Cooperative Virtual Power Plant Formation Using Scoring Rules

Valentin Robu
Electronics and Computer Science
University of Southampton, UK
vr2@ecs.soton.ac.uk

Ramachandra Kota
Secure Meters Ltd.,
Winchester, UK
rck05r@ecs.soton.ac.uk

Georgios Chalkiadakis
Electronic and Computer Engineering,
Technical University of Crete, Greece
gehalk@intelligence.tuc.gr

Alex Rogers
Electronics and Computer Science,
University of Southampton, UK
acr@ecs.soton.ac.uk

Nicholas R. Jennings
Electronics and Computer Science,
University of Southampton, UK
nrj@ecs.soton.ac.uk

Abstract

Virtual Power Plants (VPPs) are fast emerging as a suitable 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 such 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 multi-agent technology. In particular, we design a payment mechanism that encourages DERs to join CVPPs with large overall production. Our method is based on strictly proper scoring rules and incentivises the provision of accurate predictions from the CVPPs—and in turn, the 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. We show that our mechanism incentivises real DERs to form CVPPs, and outperforms the current state of the art payment mechanism developed for this problem.

1 Introduction

In recent years, a number of strands in intelligent and multi-agent systems 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 (Ramchurn et al. (2012), Dimeas et al. (2007), Kok et al. (2009), Gerding et al. (2011)). While environmental concerns are becoming increasingly important, the overriding concern of national electricity transmission network operators (termed *the Grid* herein) 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 needed.

Although reliability is easily addressed when energy is produced solely by conventional power plants (which can usually adjust their power output 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 (Pudjianto, Ramsay, and Strbac 2007), 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 visible, the uncertainty and uncontrollability of renewable energy sources inhibits individual DERs from profitably dealing with the Grid directly, or participating in the wholesale electricity market because they are often unable to meet the set generation targets. Nevertheless, the need to incorporate renewable energy resources in the existing Grid is a pressing one.

The remedy adopted by many countries is to encourage small-scale renewable energy producers with payments according to specific *feed-in tariffs*, 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.

If individual DERs could be aggregated together to form larger energy generating entities, such entities would 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 (Dimeas et al. (2007),Kok et al. (2009),Pudjianto et al. (2007)). Recently, Chalkiadakis *et al.* (2011) proposed a pricing mechanism that can be used by the Grid to promote the creation of *cooperatives* of DERs, and constitutes an alternative to feed-in tariffs. In that work, the term *cooperative VPP (CVPP)* reflects the fact that economically rational agents, representing the individual DERs, are incentivised to work together in their mutual interest. However, a critical limitation of that approach 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 production estimates are provided in the form of probability distributions, 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. Naturally, the more accurate the provided estimates, and the higher the confidence placed in those estimates, the better for the Grid scheduling activities. In contrast, if information is provided only in the form of single point estimates, and these estimates prove to be erroneous, the Grid would be forced to either dispatch conventional generators at short notice, or purchase the required energy in the balancing market at the last minute—both of which come at a high cost (Kirschen and Strbac 2004).

Scoring rules with specific properties, 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 (Savage 1971; Gneiting and Raftery 2007; Papakonstantinou et al. 2011). More specifically, scoring rules that are *strictly proper* can be employed by a mechanism designer to ascertain that agents accurately declare their privately calculated distributions, reflecting 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 *incentive compatibility* of their estimates. 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 possible.

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. 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 important contribution of this paper lies in the experimental analysis of the proposed mechanism. We base our experimental setting on 16 real-world wind farms, that are distributed around the UK. For these farms, we collect a 3-month 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 Cooperatives and the Grid

We consider a setting where several independent, distributed energy producers (DERs) that can sell their energy directly to the Grid or opt to join an agent cooperative (CVPP). The main CVPP function is to represent its DER members in interacting with the Grid; the CVPP can provide the aggregate estimate of the members' production, receive 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 (in most countries, 48 half-hour slots).

Formally, for any time 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 way a DER obtains this estimate (and its accuracy) depends on the type of generation capacity it has at its disposal. For instance, for DERs composed of (one or more) wind turbines - such as the ones considered in this paper's analysis - estimate their production $\widetilde{prod}_{i,t}$ based on an hourly wind prediction obtained for their area from the UK meteorological office (as described in detail in Section 4).

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

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

Note that in Equation 1 the normalisation (given by the denominator) is done with respect to $\widetilde{prod}_{i,t}$. This is because, from the perspective of the DER, it is the actual production $prod_{i,t}$ which 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 are much more predictable than those that use wind. Even among wind-based DERs, there may be substantial differences in prediction, because wind in some areas may be easier to predict than in others (or, simply, it may be that the meteorological office provides more accurate and timely predictions for some areas). However, if each DER takes care to account for these errors, over a large enough period of time, the long-term *average* error will be around 0.

Formally, using a typical statistical model of random errors, for any DER i we capture its *uncertainty over the errors* it expects to make in its prediction through a normal distribution $\mathcal{N}(\mu = 0, \sigma_i^2)$.[1] The standard deviation of this distribution σ_i then reflects how confident a DER is in the predictions it makes. Note these are standard deviations of the *relative* errors (as defined in Equation 1), the absolute standard deviation (in Kwh) at time t being $\sigma_i * \widetilde{prod}_{i,t}$.

Each DER determines its confidence in its prediction σ_i^2 at different times, using its own private information. For example, it can collect and use historical data of its past prediction errors $e_{i,t}$, and compute the statistical variance over these (for instance in case of wind turbines, this data is based on wind speeds and their predictions). But, more generally, it can also take into account other factors that are private information to the DER, and not known to the Grid or other parties. The challenge we address is designing a mechanism that elicits the uncertainty over these estimates truthfully.

Following the notation for individual DER production, the actual and estimated production of a cooperative C at t are denoted by $prod_{C,t}$ and $\widetilde{prod}_{C,t}$ respectively. Now, if we denote by I the set of the participating members of the CVPP, the total CVPP production at t can be computed as the sum total of its members' production—i.e., $prod_{C,t} = \sum_{i\in I} prod_{i,t}$. Further, the estimate of the CVPPs production at time t is the sum of the estimates of the individual DERs participating at time t, i.e. $\widetilde{prod}_{C,t} = \sum_{i\in I} \widetilde{prod}_{i,t}$.

The joint *relative* error in prediction of the CVPP is then:

$$\sigma_{C,t}^2 = \frac{\sum_{i\in I}\left(\widetilde{prod}_{i,t} * \sigma_i\right)^2}{\left(\sum_{i\in I}\widetilde{prod}_{i,t}\right)^2} \tag{2}$$

Recall that $\sigma_{C,t}$ values are defined for the *relative* (or proportional) errors. So, Eq. 2 is obtained from the standard statistical formula relating the absolute variances of a sum of Gaussian distributions: $(\widetilde{prod}_{C,t}\sigma_{C,t})^2 = \sum_{i\in I}(\widetilde{prod}_{i,t}\sigma_{i,t})^2$, using $\widetilde{prod}_{C,t} = \sum_{i\in I}\widetilde{prod}_{i,t}$.

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 $S(\hat{P}, x)$, specifying the reward that a forecaster agent i should receive if it reports a predicted distribution $\hat{P}$ 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. A scoring rule $S(\hat{P}, x)$ is *strictly proper* if it is such that agent i has the incentive to declare *only* his true belief P, as this is the only prediction that maximises its expected reward. Formally, if P is the true underlying distribution of the random variable x, scoring rule S is strictly proper if $S(P,x) \geq S(\hat{P}, x)$, with the equality holding if and only if $\hat{P} = P$. The scoring rule is said to be proper if $S(P,x) \geq S(\hat{P}, x)$, but the the prediction $\hat{P} = P$ is not the only one that maximises $S(\hat{P}, x)$. In our case, the use of a strictly proper scoring rule would mean 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.

[1] This does not mean, of course, that we assume the actual power outputs at different times are normally distributed, just that random relative errors, over a long enough time range, will be normally distributed around a mean of 0.

3.1 Continuous Ranked Probability Score

The traditional forms of proper and strictly proper scoring rules that appear in the literature (Savage 1971) do not satisfy our current requirements because most were not designed to work for continuous variables (like the Gaussian distribution in our case). Also, some of them 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 characteristics are possessed by the *Continuous Ranked Probability Score* (CRPS) (Matheson and Winkler 1976), which is a strictly proper scoring rule used for continuous variables. It has lately attracted renewed interest in the scoring rules literature, and has been used extensively to help qualify weather predictions (Hersbach 2000). This is the scoring rule that forms the basis of our payment mechanism. Since, as we showed in Section 2, the relative errors in predictions made by a DER i over the long term can be approximated by a Gaussian distribution centered at 0, we can use the CRPS form put forward by Gneiting and Raftery (2007) as follows. Consider a DER i (where i can also be the CVPP C itself), which reports an uncertainty over its relative prediction error of $\mathcal{N}(0, \sigma_i^2)$. Let the actual relative error observed at time t be denoted by $e_{i,t}$ (as defined in Eq. 1). Then, the CRPS score obtained by DER i at time t is:

$$\begin{aligned}&\mathrm{CRPS}(\mathcal{N}(\mu = 0, \sigma_i^2), e_{i,t}) = \\ &= \sigma_i\left[\frac{1}{\sqrt{\pi}} - 2\varphi\left(\frac{e_{i,t}}{\sigma_i}\right) - \frac{e_{i,t}}{\sigma_i}\left(2\Phi\left(\frac{e_{i,t}}{\sigma_i}\right) - 1\right)\right]\end{aligned} \tag{3}$$

where φ and Φ denote the probability density and the cumulative distribution function of a standard Gaussian variable, respectively. Since $\mu = 0$, the only report affecting the CRPS value is σ_i^2 and $e_{i,t}$. Therefore, we can simplify the notation of $\mathrm{CRPS}(\mathcal{N}(0, \sigma_i^2), e_{i,t})$ to $\mathrm{CRPS}(\sigma_i, e_{i,t})$. Interested readers can consult (Gneiting and Raftery 2007) for the proof of strict propriety and the intuition behind this function form.

3.2 Payment Mechanism from the Grid to 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, by any DER i that chooses to sell directly to the Grid). We then present the *"CVPP-to-DER"* pricing function, which is used by the CVPP to distribute the received payments internally among its members. First, we denote the electricity *base price* per kWh produced by π_B. This can be either under the direct control of the Grid, or determined directly by the electricity market. Recall that, for each settlement period t, each producer C supplies to the Grid an estimate of the energy it is going to produce $\widetilde{prod}_{C,t}$, as well as the relative average prediction error σ_C, that encodes how (un)certain the producer is on the accuracy of its predictions. The Grid will also observe the actual amount of energy $prod_{C,t}$ which is produced by C in settlement period t. As discussed in Section 2, it then computes the actual relative prediction error made by producer C in period t as $e_{C,t} = \frac{prod_{C,t} - \widetilde{prod}_{C,t}}{\widetilde{prod}_{C,t}}$. The payment from the Grid G to the CVPP C for settlement period t is then given by the function:

$$V_t^{G,C} = \text{CRPS}(\sigma_C, e_{C,t}) \times \pi_B \times \log(\text{prod}_{C,t}) \times \text{prod}_{C,t} \quad (4)$$

The function is composed of four factors, multiplied together to determine the amount of payment received for a realised production $prod_C$. The *"accuracy factor"* (represented by the CRPS function, scaled between 0 and 1) is the part which incentivises the CVPP to provide as accurate description as possible for its relative prediction error. Note that this is the only part of the function that uses the reports σ_C and $\widetilde{prod}_{C,t}$ made by the agent (through the relative error $e_{C,t}$). The value $prod_{C,t}$ is the actual production of C in settlement period t, which is independently observed by the Grid, not a report. As shown above, the CRPS part of function is a strictly proper scoring rule w.r.t. σ_C and $e_{C,t}$, and the entire payment function in Eq. 4 is an affine transformation of this rule (since it only involves a multiplication with other factors which do not depend on the reports made by the agent). Hence, Eq. 4 is also strictly proper (c.f. (Gneiting and Raftery 2007)).

Fig. 1 illustrates this accuracy factor for different values of σ_i vs. actual proportional error $e_{i,t}$, i.e. the proportional difference between the actual and predicted productions of DER i. What is interesting to observe here, however, is 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 reporting a higher σ_i (i.e. less confidence) provides a better reward.

Factor $log(prod_C)$ is a *"production factor"*, which promotes the formation of large cooperatives as the Grid requires—combined with $prod_C$, it makes the payment to C super-linear. The "accuracy" and "production" factors, along with π_B, constitute the *actual price* paid by the Grid to C. The overall Grid-to-CVPP payment is then calculated by multiplying this with the realised production $\text{prod}_{C,t}$.

3.3 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 and productions based on the formulas presented in Section 2, and get rewarded by the Grid, for each period t with payment $V_t^{G,C}$. This is then distributed by the CVPP to each member $i \in I$ (where I is the set of members) as:

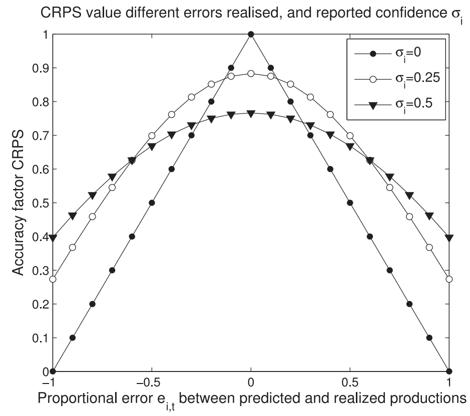

Figure 1: Accuracy factor function diagram

$$V_t^{C,i} = \frac{\text{CRPS}\left(\sigma_i, e_{i,t}\right) * prod_{i,t}}{\sum_{\forall j \in \text{I}} \text{CRPS}\left(\sigma_j, e_{j,t}\right) * prod_{j,t}} \cdot V_t^{G,C} \quad (5)$$

Eq. 5 ensures that each member is paid a weighted fraction of the total payment received by the CVPP in settlement period t, with a weight proportional to its contribution. Contributions are measured not just with respect to the actual energy outputs $prod_{i,t}$, but also to each DER's individual CRPS score, which reflects how beneficial their estimates were in terms of obtaining a better price via the CVPP wide accuracy factor in Eq. 4. Moreover, the normalization ensures the entire received payment is distributed.

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[2]. 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 study how the incentives provided by our scoring-rule based payment functions compare to a benchmark payment function which does not use probabilistic estimates.

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

[2] www.ecotricity.com

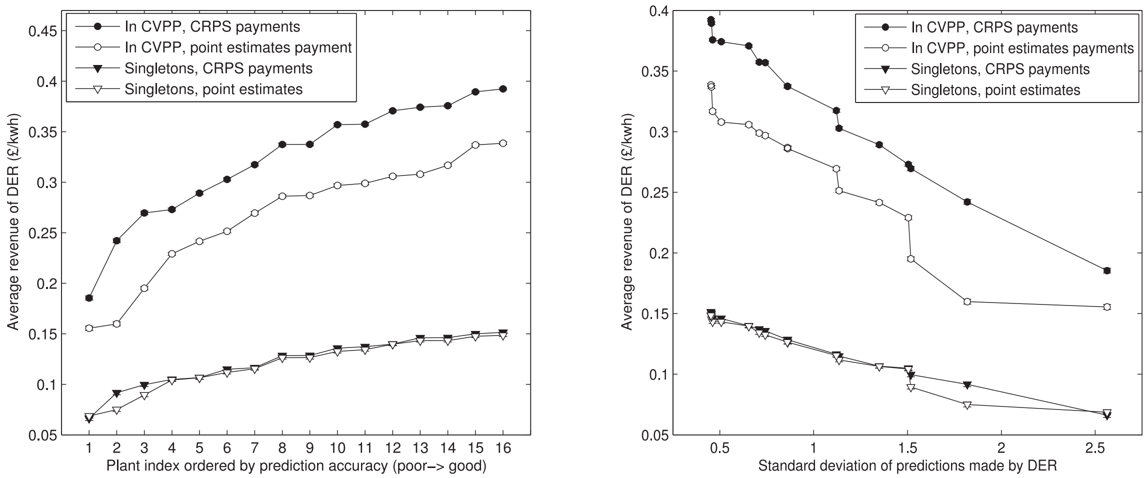

Figure 2: Average payment made to each individual farm (DER) for the entire 10 week period (in £ /kWh), for the setting where all DERs are asked to predict their production 4 hours in advance. The x-axis is ordered in two ways: (a) DERs are ordered ascendingly, in the order of their prediction accuracy (from the poorest to the best predictor), (b) DERs are ordered based on their standard prediction accuracy error σ. (hence, the order is from the best to the poorest predictor)

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 hours.

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. 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:

$$prod_{i,t}(w_t^{HH}) = \frac{NomCapacity_i}{1+e^{\alpha*(\beta-w_t^{HH})}} \quad (6)$$

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

We use a technical report from Enercon (the main producer of the wind turbines used by Ecotricity farms) (Enercon 2010) to determine the power curve values of $\alpha = 0.625$ 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 to 24 hours in advance.[3]

For the fixed price parameter π_B in Equation 4 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.

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 mean production estimate and an expected standard deviation for their prediction error.

(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

[3]Note that, in real-life, there may be other factors causing a variation in the actual power being produced besides the ones captured in Equation 6, such as losses in transformers and transmission lines, from frequency matching etc. However, these can be expected to be insignificant, and thus would not alter our conclusions.

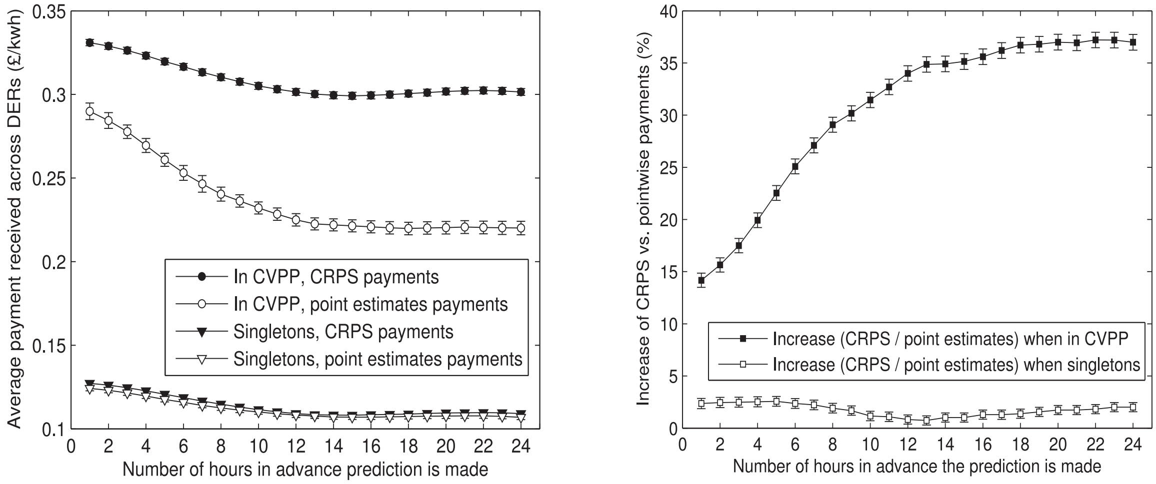

Figure 3: (a) Average revenue across all DERs (in £ /kWh), for different prediction horizons ranging from 1 to 24 hours. (b) Percentage increase in average (i.e. per kWh) revenue for scoring rules-based versus point-based estimates payment mechanism.

in a CVPP as above, but jointly provide the Grid with both a CVPP-wide mean production estimate, and an expected standard deviation for their prediction error.

In the two cases (i.e., scenarios 2 and 4) when the DERs, and respectively the CVPP, provide both expected production and standard deviation estimates, they will be paid according to the function in Eq. 4 introduced in this paper. 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 (Chalkiadakis et al. 2011) which we use as the benchmark for comparison:

$$V_t^{G,C} = \frac{1}{1+\alpha|\widetilde{\text{prod}}_{C,t} - \text{prod}_{C,t}|^{\beta}} \cdot \log(\text{prod}_{C,t}) \cdot \pi_B \cdot \text{prod}_{C,t} \tag{7}$$

Here $prod_C$ and π_B represent the actual production and the base price per kWh respectively, and have the same meaning as in Eq. 4 from earlier in this paper. However, in the point-based estimate payment function from Eq. 7, agents report *only* a single point estimate $\widetilde{prod}_C$, and not an uncertainty distribution $\mathcal{N}(0,\sigma_C^2)$ over the error it expects to make.

A crucial point for ensuring a fair comparison between the two methods is choosing the way to scale the α and β parameters of the payment function from Eq. 7. In order to have a fair benchmark, we set these parameters such that, when the DERs participate in the market as singleton producers, they receive the same payment with both the payment functions (i.e. the ones in Equations 5 and 7). In this way, we have an unbiased benchmark for comparing the effects of these functions towards incentivising CVPP formation.

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 hours in advance. This prediction horizon is often used in energy markets for short-term wind energy predictions (Giebel, Brownsword, and Kariniotakis 2003), and provides a good benchmark value for our model. Results for the 4-hours-in-advance prediction setting are shown in Figure 2. In Figure 2(a), the 16 DERs are ordered from poor predictors (high standard deviation, i.e., σ) to good predictors (low σ), while in Figure 2(b), the relationship between the SDs for different DERs and their revenues are explicitly plotted. For both cases, the prediction error σ_i of each farm or DER was computed using all the data from the 3360 half hourly intervals in our 10 week dataset.

Now, looking at the results in Figure 2, two main trends can be observed [4]. First, it is seen that, when DERs are interacting with the Grid as singleton producers and their estimates are reasonably accurate, they receive roughly the same payment from both the payment mechanisms. This is expected because of our choice of parameters for the point estimates-based payment function in Eq. 7. Note that even then, there are two poor predicting DERs, with σ of 1.5 and 1.8, respectively, for whom the revenues diverge slightly (Figure 2(b)). These two are actually better off with CRPS payments, i.e., where they report the mean and the expected standard deviations. Intuitively, the reason behind this is that with a scoring rule payment function, poor predictors are "punished" less as they also report their confidence than when they just report an inaccurate single point estimate.

The second interesting observation is that, for both types of payment functions, forming a CVPP is clearly beneficial for all the agents. However, 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. Thus, our scoring rules payment mechanism not only results in gathering more useful information for the Grid, it also provides stronger incentives for individual DERs to group together into cooperatives.

[4] Error bars in Figure 2 are not visible due to their small size.

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 hours. 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 Figure 3. As in Section 4.3, for the scenarios when DERs interact with the Grid as singletons, there is almost no difference between the revenues made (in £/kWh) for the scoring rule-based and point estimates-based mechanisms.

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. In this context, it is especially interesting to observe that the relative advantage offered by scoring rules actually *increases* as the time-horizon of prediction increases (that is, as DER agents are asked to predict much more in advance). This relative advantage between the two mechanisms may seem small in absolute terms from looking at Figure 3(a), but as Figure 3(b) shows, it actually increases from 14% to around 37%. Although space in the paper does not allow for additional graphs, we found that the underlying reason for this is that, as the CVPP predictions are attempted for a longer period in advance, they tend to become considerably less accurate. Allowing agents to report the uncertainty in their estimates allows them to avoid being punished, in settings with higher uncertainty.

5 Conclusions and Further Work

This paper develops a novel pricing mechanism to encourage the integration of renewable DERs in the existing electricity Grid, through the formation of Cooperative Virtual Power Plants (CVPPs). Our mechanism provides an alternative to unsustainable feed-in tariffs, and uses scoring rules to incentivise DERs to report not only accurate estimates of their production, but also the uncertainty in these estimates.

In future work, we plan to model CVPPs formed by a combination of renewables, such as wind, solar and tidal energy. Also, we plan to study how the widespread availability of storage would impact the formation of CVPPs. 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. In order to have meaningful results, this would have to involve the use of detailed simulation environments, built based on real-world business processes and data, but which were not readily available for this study.

Finally, we intend to enlarge our settings to consider the formation of cooperatives for demand response and demand-side management—that is, to investigate ways in which co-operatives of electricity *consumers* could form to assist the effort of achieving demand reduction.

Acknowledgements:

This work is supported by the iDEaS project (`www.ideasproject.info`).

References

Chalkiadakis, G.; Robu, V.; Kota, R.; Rogers, A.; and Jennings, N. R. 2011. 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)*, 787–794.

Dimeas, A., and Hatziargyriou, N. 2007. Agent based control of virtual power plants. In *14th Int. Conf. Intelligent System Applications to Power Systems*, 1–6.

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

Gerding, E. H.; Robu, V.; Stein, S.; Parkes, D.; Rogers, A.; and Jennings, N. R. 2011. Online mechanism design for electric vehicle charging. In *Proc. 10th Int. Conf. Aut. Agents and Multi-Agent Systems (AAMAS'11)*, 811–818.

Giebel, G.; Brownsword, R.; and Kariniotakis, G. 2003. The state-of-the-art in short-term prediction of wind power. a literature overview. In *Project ANEMOS D1.1*.

Gneiting, T., and Raftery, A. 2007. Strictly proper scoring rules, prediction and estimation. *Journal of the American Statistical Association* 102(477):359–378.

Hersbach, H. 2000. Decomposition of the continuous ranked probability score for ensemble prediction systems. *Weather and Forecasting* 15(5):559–570.

Kirschen, D. S., and Strbac, G. 2004. *Fundamentals of Power System Economics*. Wiley.

Kok, K.; Scheepers, M.; and Kamphuis, R. 2009. Intelligence in electricity networks for embedding renewables and distributed generation. In *Intelligent Infrastructures*.

Matheson, J. E., and Winkler, R. L. 1976. Scoring rules for continuous probability distributions. *Management Science* 22:1087–1095.

Papakonstantinou, A.; Rogers, A.; Gerding, E.; and Jennings, N. R. 2011. Mechanism design for the truthful elicitation of costly probabilistic estimates in distributed information systems. *Artificial Intelligence* 175(2):648–672.

Pudjianto, D.; Ramsay, C.; and Strbac, G. 2007. Virtual power plant and system integration of distributed energy resources. *IET Renewable Power Generation* 1(1):10–16.

Ramchurn, S.; Vytelingum, P.; Rogers, A.; and Jennings, N. 2012. Putting the "smarts" into the smart grid: A grand challenge for artificial intelligence. *Comm. of the ACM*.

Savage, L. J. 1971. Elicitation of personal probabilities and expectations. *Journal of the American Statistical Association* 66(336):783–801.