prediction horizon (i.e., more hours in advance). In Fig. 5(a), this advantage ranges from 7.3p/kWh vs. 7p/kWh for 1-h in advance predictions (a difference of approx. 4%), to £0.67 vs. £0.57 for 24 h in advance (a difference of approx. 17%). The underlying reason for this is that, as predictions are attempted further in advance they become considerably less accurate (i.e. they are made with a higher expected relative prediction error). The scoring rule-based mechanism allows the CVPP to truthfully report this expected loss of accuracy (by reporting a worse distribution of expected relative errors), and hence get "punished" less.

Fig. 5(b) plots the same effect in percentage (relative) terms of the increase in per-unit payment. Note that singleton agents do not benefit so much from the CRPS-based scheme (though there is a slight positive difference of about 1%), because their accuracy is too low to benefit anyway. However, grouping in a CVPP allows them to report their increasing joint uncertainty, and thus gain up to 17% more per-unit payment under the CRPS-based scheme compared to the single-point estimate scheme.

4.5. Non-redistributed payments

Finally, one aspect we studied in our experiments is how large are the payments collected by the CVPP that remain undistributed from the CVPP to individual members, due to our redistribution mechanism. Recall that Eq. (6) assures that redistribution payment function guarantees truthful reports of expected generation, but a small part of the payment may remain undistributed by the CVPP to individual DERs. In particular, Fig. 6 shows the size of this payment

Fig. 4. Average payment made to each individual farm (DER) for the 10 week period (in £/kWh), for the setting where all DERs are asked to predict their production 4 h in advance. DERs are ordered ascendingly, in the order of their prediction accuracy (from the poorest to the best predictor).

Fig. 5. (a) Average revenue across all DERs (in £/kWh), for different prediction horizons ranging from 1 to 24 h. (b) Percentage increase in average (i.e. per kWh) revenue of the scoring rules-based scheme vs. the single-estimate scheme.

in practice: it varies from 1.25% for predictions required an hour in advance to 1.55% to longer term predictions, when agents are less accurate. However, the overall conclusion is that the amount of payment collected that is kept by the CVPP (and not redistributed) is very small. Otherwise stated, the "cost" of assuring truthful report s through Eq. (6) is very small, less than 2% (note that this 1.5—2% is not an actual loss, just money that accrues to the CVPP itself, rather than the members).

5. Conclusions and further work

In this paper, we study the problem of eliciting truthful and accurate estimates from multiple distributed energy resources (DERs) and, at the same time, incentivise them to group to improve their accuracy, and sell energy together as a cooperative virtual power plant (CVPP). In order to do this, we propose a payment mechanism (an alternative to the current feed-in tariffs), in which producers are rewarded not only based on their production, but also on the prediction of future output and the uncertainty of their estimates. In order to elicit this information truthfully from potentially self-interested agents, our proposed method uses a state-of the art technique in mechanism design called strictly

proper scoring rules. Moreover, the method we develop is generic, in the sense that it does not make any assumptions on the distribution of the error vector (such as errors following a normal distribution).

In order to validate our method, we use a large-scale dataset from the 16 farms of Ecotricity, a large UK renewable producer, with outputs and predictions collected over a 10-week period. We show that, first, the proposed payment scheme is successful in incentivising agents to group and sell electricity as a CVPP. Second, the scoring rule based method outperforms a benchmark that only uses a single-point prediction, by a factor ranging from 4% to 17% in per-unit revenue terms, depending on the prediction horizon required.

In future work, we plan to model CVPPs formed by a combination of renewables, such as wind, solar and tidal energy, in areas where several of these are available (this study focused on the UK, where wind energy is the main renewable generation resource). Furthermore, we would like to experimentally measure the financial and technological benefits to the Grid due to the better scheduling given the increased CVPP reliability. Finally, we intend to enlarge our settings and apply similar mechanisms to enable the formation of cooperatives for demand response and demand-side management—that is, to investigate ways in which cooperatives of electricity consumers could form to assist the effort of achieving demand reduction, through demand response techniques. Initial work on applying cooperatives for collective action on the demand side was started by Kota et al. [20] and by Shafie et al. [21]; however these works do not consider the use of scoring rules to incentivise accurate forecasts. One work that applies a similar technique (based on scoring rules), to the problem of demand consumption shifting, rather than generation, is Akasiadis and Chalkiadakis [22], whose approach was partially inspired by the conference version of our work.

Acknowledgements

References

[1] Chalkiadakis G, Robu V, Kota R, Rogers A, Jennings NR. Cooperatives of distributed energy resources for efficient virtual power plants. In: Proc. Of the 10th int. Conference on autonomous agents and multiagent systems (AAMAS 2011); 2011. p. 787—94.
[2] Robu V, Kota R, Chalkiadakis G, Rogers A, Jennings NR. Cooperative virtual
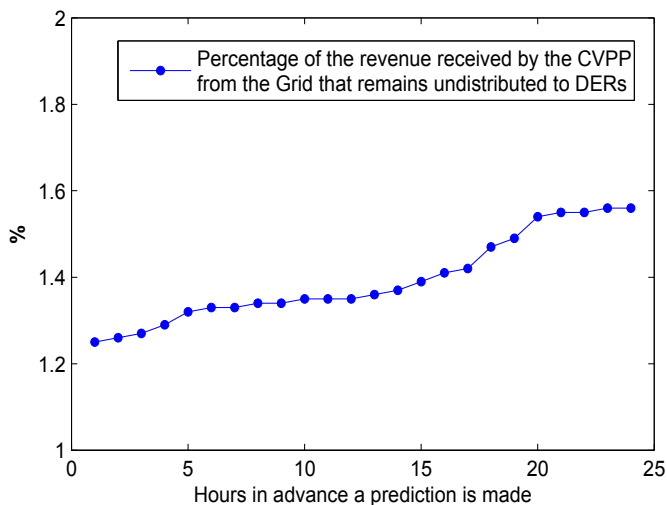
Fig. 6. Percentage of the total payment received by the CVPP from the Grid that is not distributed to member DERs, for the different prediction scenarios.

power plant formation using scoring rules. In: Twenty-sixth int. Conference on artificial intelligence (AAAI'12); 2012. p. 1—6.

[3] Ramchurn S, Vytelingum P, Rogers A, Jennings N. Putting the "smarts" into the smart grid: a grand challenge for artificial intelligence. Commun ACM 2012;55(4):86—97.

[4] Kok K, Scheepers M, Kamphuis R. Intelligence in electricity networks for embedding renewables and distributed generation. In: Intelligent infrastructures, vol. 42; 2010. p. 179—209.

[5] Pudjianto D, Ramsay C, Strbac G. Virtual power plant and system integration of distributed energy resources. IET Renew Power Gener 2007;1(1):10—6.

[6] Robu V, Gerding EH, Stein S, Parkes DC, Rogers A, Jennings NR. An online mechanism for multi-unit demand and its application to plug-in hybrid electric vehicle charging. J Artif Intell Res 2013;48:175—230.

[7] Ma H, Robu V, Li N, Parkes DC. Incentivizing reliability in demand-side response. In: Proc. Of 25th int. Joint conf. On artificial intelligence, IJCAI; 2016. p. 352—8.

[8] Asmus P. Microgrids, virtual power plants and our distributed energy future. Electr J 2010;23:72—82.

[9] Tarroja B, Mueller F, Eichman J, Samuelsen S. Metrics for evaluating the impacts of intermittent renewable generation on utility load-balancing. Energy 2012;42(1):546—62.

[10] Sijm JPM. The performance of feed-in tariffs to promote renewable electricity in european countries., the Energy Centre of The Netherlands. 2002. Report ECN-C-02—083.

[11] Dimeas A, Hatziargyriou N. Agent based control of virtual power plants. In: 14th int. Conf. Int. Appl. To power systems; 2007. p. 1—6.

[12] Vasirani M, Kota R, Cavalcante R, Ossowski S, Jennings N. An agent-based approach to virtual power plants of wind generators and electric vehicles. IEEE Trans Smart Grid 2013;4(3):1314—22.

[13] Savage LJ. Elicitation of personal probabilities and expectations. J Am Stat Assoc 1971;66(336):783—801.

[14] Matheson JE, Winkler RL. Scoring rules for continuous probability distributions. Manag Sci 1976;22:1087—95.

[15] Gneiting T, Raftery A. Strictly proper scoring rules, prediction and estimation. J Am Stat Assoc 2007;102(477):359—78.

[16] Hersbach H. Decomposition of the continuous ranked probability score for ensemble prediction systems. Weather Forecast 2000;15(5):559—70.

[17] Enercon. Enercon wind energy converters - product overview (e70). 2010.

[18] Peterson E, Hennessey J. On the use of power laws for estimates of wind power potential. Appl Meteorol 1978;17:390—4.

[19] Giebel G, Brownsword R, Kariniotakis G. The state-of-the-art in short-term prediction of wind power. a literature overview. In: Project ANEMOS D1, vol. 1; 2003. p. 1—36.

[20] Kota R, Chalkiadakis G, Robu V, Rogers A, Jennings NR. Cooperatives for demand side management. In: 7th conf. On prestigious appl. Of int. Syst. (PAIS @ ECAI); 2012. p. 969—74.

[21] Shafie-khah M, Moghaddam M, El-Eslami M, Andebili M. Modeling of interactions between regulations and behavior of plug-in electric vehicle aggregators in a virtual power market. Energy 2012;40(1):139—50.

[22] Akasiadis C, Chalkiadakis G. Agent cooperatives for effective power consumption shifting. In: In proc. Of the 27th AAAI conference on artificial intelligence, (AAAI-2013); 2013. p. 1263—9.