Analyzing and improving the energy balancing market in the Power Trading Agent Competition

Mohammad Ansarin & Wolfgang Ketter
Erasmus University Rotterdam
Rotterdam, Netherlands
{ansarin, wketter}@rsm.nl

John Collins
University of Minnesota, Minneapolis,
MN, United States
jcollins@cs.umn.edu

Abstract—Widespread adoption of sustainable energy sources is driving electricity grid operators to supplement hierarchical control regimes with market-based control that better motivates stakeholder involvement. However, to prevent market failures, such controls require testing before real-world implementation. The Power Trading Agent Competition is a competitive simulation of distribution grids that mirrors real-world scenarios and tests alternative policy and business scenarios. In Power TAC, broker agents acquire energy through bidding in a forward wholesale market to satisfy their customers overall demand on an hourly basis. In addition, a balancing market is intended to resolve real-time energy imbalances caused by broker prediction errors using demand response resources. As part of the annual alignment process, we discovered that brokers in the 2015 competition were persistently buying insufficient energy on the wholesale market to satisfy their customer demand. Instead, the balancing market made up the deficit, charging brokers a premium over the wholesale price. Also, demand response resources were heavily underused. We studied the economic impact of this systematic imbalance on brokers and discovered that they were behaving rationally, given the prices they faced in the two markets. We present the process and results of this analysis, and show how the balancing markets pricing mechanism can be adjusted for the 2016 competition to make it rational for brokers to achieve an overall neutral imbalance.

I. INTRODUCTION

Modern society depends heavily on reliable electric power [1], but our existing electricity infrastructure presents significant sustainability challenges [2]. Traditional approaches to grid management rely heavily on top-down control and fast-response, low-efficiency gas turbines [3]. These systems were not designed to work with large amounts of weather-dependent renewable energy sources like wind turbines and solar panels [4].

Many approaches have been put forward for revising the organization and management of electric power systems to enhance sustainability. However, it is difficult and risky to experiment with various policy prescriptions in the real world; witness the disaster visited on California in 2000 by an ill-considered market introduction [5]. For this reason, many groups have turned to simulation studies to evaluate market designs [6], policy prescriptions [7], and grid operations under high penetrations of variable-output renewable energy sources [8]. Some research also focuses on the introduction of demand response (DR) resources as a means of mitigating renewable energy volatility [9], [10]. In such studies, a simulation is typically carried out by a single researcher or research group. When entities in such a simulation must compete with each other across multiple markets, it can be quite difficult to ascertain that the simulation developers have considered all the ways competitors might exploit market design features. In such cases, a method known as Competitive Benchmarking (CB) offers much promise [11].

CB has three elements: a Platform, typically in the form of a competitive simulation environment that serves as a shared research foundation among a number of participating researchers [12]; periodic Alignment to refine elements of the platform and synchronize the platform with important elements of the real-world problem area being studied; and a Process that integrates contributions from stakeholders and participating researchers, and runs periodic competitions to assess the performance of the platform and competing agents contributed by researchers.

The Power Trading Agent Competition\(^1\) (Power TAC) [13] began in 2010 as a CB project to study retail future retail electric power markets, with a focus on improving the match between consumption and production (especially from distributed renewables) through market-based controls. Autonomous retail broker agents (or simply “brokers”) compete with each other to make profits by serving customers, buying energy in a wholesale forward market and selling it to customers through tariff contracts. Brokers are required to balance wholesale energy purchases with the consumption (and production) of their customers on an hourly basis; any residual imbalances are resolved through a balancing market. In this market, brokers can offer customer demand-response capacity [14], such as interruptible heating and cooling units and electric-vehicle batteries. The remaining imbalance is cleared using wholesale resources at a price computed as a premium over the “spot” price, the highest price at which the forward market has cleared for delivery in the current hour.\(^2\)

The Power TAC platform is designed to guide real-world decision-making by providing an environment for quantitative research. Previously, two common lines of research using Power TAC’s environment involved markets analysis [18] and autonomous retail broker agents [19], [20], [21]. It is central

---

1[^1]: see www.powertac.org

2[^2]: This is similar to the scheme used in the Nord Pool regulating market [15], details are outlined in [16] and in the Power TAC game specification [17]
to the applicability of these studies that the platform they are based upon holds relevance to the real world and also retains credible dynamics. This can be ensured through regular alignment events.

In the 2015 competition, competing brokers were generally not offering customer demand-response capacity within the balancing market. As part of the 2016 alignment, it would be interesting to determine whether this was a rational response to the market conditions they faced. Another research question arose during this inquiry: whether brokers found it rational to carry negative imbalances, thereby requiring the balancing market to systematically supply a portion of the energy their customers were using.

We can summarize our results in the following: (a) We determined that there was a strong bias toward negative imbalance, and that brokers indeed found it rational to purchase through the forward market less than the full quantity their customers were consuming. (b) Due to limited profitability, only 1 of the 11 competing brokers ever offered customer demand-response resources into the balancing market. (c) We identified the balancing-market premium as a key parameter that drives broker responses to the interaction of the forward and balancing markets, and we make recommendations on how this premium might be adjusted.

The following section provides more detail on the Power TAC markets. Section III explains the alignment process by detailing the measures taken to understand and remedy the problem. We conclude by summarizing our findings, future implementation goals, and possible future work in Section IV.

II. THE POWER TAC MARKETS

In the 2015 Power TAC Finals competition, the simulation modeled a small city’s electricity distribution market over a period of a little over 8 weeks. The city contains a multitude of customer models, from residential “prosumers” who both produce and consume energy, to industrial consumers with very stable hourly energy use. To focus on the economic aspects of market design, Power TAC simplifies most physical limitations of electricity grids. Since the small city is designed to house a large number of renewable energy sources, its energy mix tends to be strongly weather dependent. Hence, as an information input to create a realistic weather environment, real-world weather data and forecasts from an undisclosed location are injected into the simulation scenario and provided to brokers.

Electricity cannot be easily stored for later use; what is drawn from the grid in some location must be injected elsewhere. In Power TAC, retail brokers are required to purchase the energy their customers consume in each one-hour timeslot [22]. However, customer demand is stochastic, and the customer population includes a large quantity of distributed solar photovoltaic installations. Brokers can predict the solar output in their portfolios using weather forecasts. However, forecasts in Power TAC are not entirely reliable, because weather forecasts and current weather conditions are drawn from real-world data. Hence, brokers are faced with a prediction task in matching supply (mainly from wholesale market) and demand (mainly from retail market), the errors of which are meant to be solved in the balancing market. Since these markets deal in the same commodity, brokers have to make choices about how to interact with them [23].

A. Wholesale Market

The Power TAC wholesale market is a simple call market [24]. Offers (bids and asks) for delivery of energy in a given hour timeslot are cleared once in each of the previous 24 timeslots. In other words, energy for delivery in hour 50 can be traded 24 times, from hour 26 to hour 49. From the viewpoint of the market, offers arrive with delivery times from one hour to 24 hours in the future; the market sorts them by timeslot and clears each timeslot separately. Offers can be market orders that specify a quantity to be cleared at any price, or limit orders that specify quantity and price. Bids and asks are matched as long as there remain bids with prices at least as high as the lowest remaining un-matched ask. The last matched bid/ask pair (highest matched ask and lowest matched bid) set the clearing price for all transactions. Deferred transactions are generated for the matched offers, and are sent to the parties who submitted them. Payment is due on delivery. All participants receive an order book listing the clearing price, the quantity traded, and the quantities and prices of the unmatched offers. This allows brokers to estimate the prices they might see in the next round.

In addition to the retail broker agents, there are two other types of market participant:

- Gencos are wholesale energy suppliers. In the 2015 competition, all of the supply came either from the brokers themselves, or from an abstract Genco that generates a large number of asks with quantities and prices that produce a supply curve that is approximately quadratic when sorted by price. The coefficients of the quadratic vary over time through a mean-reverting random walk, and the bid quantities are selections from a normal distribution. The Genco keeps track of its commitments, and so at each market clearing it repeats the uncleared offers from the previous clearing, with some additional perturbation.
- An outside Buyer represents parties outside the scope of the simulation who also trade in the market. Their primary purpose is to provide liquidity to broker agents who wish to sell energy, and they also add some uncertainty to energy prices.

Brokers are free to learn and exploit price patterns in the forward market, buying energy for future delivery and then selling it back before the delivery time arrives, or even short-selling as long as they restore their positions before delivery.

B. Balancing Market

The balancing market operates as a single auction call market for demand response resources. Brokers are allowed to bid utilization of their DR resources for up-regulation and down-regulation; any additional imbalances are processed through a
mechanism similar to the Nordpool regulating market [15]. This market clears offers through a VCG mechanism, an incentive-compatible mechanism; as a result, it is rational for brokers to set bid prices equal to their costs [16]. Clearing prices are determined by the values of $P^+$ and $P^-$, by the slopes of the market’s price functions, and by the size of the total imbalance. Hence, broker’s offers will either fail to clear, or will clear at prices higher than or equal to bid prices.

In each timeslot, the overall imbalance is either positive or negative. A negative imbalance happens when brokers require extra energy to supply for underestimated demand, since they have purchased too little for delivery in the current timeslot. In this case, which we call up-regulation, we consider the price for energy to be positive, because the broker has to pay. On the other hand, a positive total imbalance entails that brokers have purchased too much energy for delivery, and therefore require down-regulation. Essentially this means the energy must be sold back or otherwise absorbed within the grid on short notice. The surplus can be used to charge batteries, add heat to water heaters, pump water uphill, or reduce the output of a generator somewhere, but it cannot be sold for the same price at which it had been purchased. To determine the price, the lowest price at which energy was sold in the forward market for the current timeslot is taken as spot sale (Figure 1). The down-regulation spot price $\text{spot}_{\text{sale}}$ is multiplied by a premium ratio $\phi^-$ to compute $P^-$, the zero-quantity price for down-regulation. The final down-regulation price is then $P_{\text{down}} = P^- + q_{\text{down}} \phi^-$ where $q_{\text{down}}$ is the surplus energy (in MWh), and $\phi^-$ is a constant. As a result, the price escalates to the disadvantage of brokers as imbalance increases in either direction. Up regulation is treated in an analogous fashion, except that the spot price is the highest price at which energy was purchased in the current timeslot, and $P^+$ is the spot price multiplied by a premium ratio larger than 1.

The premium ratios between spot prices and the zero-quantity balancing prices were $P^+_\text{+} = 1.1$ and $P^- = 1/1.1$ in the Power TAC 2015 Finals. The slopes were $\phi^+ = 1e-6 \text{ €/kWh}$ and $\phi^- = -1e-6 \text{ €/kWh}$. These parameters were chosen to roughly replicate the balancing market mechanism used in the Nord Pool regulating market and to ensure that brokers will always see a better deal in the forward market.

III. Aligning Incentives Between Markets

As part of the CB process, the Power TAC platform is annually re-aligned to improve its research value. This may involve incorporating real-world data and trends, adding or updating models, and making improvements based on analyses of agent and model performance over previous competitions. For example, the 2015 re-alignment added three new customer models. Two of these models, electric vehicles and electric fork-lift trucks, offer demand response capacity that brokers can utilize within the balancing market. The third, and the most influential in terms of market dynamics, was the addition of a large population of small retail solar producers. Occasionally, this population has sufficient capacity to offset all of the customer demand in the most extreme cases of high sunlight and low demand (see Figure 2).

A. Impact of Solar Capacity

The high volatility and uncertainty of the solar producer model’s output introduced a new tension in Power TAC’s simulated grid. While solar production through the retail market is a cheap alternative to wholesale market energy purchases, brokers face the challenge of predicting PV output and treating
the side effects of prediction errors. In Figure 2 we can see the impact of solar production in the retail market. The upper plot shows percentile contours of daily net demand (consumption less retail production) over all 230 games of the 2015 finals. The lower plot shows daily consumer consumption. There is a small amount of wind production in the model, but its variability is fairly constant. The high variability we see in the middle of the day comes from solar production. Thus, during the noontime hours, when solar production is highly volatile, brokers are prone to high output prediction errors, which would lead them to over-supply or under-supply their portfolio’s demand for those hours. This imbalance in supply and demand is managed in the balancing market.

We next investigated how brokers were being charged for this imbalance. In Figure 4 we see a scatter plot of broker profits vs. the broker’s RMS imbalance for each of the 230 games in the 2015 Finals. This plot again confirms individual brokers’ imbalance numbers are abnormally high. More importantly, we see little correlation between the imbalance performance of any brokers and their overall profits. This is not what we hoped to see.

Through investigating energy trades within the two markets, we noticed a strong presence of negative imbalance events in the balancing market (about 60% of imbalance cases). Also, the amount of imbalance was much more that what we see in real world data. The mean imbalance across all timeslots in all games from the 2015 Finals was −4.111 MWh (with a standard deviation of 10.76 MWh), which is over 10% of mean net demand. Given that imbalance presumably arises from a broker agent’s inability to predict supply and demand in its portfolio, we expect imbalance to have significant correlations with unpredictability of supply and demand. Deviations in solar production and retail consumption, the main two sources of energy unpredictability, indeed seem to have strong correlation with imbalance (Figure 3). A Pearson’s correlation test showed a coefficient of −0.5999 and for net imbalance versus the standard deviation of consumption, and 0.5771 for imbalance versus standard deviation of production (p-values are 0.0019 and 0.0032, respectively). Thus, imbalances were dependent on production and consumption predictability, whereas in 2014, they were mostly dependent on only consumption predictability.

A lack of correlation between broker imbalance and profits suggested that the balancing market may be priced too cheaply; that is, the brokers may have little economic incentive to balance their energy supply and demand. In other words, brokers may have found it profitable to acquire a portion of their energy needs through the balancing market. The next step is to understand the economic incentives faced by brokers with regard to energy imbalances.

We investigated demand response resource (DRR) bids within the balancing market to understand how the market was being used. We noticed that out of eleven participants, only one (CrocodileAgent [18]) offered DRR into the balancing market. Although the quantities were small, it earned a 44% margin when these offers were exercised. The surplus available to brokers for DRR offers is the difference between what brokers earn from the balancing market and the price brokers offer to their customers for the use of these resources. In addition to again indicating that balancing is too cheap, this disregard for DRR forced nearly all energy balancing to be performed by the regulating market entity (RM).

Based on these two observations, we determined that the RM entity was offering energy trade at prices too cheap to suitably demotivate imbalances and motivate DRR use. Mathematically speaking, the ratios $P^+$ and $P^−$ proved to be too close to 1.0 to provide the correct economic incentives. The problem is that while the balancing market cost is always less attractive than what brokers have already paid for energy being supplied to customers, it is not necessarily less attractive than

![Figure 3. Net imbalance and the standard deviations of imbalance, consumption, and production per hour of day aggregated over all games in the 2015 Finals.](image)

![Figure 4. Broker total profits in each game versus root mean squared imbalance for the 2015 Finals games.](image)

**B. Balancing Market Pricing**

We next investigated how brokers were being charged for this imbalance. In Figure 4 we see a scatter plot of broker profits vs. the broker’s RMS imbalance for each of the 230 games in the 2015 Finals. This plot again confirms individual brokers’ imbalance numbers are abnormally high. More importantly, we see little correlation between the imbalance performance of any brokers and their overall profits. This is not what we hoped to see.

We investigated demand response resource (DRR) bids within the balancing market to understand how the market was being used. We noticed that out of eleven participants, only one (CrocodileAgent [18]) offered DRR into the balancing market. Although the quantities were small, it earned a 44% margin when these offers were exercised. The surplus available to brokers for DRR offers is the difference between what brokers earn from the balancing market and the price brokers offer to their customers for the use of these resources. In addition to again indicating that balancing is too cheap, this disregard for DRR forced nearly all energy balancing to be performed by the regulating market entity (RM).

Based on these two observations, we determined that the RM entity was offering energy trade at prices too cheap to suitably demotivate imbalances and motivate DRR use. Mathematically speaking, the ratios $P^+$ and $P^−$ proved to be too close to 1.0 to provide the correct economic incentives. The problem is that while the balancing market cost is always less attractive than what brokers have already paid for energy being supplied to customers, it is not necessarily less attractive than...
what brokers *would have to pay* for the energy needed to clear the imbalance through the forward wholesale market. This is because the supply curve – the relationship between quantity and price over all the asks (offers to sell) in the forward market – is roughly quadratic to model the effects of merit order and congestion pricing in real-world markets [25]. Since brokers can observe the order books, they may have been able to predict that energy trades through the forward market would have been more expensive than expected balancing market prices.

### C. Incentive Alignment

One method to make RM balancing less profitable than trading in the forward market, is to price balancing according to how much it may have cost were it to be solved by the forward market. To calculate this, we computed how much the energy imbalance would have cost for each broker, if all brokers had predicted accurately and traded all energy imbalance at the last possible moment in the forward market, prior to entering the balancing market. Next, we calculated the $P^+$ and $P^-$ values that would have produced these prices in the balancing market, and the ratios by which $P^+$ and $P^-$ would have to change to correct the interaction between the balancing and forward markets.

We ran these calculations on all 230 games in the 2015 Finals competition, over all timeslots and all brokers within each game.\(^4\) We ignored cases in which a broker’s imbalance was the opposite sign of the total imbalance, as the handling of these scenarios did not depend on the market mechanism that dissuades broker imbalance. A few timeslots for which data was not adequately recorded were also pruned; there were too few to affect the results.

Figure 5 shows how much the premiums need to be changed to remove the advantage brokers can achieve by relying on the balancing market to resolve their imbalances. The 95th percentile values of $P^-$ and $P^+$, are values for which 95% of brokers’ imbalance cases have prices in the wholesale market equal to or better than the prices for the same energy imbalance in the balancing market. Thus, they indicate the ratios that would encourage balancing through the forward market in 95% of cases in which a broker’s imbalance is similar to overall imbalance. We posit that the ratios at this value level would satisfy our alignment goals with little negative repercussions. The 95th percentile values over all timeslots for the $P^-$ and $P^+$ ratios are found to be 0.096 and 2.52, respectively.

Increasing $P^+$ by 2.52 would suggest that in 95% of cases, the wholesale market would either be economically preferable or as good as the balancing market’s regulation prices. Similarly, multiplying the $P^-$ ratio by 0.096 would mean that in only 5% of cases where total imbalance is positive, would it be rational to carry surplus energy into the balancing market.

\(^4\)All Power TAC 2015 Finals game logs are available at http://xlarge.rsm.nl-finals_2015_04/. Tools used for data extraction can be found at www.github.com/powertac/powertac-tools

Within the $P^+$ and $P^-$ ratio figures, we can observe a clear diurnal pattern. Also, in the top plot of Figure 5 the ratio appears to stabilize after the first days. At least part of this pattern is likely due to the high churn in the tariff market over the first few days of a game as brokers compete for customers. It is probably also a consequence of agent learning as they build their portfolios. This learning behavior could also aid agents in predicting potential market costs and balancing costs, finding cases in which the latter is cheaper, and thus continuing to exploit price differences across these markets in some timeslots. Regarding the diurnal fluctuation, the $P^+$ ratio values appear to be high where solar output is most unpredictable (and highest) and in peak demand hours (Figure 6). Note that the peaks match well with the high-variability hours shown in the top plot of Figure 2. The peaking regions could possibly be predicted by brokers, enticing them to forgo some energy trade in the wholesale market in favor of the balancing market’s potential prices once the latter has closed. However, with a $P^+$ ratio of 2.52, this does not appear to contribute significantly enough to warrant concern. The $P^-$ ratio plots did not have any recognizable patterns.
IV. CONCLUSIONS AND FUTURE WORK

The Power TAC simulation platform is designed to be a Competitive Benchmarking testbed for viable policy and business scenarios in the electricity distribution grid. It also seeks to provide an environment within which quantitative research can be conducted. However, such an environment demands routine alignment to sustain research value and retain relevance to practical situations.

Alignment changes in the 2015 competition year exposed a maladjustment in the Power TAC balancing market. In the Power TAC 2015 Finals competition, the broker agents significantly underused demand response resources and routinely passed excess demand to the balancing market. We calculated a possible adjustment of the balancing market’s pricing mechanism that would make the balancing market generally unfavorable to the wholesale market. This adjustment will also motivate brokers to utilize demand response resources, since brokers can expect higher profits from balancing orders.

Through implementation in the 2016 competition, this adjustment can reconfigure Power TAC’s balancing market to provision balancing for a customer base that incorporates recent energy technology changes. This may hold strong relevance for many regions, such as Germany and the US state of California, which are transitioning from fossil fuels to an energy system based on clean but occasionally unpredictable renewable energy sources.

Possible future work on Power TAC may focus on capacity controls for peak demand and tariff schemes for demand response resources.

REFERENCES