System imbalance from solar energy trading

Tomas Landelius
Swedish Meteorological and
Hydrological Institute (SMHI)
601 76 Norrköping, Sweden
Email: tomas.landelius@smhi.se

Sandra Andersson
Swedish Meteorological and
Hydrological Institute (SMHI)
601 76 Norrköping, Sweden
Email: sandra.andersson2@smhi.se

Roger Abrahamsson
Tekniska Verken Linköping Nät AB
Brogatan 1
582 78 Linköping
Email: roger.abrahamsson@tekniskaverken.se

***Abstract*—The aim of this paper is to study how the use of an advanced trading method, like the optimal quantile strategy, may effect the balance between generation and consumption at the power system level when trading solar PV power on the Nordic Power Exchange. In order to do this we first developed a set of PV power forecast models. Numerical weather prediction data together with power measurements at 210 PV installations, in the regional network operated by Tekniska Verken Linköping Nät AB, were used for estimation and evaluation. Linear and non-linear regression, the latter in terms of an artificial neural network, both resulted in an RMSE, normalized with installed power, of about 6 %. Second we used the neural network to perform a three month simulation experiment on the Nord Pool Elspot day-ahead market. Strategies based on deterministic forecasts were compared with the use of the optimal quantile, based on ensemble forecasts of the power probability distribution. The optimal quantile strategy resulted in an increased revenue of around 2 % but also in an increased imbalance between contracted and produced energy of almost 20 %. The imbalanced part of the power production for the optimal quantile strategy was about one third. A similar study, on trading wind power with the same strategy from a hypothetical plant on the Nord pool market, showed that about half of the traded energy became imbalanced.**

I. Introduction

The solar energy resource is fluctuating, non-trivial to predict and only dispatchable at high costs. Hence, producers of solar power participating in liberalized electricity markets are subject to regulation costs. On such markets, energy bids are proposed in advance, and the bidders are then charged for any imbalance between the actual production and the bids. With more variable solar power being urged to enter the market, imbalances from individual traders may add up and result in increased imbalances between production and demand at the power system level. This could cause challenges to the integration of solar energy into the power system.

It is clear that solar power forecasts are essential when trying to maximize ones benefits on such a market. Previous studies on wind energy trading show that the use of the so-called optimal quantile strategy (OQS) outperforms strategies based on deterministic point forecasts [1]. OQS uses probabilistic energy forecasts together with forecasts of the prices for up and down regulation in order to minimize the expected imbalance costs for the trader. However this can, somewhat counter-intuitive, result in an increase in the actual imbalance between bids and production. If many wind energy producers would employ this same strategy it might no longer be optimal. The reason is that it is based on the price-taker assumption where the producer is without market power. Moreover, if the strategy became commonplace, it may even act destabilizing on the power system [2].

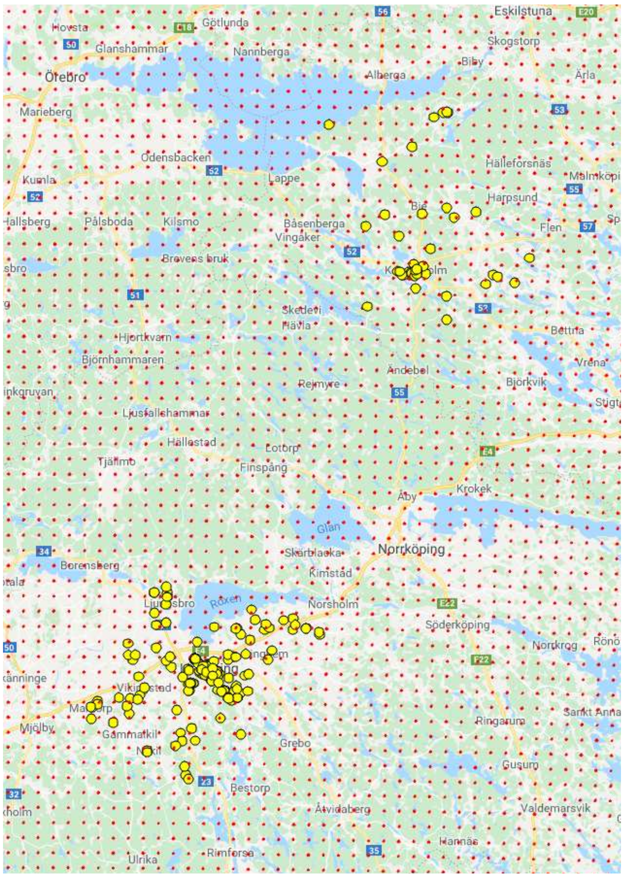


Fig. 1. NWP grid points (red) and PV installations (yellow) in the network managed by Tekniska Verken AB. Map created with the Python package `gmplot` using Google Maps (2018) as background.

In this paper we investigate how pronounced this effect may be when trading solar power using the OQS. For this purpose we take on the role of a fictitious market participant trading the available solar power from the Swedish regional network operated by Tekniska Verken Linköping Nät AB on the Elspot day-ahead market. In this network there are close to 300 solar power installations with a total nominal effect of about 6 MW.

In section II we present the data used for this study and in section III we develop a linear regression model and an artificial neural network (ANN) for prediction of the total net solar power production in the regional network. The model input consists of measured net load and solar power production from the previous day along with probabilistic

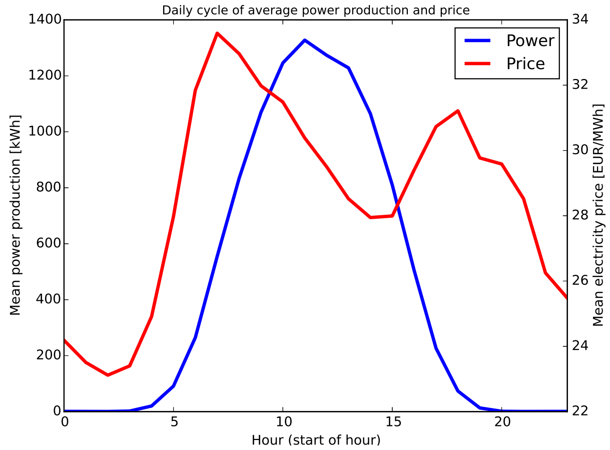


Fig. 2. Daily cycles of average net PV power production in the study area along with average prices for the Elspot area SE3 during April - June 2017.

km scale numerical weather prediction (NWP) ensemble data. Here we also revisit the formulation of the optimization problem facing the market participant and how it can be solved with the OQS.

A market simulation experiment using the ANN forecasts together with the OQS is presented in section IV-B. Here deterministic trading strategies are evaluated in competition with the probabilistic OQS. Forecasting the regulation costs necessary for the OQS is beyond the scope of this paper. Instead we use either persistence or a perfect forecast for this purpose. The perfect forecast is given by regulation costs data from the Nord Pool archive. This provides us with reasonable lower and upper bounds on the performance of the strategy with respect to the market information.

In section IV we summarize the results from the model and strategy evaluations. The paper ends with section V on discussions and conclusions.

II. Data

This section describes the three data sets that were used for this study; measured net power production and consumption, NWP ensemble forecasts and Elspot prices for the corresponding dates. All data have hourly time resolution and refer to the time periods March - October 2016 and April - June 2017. The data was divided into one training set, consisting of the data from 2016, and one set with the data from 2017 used for model evaluation and market simulation experiments.

A. Power measurements

The electricity measurements were provided by Tekniska verken that is responsible for the electrical grid in the Municipality of Linköping, parts of the Municipality of Mjölby and large parts of the Municipality of Katrineholm in Sweden. Hourly measurements of net PV power production and net consumption are available from 220 sites with PV installations in the Linköping net, and 51 sites in the Katrineholm network. For each installation information about the geographical location and the installed effect is also known. Measurements of the gross PV production are not mandatory in Sweden. Such information is only available for a very limited number of sites and was not used in this study.

Some of the sites were excluded due to limited data availability. In the end, data from 210 power plants with a total installed effect of 3.75 MW was used. These installations are located on 125 households (1030 kW), 54 apartment complexes (1495 kW), 10 office buildings (515 kW) and 21 farms (710 kW) and are scattered over an area of about 100 x 150 km^2. They all reside inside the Elspot area SE3, see figure 1. The daily cycle of the net power production from the 210 installations is shown with a blue line in figure 2. The average net production peaks at about 35 % of the installed effect.

B. MetCoOp Ensemble Prediction System (MEPS)

MetCoOp is a cooperation around operational NWP production between Sweden, Norway and Finland. MEPS is developed in the framework of the shared Aire Limite Adaptation Dynamique Developpement InterNational (ALADIN) - High-Resolution Limited-Area Model (HIRLAM) NWP system. This system can be run with different configurations and here version 40h1.1 of the so-called ALADIN-HIRLAM Regional Meso-scale Operational NWP In Europe-Application of Research to Operations at Mesoscale (HARMONIE-AROME) is used [3], [4]. The main components of the ALADIN-HIRLAM NWP system are surface data assimilation, upper-air data assimilation and the forecast model for the forward time integration. MEPS provides both global (GHI) and direct normal irradiance (DNI) solar radiation fluxes at the surface. The radiation model is based on the Morcrette radiation scheme from cycle 25R1 [5] and uses the Rapid Radiative Transfer Model of Mlawer et. al [6].

The MEPS model domain is made up of 900 x 960 points with 2.5 km grid spacing and 65 levels covering the Nordic region. The grid is defined by a Lambert projection with the center at 63.5°N and 15°E. An example showing the layout of the gridpoints within the region of interest for this study is given in figure 1. The ensemble consists of one control (deterministic best estimate) plus nine perturbed ensemble members. The SLAF method [7] is used to produce initial and boundary perturbations from ECMWF deterministic forecasts using a lagging technique.

Unfortunately only the control member is archived operationally at SMHI. However, all MEPS member forecasts started at 00 UTC with a length of 24 hours and hourly time steps were temporarily archived during April - June 2017. These have been used for the present study along with archived control member forecasts from April - October 2016. The parameters of interest here are the accumulated hourly values of GHI and DNI together with instantaneous values of two metre temperature and 10 metre wind speed.

C. Spot prices and regulation costs

As noted earlier, all of the PV power plants in the network under investigation are located within the price area SE3 of Nord Pool's Elspot market. For the trading simulations we downloaded historical data for the Elspot day-ahead market from `https://www.nordpoolgroup.com`. More specifically we obtained hourly data regarding the Elspot price and the up and down regulating prices (€/MWh) for the SE3 area during the years 2016 and 2017. The price data refers to hourly intervals.

The daily cycle of the Elspot price during April - June 2017 is illustrated with the red line in figure 2. It shows that the timing imbalance between peak demand and PV power production is also present in the SE3 area of the Elspot market given the present PV installations. The morning and afternoon peaks in the demand should also be applicable to our study where about two thirds of the installed power is attributed to PV-installations on households and residental buildings.

III. Methods

In order to asses if and how different trading methods affects the imbalance we set up a test case during April - June 2017. To perform a market simulation we need to be able to forecast the total net solar PV power production from the portfolio of installations and apply a trading strategy to place bids on the market. In this section we describe three different forecast methods (persistence, linear regression and an ANN) for the power production. We also briefly describe the trading problem and the optimal quantile strategy that was used to derive optimal bids based on information from the ensemble forecasts.

A. Predicting net PV power production

As a hypothetical actor on the Nord Pool Elspot market, specialized on trading solar power, we need to be able to forecasts the production for the day-ahead market. Here we use an approach similar to that in our previous study where gross and net PV production were predicted at individual sites [8]. NWP information from the MEPS, along with measurements of net production and net consumption from the previous day were used as input data for the forecast. To make things simple we did not distinguish between weekdays and holidays. Furthermore we assumed that we have access to the NWP forecast issued at 00 UTC and all measurements up to the hour when the bid is placed. In reality the bids should be placed at 12 UTC the previous day and a more realistic experimental set-up will be the subject for a future study.

1) Persistence: The persistence model does not involve any parameter estimation and the input is simply the same as the output - the observed total net PV power production value for the corresponding hour from the previous day.

2) Linear regression: The input vector (x) to the linear regression consist of the NWP forecast data (GHI, DNI, two metre temperature and 10 metre wind speed for the next 24 hours as well as for the past 24 hours) together with measured net production and consumption from the previous day. The idea behind this is that model should be able to make a connection between yesterdays state of the atmosphere (containing information about the gross production) and the net production and consumption during the previous day. We also included the cosine of the solar zenith angle to provide some information about the time of the day and the season. It should also help to account for the effect of PV on tilted surfaces. A constant was added to the input vector in order for the model to be able to add a bias.

Hourly NWP data for the GHI and DNI were obtained by taking differences between accumulated forecasts from $00 + LL + 1$ and $00 + LL$ where LL is a forecast length between one and 24 hours. Hence forecasts for the time intervals $00 - 01$ UTC, ..., $23 - 24$ UTC were obtained. The two metre temperature and 10 metre wind speed are available as instantaneous values. Here we used the values assigned to the start of the interval. The measured power production and consumption refer to the same intervals as the GHI and DNI forecasts.

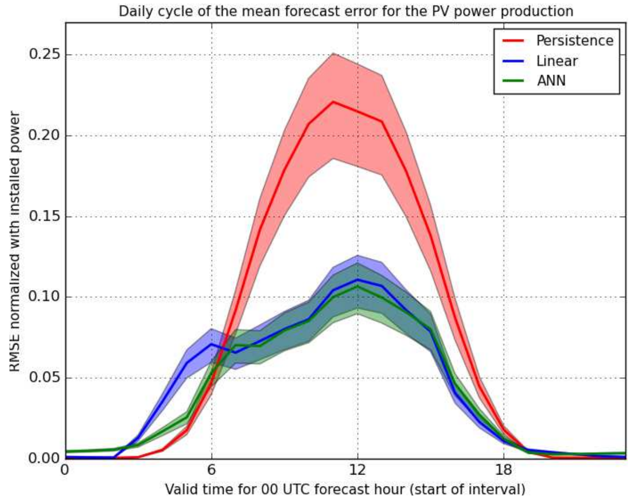


Fig. 3. Daily mean RMSEn cycle for the PV power predictions (persistence: red, linear model: blue, ANN: green). The shaded areas illustrate the 95 % confidence intervals.

Instead of modelling the power production at each individual site and then accumulate, we used a single input vector obtained as a weighted sum of the input vectors for each site. The weights are given by the relative installed power at each site:

$$x = \sum_k p_k^n x_k / \sum_k p_k^n. \tag{1}$$

Here p_k^n and x_k refer to the installed power and the input vector for a given site. A linear least squares method was used to estimate the model parameters (Python function `numpy.linalg.lstsq`):

$$w^\star = \operatorname*{argmin}_w \sum_d \sum_t (w^T x - p_o(d,t))^2, \tag{2}$$

where $p_o(d,t)$ denotes the observed sum of the net power production at all sites at time t of the day d. Data for the estimation was taken from the time period April - October 2016 when only the MEPS control member was available.

3) Artificial neural network: The use of machine learning has a long history within the area of energy forecasting [9]. In this study we used an off the shelf ANN from TensorFlow [10] to see if it could perform better than the linear model.

We employed the standard feedforward ANN (`DNNRegressor`) from TensorFlow with a three layer feedforward topology using one input, one hidden, and one output layer. In our set-up it had 11 inputs, 32, 64 or 96 nodes in the hidden layer and 1 node in the output layer.

Determining the number of hidden neurons in the hidden layer(s) is a trade off between the networks ability to generalize from the training data and its representative power. As a guide we used an empirical relation for the number of hidden layer neurons: $n_{hl} = 0.5(n_{in} + n_{out}) + \sqrt{n_{train}}$ [9]. Here

n_{in}, n_{out} and n_{train} denotes the number of input, output and the size of the training data set respectively. In our case the number of inputs equals 11 and we have one single output (the total net production). The number of cases in the training data set was 5136. Hence the suggested number of nodes in the hidden layer becomes 78. To check the robustness of this choice we decided to try with 32, 64 and 96 neurons in the hidden layer. The network with 32 neurons performed slightly better than the others on the evaluation data set so the results presented here are based on the outputs from that network.

The ANN model use the same input as described for the linear regression approach. In order to harmonize the amplitudes of the input variables we removed the mean (m) and normalized the input vector by multiplying it with the inverse of the square root sample covariance matrix (C), calculated from the training data set. The TensorFlow minimization algorithm then finds the solution to

$$w^\star = \underset{w}{\operatorname{argmin}} \sum_d \sum_t (w_o^T f(w_h, C^{-\frac{1}{2}}(x-m)) - p_n^o(d,t))^2 \tag{3}$$

using an iterative procedure. Here we used data from the time period April - October 2017 for estimation and from April - June 2017 for evaluation. In both cases NWP data from the MEPS control member was used.

B. Electricity trading

We now turn to the problem of trading PV power using the model presented in the previous section. First we formulate the trading problem in terms of a cost function and then describe strategies for how to make bids that minimize it. For this presentation we use a notation and problem formulation that is close to that in Pinson et. al [1].

1) Cost function: On the Elspot day-ahead market, power producers have to announce the amount of energy they are willing to deliver over the following trading period for every unit of time. Here we restrict ourselves to a trading period of 24 hours and hourly time units, each referred to as a program time unit (PTU).

For any PTU (at time t) during the following day, market participants have to propose an amount of contracted energy $E_c(t)$. Now the revenue from proposing an amount of $E_c(t)$ but in fact producing the amount $E_o(t)$ is given by

$$R(t) = \pi_c(t)E_c(t) + T_c(t). \tag{4}$$

Here π_c and T_c denote the spot price and imbalance cost respectively. The imbalance d_o is given by the difference between the actual and contracted energy amounts:

$$d_o(t) = E_o(t) - E_c(t). \tag{5}$$

The regulating power market is managed by the Tranmission System Operator (TSO). In case consumption exceeds generation, the TSO ensures that the producers deliver more electricity to the grid. When generation exceeds consumption its the other way round and the TSO must see to that producers reduce the generation. These procedures to regulate imbalances at the system level are called up- and down-regulation. During an hour with up-regulation on the Elspot market, producers producing too much will only get paid the market price. On the other hand, producers producing too little during up-regulation will instead be charged the up-regulating price (higher or equal to the market price). During hours with down-regulation, producers producing too much will get paid the down-regulating price (lower or equal to the market price). Producers producing too little will in this case be charged the full market price.

The imbalance cost can hence be written as

$$T_c(t) = \begin{cases} \pi^+(t)d_o(t)\ , & d_o(t) \geq 0 \\ \pi^-(t)d_o(t)\ , & d_o(t) < 0 \end{cases} \tag{6}$$

where π^+ and π^- are the prices associated with positive (down-regulation) and negative (up-regulation) imbalances respectively.

Let us introduce regulation unit costs in terms of differences between imbalance costs and the spot price π_c as

$$\pi^{u+}(t) = \pi_c(t) - \pi^+(t) \tag{7}$$

$$\pi^{u-}(t) = \pi^-(t) - \pi_c(t). \tag{8}$$

The revenue can then be re-formulated as a sum of the energy actual produced minus the regulation costs

$$R(t) = \pi_o(t)E_o(t) - T_o(t), \tag{9}$$

where the cost is now expressed in terms of the unit regulation prices:

$$T_o(t) = \begin{cases} \pi^{u+}(t)d_o(t), & d_o(t) \geq 0 \\ -\pi^{u-}(t)d_o(t), & d_o(t) < 0 \end{cases} \tag{10}$$

Using this formulation the first terms now corresponds to the revenue from using a perfect forecast, independent of the bidding strategy. Hence, maximizing the revenue then means minimizing the cost T_o.

The revenue can also be expressed in terms of a performance ratio. This measure will be used when evaluating different bidding strategies in the next section. The performance ratio is calculated for a given number of days and PTUs by normalizing the actual revenue by the revenue had one used perfect forecasts:

$$\gamma = 1 - \frac{\sum_d \sum_t T_o(d,t)}{\sum_d \sum_t \pi_c(d,t)E_o(d,t)} \tag{11}$$

Note that the performance ratio will have the property that $\gamma \in (-\infty, 1]$ where 1 will correspond to the performance of bidding with a perfect forecast.

2) Trading strategies: Determinisitic or point prediction methods forecast the expected power production. Their output can be seen as a measure of the average power output, e.g. during a PTU of one hour as is the case here. Hence we can use the power forecasts as forecast of the energy for a PTU with a one hour duration.

With access to ensemble forecasts we can also take a probabilistic approach. The imbalance in equation 5 will then be treated as a realization of a random variable. Instead of a point forecast, E_c, we will use an estimate of the probability for a given energy production at a given time, $P(E,t)$.

With a probabilistic approach, minimization of the imbalance cost T_o in equation 10 is replaced with the minimization of its expectation value

$$E_c(t) = \underset{E}{\operatorname{argmin}} \int_0^\infty T_o(t)P(x,t)dx \tag{12}$$

where we use $d_o(t) = x - E(t)$ in the calculation of $T_o(t)$.

This minimization problem is known as a standard linear terminal loss problem. The reader is referred to statistical textbooks for the derivation of its solution, e.g. [11]. The bid that minimize the expected loss is given by

$$E_c(t) = F_E^{-1}\left(\frac{\pi^{u+}(t)}{\pi^{u+}(t) + \pi^{u-}(t)}\right). \quad (13)$$

Here F_E^{-1} is the inverse cumulative distribution, or quantile function, of the random variable E. This is the reason why this bidding strategy is known as the optimal quantile strategy. In our case we had 10 ensemble members providing information about the distribution $P(E)$ and we used the Python function `percentile` to obtain the optimal quantile in equation 13. Note that this means that the optimal bid is in general not given by a deterministic forecast since such a forecast only provides an estimate of the expectation value of the produced energy.

Moreover, to obtain the optimal bid one must not only have a forecast for the quantile function but also of its argument, the percentage point, given by the ratio

$$r(t) = \frac{\pi^{u+}(t)}{\pi^{u+}(t) + \pi^{u-}(t)}. \quad (14)$$

When both the down and up regulation prices equal the spot price we end up with a zero in the denominator. In these cases we set $r(t) = 0.5$ in order not to favour a certain direction of the imbalance. Note that the percentage point will almost only takes on the values 0, 0.5 and 1 since most of the time, either the up or the down regulation price equals the spot price.

To illustrate the workings of this strategy, consider the case when the unit cost for producing too much (positive imbalance) is much higher than the cost for producing to little (negative imbalance). In such a situation where $\pi^{u+} \gg \pi^{u-}$, the percentage unit will be close to 1. The OQS then suggests us to contract as much energy as possible in order to end up producing less than the contracted amount and obtain a negative imbalance. When the cost for producing too little outweighs the cost for producing too much, the percentage unit will be close to zero and the situation is reversed.

Forecasting the percentage point is beyond the scope of this paper. Instead we used either a perfect forecast or the average percentage point during the previous day when comparing OQS to strategies based on deterministic forecasts. We did look at the daily and yearly cycles of the percentage point ratio during the year 2017 but found no significant patterns.

TABLE I

UPPER AND LOWER 95 % CONFIDENCE LEVELS FOR THE MEAN SQUARED CORRELATION VALUES AND NORMALIZED (WITH INSTALLED POWER) RMSE AND BIAS FOR MODELED PV POWER PRODUCTION USING PERSISTENCE, LINEAR REGRESSION AND ANN.

	Persistence	Linear	ANN
Squared correlation	0.56, 0.61	0.86, 0.88	0.87, 0.89
RMSEn [%]	11, 12	5.8, 6.1	5.5, 5.8
BIASn [%]	-0.51, 0.44	0.11, 0.61	0.15, 0.62

TABLE II

SUMMARY OF DETERMINISTIC AND OQS STRATEGIES USED FOR COMPARISONS IN THE MARKET SIMULATION EXPERIMENT.

Deterministic strategies	Optimal quantile strategies
Persistence for p	Ensemble for p, persistence for r
Control member for p	Ensemble for p, perfect forecast of r
Ensemble mean for p	
Perfect forecast of p	

IV. RESULTS

In this section we first present results from a comparison of the different models for predicting net PV power production, followed by findings from a test case where bidding strategies based on OQS and deterministic forecasts were compared.

A. *Predicting net PV power production*

In order to evaluate the performance of the different models for prediction of the net PV power production we looked at the bias and root mean squared error (RMSE) normalized by the installed power (RMSEn) and calculated the square of the Pearson's correlation coefficient between the modelled and observed net PV power production.

The daily cycles of the mean RMSEn when using the linear regression and the ANN models are illustrated in figure 3 along with the 95 % confidence intervals. The error corresponding to a persistence model is included for comparison. The error is shown as a function of forecast length. All forecasts are initialized at 00 UTC. Both the linear and the ANN models perform significantly better than persistence throughout most of the day. However, the linear model is somewhat worse during the morning hours. Aside from that the linear and ANN models show similar scores as summarized in table I. The results are in line with those from our previous study on net PV power prediction at individual sites [8]. Even though the ANN did not significantly outperform the linear model we decided to use it in the trading test case described in the next section.

B. *Trading test case*

Solar power is neither dispatchable nor easy to predict. Hence, trading solar energy will result in imbalances when either too much or too little energy is contracted during day-ahead trading. In order to evaluate the OQS strategy and investigate its effect on the resulting imbalances we set up a market simulation experiment. The simulation was run with hourly time resolution for the period April - June 2017 since this was the time period for which we had available ensemble NWP data. Corresponding market data for the Elspot SE3 area was obtained as described earlier.

The assumptions were the same as in the previous study for wind energy [1]. We review them in brief here. First, potential effects related to solar energy penetration in the market are neglected, i.e. the amount of solar energy that enters the power system is so small that it does not affect the spot price. Second, the solar power trader is considered as a price-taker and again too small to affect neither the spot nor the the imbalance price. Third, the trader does not do any corrective actions during the intra-day market.

In this experiment, the solar power producer participates in the market with the aim of maximizing the expected revenue during the three months. This then translates into minimizing penalties from up and downward regulation and that the optimal strategy is given by the solution to equation 13. However, the solution of this equation calls for a forecast of the percentage point r. In this experiment we compared two such forecasts. To get optimistic and pessimistic bounds on the performance of the OQS strategy we assumed either to have a perfect forecast, r_o, or a persistence like forecast given by its mean value during the previous day, r_{-1d}. These two OQS strategies were then compared with four strategies based on deterministic forecasts resulting in a comparison of six different alternatives as outlined in table II. The performance of these strategies during the market simulation experiment is summarized in table III.

Trading using the deterministic power forecast based on persistence results in more or less no difference between the positive and negative imbalances. Using the control member or the ensemble mean (both based on the ANN) also results in about equal amounts of surplus and shortage as shown in table III. This is in line with the model performance results related to bias presented in table I.

The strategy based on persistence contracts a similar amount of energy as that based on a perfect forecast. However, almost half of the production becomes imbalanced with similar amounts in surplus and shortage. Using persistence one trades at the highest average energy price but still ends up with the worst results, whichever performance measure we look at. Using a good forecast pays off in terms of greater revenue for the trader since it decreases the financial risks related to regulation costs. Note however, that the strategy resulting in the largest revenue (OQS with perfect forecast for r) actually both contracts the largest amount of energy and at the same time results in a larger imbalance (32 %) than any of the other non-persistence strategies.

Figure 4 illustrates the increase in accumulated revenue between the strategy based on a perfect forecast, the two OQS versions and the ensemble mean compared to that from a strategy based on persistence only. Note that there are flat periods where persistence works well and abrupt steps where such an assumption breaks down. There is a large gap between the end result from the two OQS strategies indicating that the quality of the forecast of the percentage point r is essential.

The average spot price during April - June 2017 was 28.51 €/MWh. However, due to the timing imbalance between peak demand and PV power production, illustrated in figure 2, the highest average price per produced MWh a solar producer will obtain is something different. With a perfect forecast of the solar power, the average price equals 28.75 €/MWh while the best strategy trades at, 29.95 €/MWh and persistence at the highest average, 30.29 €/MWh.

Reducing imbalances does not necessarily result in increased revenues. It is actually the other way round, unless one have a perfect forecast. Both of the optimal quantile strategies perform better than the deterministic strategies based on the control member or the ensemble mean. OQS with a perfect forecast for r produce more imbalanced energy (32 %) than its deterministic counterparts while OQS

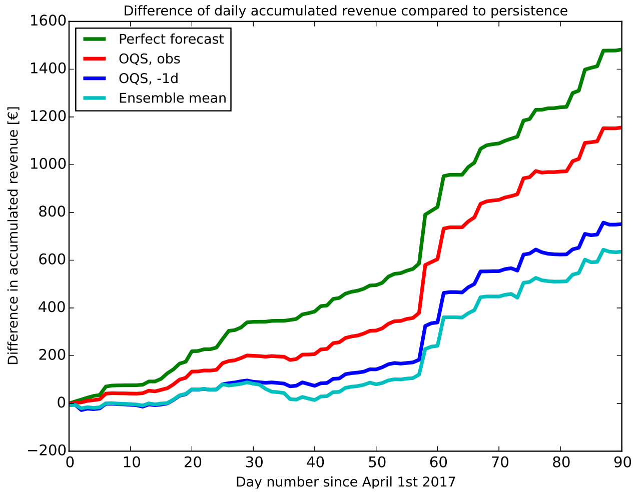


Fig. 4. Difference in accumulated revenue from trading strategies based on perfect forecasts (green), OQS with actual information about the regulation prices (red), OQS based on a mean value of yesterday's regulation prices (blue) and using the ensemble mean forecast (cyan) compared to the revenue obtained using a persistence strategy.

with persistence for r results in almost the same amounts of imbalanced production. The reason for this is that the OQS trades in such a way that an imbalance risk for any given time is matched with its corresponding regulation cost. Hence, the average unit regulation cost for the optimal quantile strategies are the lowest.

During the three month market simulation, the average unit costs for down-regulation (positive imbalances) were higher than for up-regulation (negative imbalances). This was true for all the non-perfect trading strategies. As a consequence a trader would prefer to make bids that are more likely to result in a shortage than in a surplus as confirmed by the figures in table III. This situation is desirable for the TSO since the relation between the up and down-regulation prices reflects how sensitive the TSO is to negative and positive system imbalances respectively.

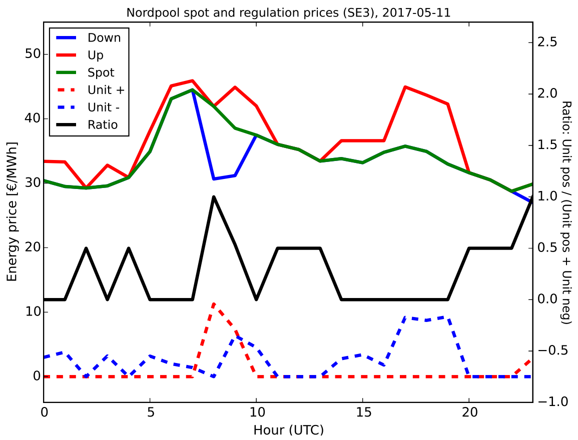
Figure 5a shows an example with spot and regulating prices for May 11:th 2017 along with the corresponding percentage point (black line). Note that the percentage point only takes on the values 0, 0.5 and 1 as pointed out earlier. The result using the OQS and the observed percentage point (perfect forecast) for this date is illustrated in figure 5b. Note how the forecast switches between the minimum (r= 0), the median (r= 0.5) and the maximum (r= 1.0) percentiles of the forecasted solar power distribution.

V. Discussion and Conclusion

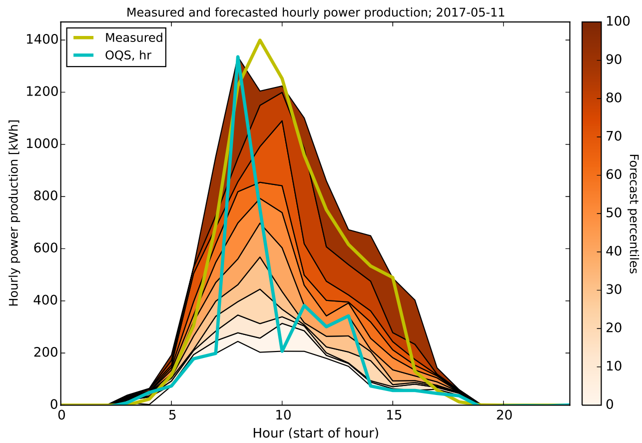
In this paper we developed and evaluated a linear and an ANN model for day-ahead prediction of the total net solar power production from 210 PV installations in an area of about 100 x 150 km^2 in the Southern part of Sweden. The linear model performed almost as well as an off the shelf ANN from TensorFlow. In terms of RMSE, normalized with the installed power, the linear model and the ANN ended up with values of 5.9 and 5.6 % respectively. Some degradation in the linear model could be seen during the first six hours of the forecast. Previous studies have shown that the choice between different non-linear models is not critical [12].

TABLE III
SUMMARY OF THE RESULTS FOR DIFFERENT STRATEGIES BEING EMPLOYED DURING THE MARKET SIMULATION EXPERIMENT.

	Persistence	Control mbr	Ensemble mean	OQS, r_{-1d}	OQS, r_o	Perfect forecast
Contracted energy (MWh)	968.4	933.6	932.6	975.8	985.8	965.4
Surplus (MWh)	233.2	148.6	148.0	124.6	144.6	0.0
Shortage (MWh)	236.2	116.8	115.1	135.0	165.0	0.0
Total revenue (10^3 €)	27.76	28.40	28.39	28.51	28.91	29.24
Down-regulation costs (€)	922.0	555.3	544.5	433.8	229.8	0.0
Up-regulation costs (€)	560.7	282.9	302.5	297.5	97.1	0.0
Avg. down-regulation unit cost (€/MWh)	3.57	3.30	3.23	3.02	1.82	0.0
Avg. up-regulation unit cost (€/MWh)	1.68	2.07	2.08	1.88	0.99	0.0
Avg. regulation unit cost (€/MWh)	2.66	2.74	2.71	2.45	1.39	0.0
Avg. energy price (€/MWh)	30.29	29.42	29.41	29.53	29.95	28.75
Performance ratio (%)	94.93	97.13	97.10	97.26	98.88	100.0
Imbalanced part of production (%)	48.62	27.50	27.25	26.89	32.06	0.0



(a) Elspot prices, May 11 2017



(b) Power forecast, May 11 2017

Fig. 5. Example from May 11:th 2017. In (a) the hourly Elspot price in the SE3 area (solid green) along with down (solid blue) and up (solid red) regulation prices and regulation unit costs for positive (dashed red) and negative (dashed blue) imbalances. The ratio between the unit regulation cost for positive imbalance and the sum of the unit regulation costs for positive and negative imbalances is shown with a solid black line. In (b) the measured power production (yellow) and bid based on the optimal quantile strategy using actual hourly values of r (cyan) along with forecasted quantiles for the power production from the NWP ensemble system. All forecasts are issued at 00 UTC.

Hence, switching from an ANN to another non-linear black box model will probably not result in significantly better performance. However, better performance could perhaps be achieved if a recurrent structure is tried, exploiting the correlation in time between the forecasted values.

The input consisted of information about measured net production and net consumption from the previous day along with an NWP forecast for the next 24 hours. In a real situation the forecast has to be available well before the Elspot market closes at noon on the day prior to delivery. Hence the forecast horizon needs to be stretched to 42 hours using the forecasts started from 06 UTC. Such considerations should be the subject of further studies. What to include in the input vector is another question for further investigations. Here we picked information we thought was reasonable. No evaluation was made regarding how useful different input parameters were for the prediction. The reason why the linear model works well could be that the task is to forecast total power production over a quite large area. In this case the response to the input should become smoother than if the forecast would refer to an individual site and maybe this renders the problem more linear. It would also be interesting to see if switching NWP data from MEPS, with its 10 members at 2.5 km grid resolution, to the 50 member IFS ensemble from ECMWF at 18 km resolution, affects the performance of the power forecast. A previous study has shown that the ensemble mean from MEPS performs better than that of IFS when it comes to predicting GHI and DNI over Sweden, [13]. Such a study should ideally be carried out over a longer time period than three months.

The ANN model for power prediction was then used in a market simulation experiment with the purpose to evaluate the OQS strategy, based on a probabilistic view and ensemble forecasts, and investigate its effect on the resulting imbalances. The simulation was run for the period April - June 2017 using power measurements from the regional network described above along with electricity price and regulation cost data for the Elspot SE3 area. The solar power producer was assumed to be a price-taker, without market influence, participating in the market with the aim of maximizing the expected revenue. Using the OQS, either with a perfect or a persistence forecast of the regulation costs, resulted in increased revenues compared to a deterministic strategy based on the ensemble mean. However, the use of OQS also resulted in more imbalanced energy being produced.

Even if the previous study by Pinson et. al [1] on using OQS for trading wind power was made for a different time period and a different Nord Pool area it could still be interesting to compare some figures. In case of wind energy the use of a good deterministic forecast reduced the imbalanced part from 74 % to 41 %. For the solar case the corresponding drop was from 49 % to 27 %. The reason for the difference in offset is probably that the diurnal structure of the solar resource makes it easier to predict. The corresponding figures for the increase in performance ratio are 79 % to 87 % for wind and from 95 % to 97 % for the solar case. Turning to the probabilistic case, the imbalanced part then increased to 55 % for wind (r from last quarter) and to 32 % for solar (perfect forecast of r). In this case the performance ratio increased to 92 % and 99 % respectively.

With a good prediction of the relation between prices for up- and down-regulation (the percentage point) a trader using the OQS strategy will generate imbalances that are in line with the need of the TSO, since this relation reflects the sensitive of the TSO to the sign of the system imbalance. However, this result is only true as long as the actors using this strategy do not affect the market. If a vast majority of the wind and solar energy producers would employ this same strategy it might no longer be optimal, given that their share of the production is non-negligible. Their prediction of regulation prices may suffer from the same systematic errors and hence result in all producers opting for the same sign for the imbalance. In such a situation the use of the OQS strategy may instead act destabilizing on the power system.

To conclude, we have shown that a linear or ANN model can be used to predict the total net power production in a regional network on the 100 km scale with a normalized RMSE of about 6 %. Moreover, such a model can be used in conjunction with an optimal quantile trading strategy and NWP ensemble data to increase the revenue with about 2 % (if actual regulation prices are used) compared to a deterministic strategy based on the control member or the ensemble mean. This comes at the expense of increased imbalances compared to trading with a deterministic power forecast. The effect ranges from negligible, when assuming persistence for the regulation prices, to an increase in the amount of imbalanced production of almost 20 % if the actual regulation prices can be predicted.

ACKNOWLEDGMENT

The authors greatly acknowledge the Swedish Energy Agency that funded this work via grant 43231-1 as part of the national research program “El från solen”. Our colleague Magnus Lindskog at SMHI is also acknowledged for temporary archiving all the MEPS members during April - June 2017. Nord Pool is acknowledged for making their historical data available on-line.

REFERENCES

[1] P. Pinson, C. Chevallier, and G. N. Kariniotakis, “Trading wind generation from short-term probabilistic forecasts of wind power,” *IEEE Transactions on Power Systems*, vol. 22, no. 3, pp. 1148–1156, Aug 2007.

[2] M. Zugno, T. Jnsson, and P. Pinson, “Trading wind energy on the basis of probabilistic forecasts both of wind generation and of market quantities,” *Wind Energy*, vol. 16, no. 6, pp. 909–926, 2012. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/we.1531

[3] L. Bengtsson, U. Andrae, T. Aspelien, Y. Batrak, J. Calvo, W. de Rooy, E. Gleeson, B. Hansen-Sass, M. Homleid, M. Hortal, K.-I. Ivarsson, G. Lenderink, S. Niemelä, K. P. Nielsen, J. Onvlee, L. Rontu, P. Samuelsson, D. S. Munoz, A. Subias, S. Tijm, V. Toll, X. Yang, and M. O. Koltzow, “The HARMONIE-AROME Model Configuration in the ALADIN-HIRLAM NWP System,” *Monthly Weather Review*, vol. 145, no. 5, pp. 1919–1935, 2017.

[4] M. Mller, M. Homleid, K.-I. Ivarsson, M. A. . Kltzow, M. Lindskog, K. H. Midtb, U. Andrae, T. Aspelien, L. Berggren, D. Bjrge, P. Dahlgren, J. Kristiansen, R. Randriamampianina, M. Ridal, and O. Vignes, “Arome-metcoop: A nordic convective-scale operational weather prediction model,” *Weather and Forecasting*, vol. 32, no. 2, pp. 609–627, 2017. [Online]. Available: https://doi.org/10.1175/WAF-D-16-0099.1

[5] J.-J. Morcrette, H. W. Barker, J. N. S. Cole, M. J. Iacono, and R. Pincus, “Impact of a New Radiation Package, McRad, in the ECMWF Integrated Forecasting System,” *Monthly Weather Review*, vol. 136, no. 12, pp. 4773–4798, 2008.

[6] E. J. Mlawer, S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, “Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave,” *Journal of Geophysical Research: Atmospheres*, vol. 102, no. D14, pp. 16 663–16 682, 1997.

[7] J. A. Garca-Moya, A. Callado, P. Escriba, and C. Santos, “SLAF implementation in HarmonEPS: First results,” in *ALADIN-HIRLAM The 25th Workshop All Staff Meeting*, April 2015. [Online]. Available: http://www.umr-cnrm.fr/aladin/IMG/pdf/slaf.pdf

[8] T. Landelius, S. Andersson, and R. Abrahamsson, “Modelling and Forecasting PV Production in the Absence of Behind-the-Meeter Measurements,” in *35th EU PVSEC*, September 2018.

[9] S. A. Kalogirou, “Artificial neural networks in renewable energy systems applications: a review,” *Renewable and Sustainable Energy Reviews*, vol. 5, no. 4, pp. 373 – 401, 2001.

[10] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, “TensorFlow: Large-scale machine learning on heterogeneous systems,” 2015, software available from tensorflow.org. [Online]. Available: https://www.tensorflow.org/

[11] J. W. Pratt, H. Raiffa, and R. Schlaifer, *Introduction to Statistical Decision Theory*. The MIT Press, 1995.

[12] L. Martn, L. F. Zarzalejo, J. Polo, A. Navarro, R. Marchante, and M. Cony, “Prediction of global solar irradiance based on time series analysis: Application to solar thermal power plants energy production planning,” *Solar Energy*, vol. 84, no. 10, pp. 1772 – 1781, 2010.

[13] T. Landelius, M. Lindskog, H. Körnich, and S. Andersson, “Short-range solar radiation forecasts over sweden,” *Advances in Science and Research*, vol. 15, pp. 39–44, 2018. [Online]. Available: https://www.adv-sci-res.net/15/39/2018/