32nd European Modeling & Simulation Symposium
17th International Multidisciplinary Modeling & Simulation Multiconference

ISSN 2724-0029 ISBN 978-88-85741-44-7 © 2020 The Authors.
DOI: 10.46354/i3m.2020.emss.045

Trading Electricity Markets Using Neural Networks

Laura Pozzetti[1] and John Cartlidge[1,*]

[1]Department of Computer Science, University of Bristol, Bristol, UK

[*]Corresponding author. Email address: john.cartlidge@bristol.ac.uk

Abstract

We tackle the problem of developing an automated trading strategy to profit in the British intraday continuous electricity markets. We first train a feedforward neural network to predict one hour ahead total electricity transmission system demand. In live testing to ensure no look-ahead bias, we present forecast results with accuracy better than National Grid's own demand forecasts. We then train a second feedforward neural network, using our demand forecast as an input to the network, to predict one hour ahead net imbalance volume (NIV), and use this predicted NIV as a trading signal to buy and sell 30-minute electricity contracts. In live testing, between 09 March and 22 March 2020, the trading algorithm made 599 simulated trades, with 431 trades returning a profit (an accuracy of 72%). These results demonstrate the potential of neural network driven automated trading strategies to make significant risk-adjusted excess returns (i.e., profits) in the intraday electricity markets.

Keywords: Algorithmic trading; energy trading; forecasting demand; forecasting imbalance volume; intraday markets

1. Introduction

In March 2001, *New Electricity Trading Arrangements* (NETA) were introduced in England and Wales to create a market where electricity is traded like any other commodity. In the following year, market efficiencies caused wholesale prices to fall by 20%, while spreads between buy and sell prices reduced dramatically (from £70/MWh to £17/MWh), making it much cheaper for electricity generators (particularly those less able to predict output, e.g., wind farms) to balance their positions (Ofgem, 2002).

On 05 November 2015, the *Balancing and Settlement Code* (BSC) modification P305 introduced a number of changes to the calculation of *energy imbalance* prices to improve the ability for electricity markets to better balance supply and demand (Ofgem, 2015, 2018). As a result of P305, there has been a significant rise in the number of speculators looking to profit on short-term price changes (Tribe, 2017, p.15), leading to a rapid growth of market trading volumes (EPEX, 2019).

Electricity markets in Great Britain (GB) are now mature trading venues where considerable profits can be made from correctly predicting energy imbalances. During any given period, if trader T predicts that there is going to be excess power generated (a *long* period), then there is an expectation that prices will *fall*. Therefore, T should sell power at the start of the period (at price p_s), and then buy it back from the National Grid at the imbalance price (p_i). If the prediction is correct, then $p_s > p_i$, giving T a profit of $p_s - p_i$. Conversely, if T predicts a shortage of power (a *short* period), then T should buy at the start of the period (p_b) and sell back at the imbalance price (p_i), giving a profit of $p_i - p_b$.

In this paper, we introduce feedforward neural network models for predicting intraday demand in GB electricity markets, using ELEXON's open-source *Balancing Mechanism Reporting Service* (BMRS) as data input. We demonstrate that our model forecasts electricity demand with significantly lower error than the demand forecast supplied by National Grid (the owners of the high-voltage electricity transmission network in England and Wales, and the company responsible for balancing system supply and demand). We next use our demand forecasts as an input to a second feedforward neural network model to predict *net imbalance volume* (NIV) one hour ahead. Finally, we present a simple directional algorithm to trade 30-minute contracts in the continuous intraday market, using our NIV forecasts. In live testing between 09 March 2020 and 22 March 2020 (making a prediction in real time, every 30 minutes, to ensure no look-ahead bias), the trading algorithm suggested 599 trades, and was correct on 431 occasions (i.e., 72% long/short trading accuracy). This is strong evidence that neural networks can be used to significantly improve electricity forecasts in GB markets; and the first published result demonstrating forecasts can be profitably traded in this market. This paper describes research originally performed by Pozzetti for her BSc thesis, which was supervised by Cartlidge. For further details, see Pozzetti (2020).

2. Background

In Great Britain (GB), electricity is transmitted from where it is produced to where it is needed through the *transmission system*, a network of high voltage electricity wires that extend across the entire country and nearby offshore waters. Electricity can't be cheaply stored in large amounts, so supply and demand must be matched at all times. National Grid is in charge of balancing supply and demand.

The transmission system supplies electricity to *distribution networks*, which are lower voltage networks that deliver electricity to local customers and households. The transmission system is connected to distribution networks through *grid supply points* (see Figure 1). The transmission system is also connected to neighbouring countries through physical links called *interconnectors*, which allow the transfer of electricity across borders. GB is currently connected to four countries: to France, Ireland, Belgium, and the Netherlands. There are also plans for interconnectors with Norway and Denmark. The sum of the individual demands from all the distribution networks and interconnectors makes up the total *transmission system demand*.

The electricity wires in the UK were built to transmit electricity at a frequency of around 50 Hz. A frequency that is too high or too low can damage the transmission infrastructure and cause a power blackout across the entire country. National Grid aims to always keep the frequency at 50 Hz, with an operational limit of ± 0.2 Hz and a legal limit of ± 0.5 Hz. National Grid is regulated by Ofgem, the Government's Office For Gas and Electricity Markets.

An increase in generation, or a decrease in demand, causes the frequency to *increase*. Conversely, a decrease in generation, or an increase in demand, causes the

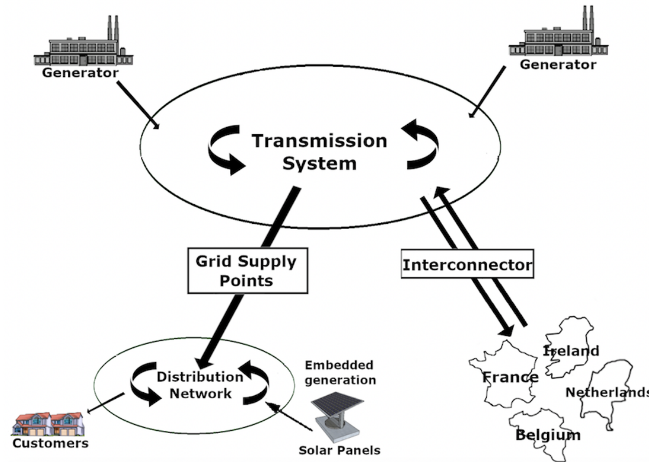

Figure 1. The transmission system (TS) delivers electricity to local distribution networks, which supply electricity to individual customers. The TS connects to foreign countries through interconnectors.

frequency to *decrease*. To regulate the frequency of the transmission wires it is necessary to balance electricity demand and supply at all times. In order to balance the system, National Grid is in constant communication with generators and consumers connected to the transmission system. All entities follow a set of regulations called the *Balancing and Settlement Code* (BSC), which is administered by ELEXON.

The electricity market is divided into half-hour periods, called *settlement periods*. Each day has 48 settlement periods, starting at midnight. To balance the system, each settlement period National Grid needs to know what generators intend to generate and what consumers intend to consume. National Grid needs this information before the start of the settlement period, so that it can understand the transmission system imbalance, plan how to balance it, and take balancing actions.

Generators and suppliers submit their planned generation or consumption—called *physical notifications* (PNs)—to National Grid for each settlement period. Positive PNs mean power generation, while negative PNs mean power consumption. One hour before the start of each settlement period (called *gate closure*) the PNs of parties are frozen. At this point PNs become final, and are called *final physical notifications* (FPNs). Parties must try to adhere to submitted FPNs and should only deviate from them at the instruction of National Grid.

Along with their intended generation/consumption, parties must submit their maximum possible generation/consumption level, called *maximum export/import levels* (MELs/MILs). MELs can be submitted at any time, even past gate closure, to inform National Grid of a sudden trip (i.e., when MEL is *lower* than the FPN).

Finally, parties have to submit notices to say how much it would cost for them to deviate from their final

Table 1. BMRS Data Description

ID	Feature	Comments
1	Month	Values: 1–12
2	Day of Week	Values: 1–7
3	Settlement Period	Half-hour intervals. Values: 1–48 (50 when clocks change due to daylight saving).
4	Final Physical Notification (FPN)	Intended net generation of all power generators connected to the transmission system.
5	Wind FPN	Intended net generation of UK wind farms only.
6	Interconnector FPN	Net import/export from international markets.
7	Generator Trips	Estimated energy not produced due to breakdowns (where MEL<FPN).
8	Wind Forecast	National Grid's forecast of energy produced from wind farms.
9	Intraday Solar Forecast	National Grid's forecast of energy produced from solar panels.
10	Temperature Forecast	National Grid's forecast of average UK temperature.
11	Day-ahead Demand Forecast	National Grid's transition system demand forecast a day in advance.
12	Day Demand Forecast	National Grid's forecast of transmission system demand 15 minutes before next period starts.
13	Rolling System Demand	Actual metered transmission system demand, calculated at end of period.
14	Net Imbalance Volume (NIV)	Net sum of all balancing actions taken by National Grid during period.
15	Market Index Price	Average price of all half-hourly contracts traded during settlement period.
16	Imbalance Price	Price used to settle the difference between contracted generation/consumption during period.

BMRS data available from ELEXON at: `https://www.bmreports.com`

physical notification. These notices are called *bids* and *offers*, and are always submitted in pairs. An offer is a proposal to increase generation or reduce demand for a price. A bid is a proposal to reduce generation or increase demand for a price.

Following gate closure, National Grid is able to evaluate/predict the net imbalance of the transmission system. National Grid does this by assessing the FPNs of the generators and suppliers and compares that assessment to its own demand forecasts for the settlement period. National Grid then assesses all the bids and offers for the settlement period and chooses the ones that best satisfy the balancing requirements of the transmission system. When a bid or offer is chosen, it is referred to as a *bid/offer acceptance* (BOA). This process is called the *balancing mechanism*.

The *net imbalance volume* (NIV) of a settlement period, together with a number of other factors, will determine the so-called *imbalance price* of that period. Imbalance price is *high* in a short period (as more generation is required) and *low* in a long period (as a reduction in generation is required). Parties who have under-generated or over-consumed compared to their FPNs, will have to *buy* that shortfall of energy at the imbalance price. Parties who have over-generated or under-consumed will have to *sell* that extra energy at the imbalance price.

This system allows speculative traders to buy and sell power contracts without producing or consuming any energy. For example, trader T can buy 50 MW of power for a settlement period without consuming any energy. The contracted amount for that period is 25 MWh and the consumed amount is 0 MWh. Therefore, T is required to sell that energy back to National Grid at the imbalance price. The *spread* (i.e., the difference) between the market price and the imbalance price determines trader T's profit or loss.

Speculative traders attempt to predict the energy imbalance of a future settlement period. If trader T predicts that there is going to be excess power (a long period), T should sell power for the period, and then buy it back from national grid at the imbalance price. If the prediction is correct, the imbalance price will likely be lower than T sold it for, so T can buy back at a profit. This benefits traders, but is also beneficial for National Grid, as it is cheaper to buy power at the last minute from T than it would be to buy power from a generator. Therefore, these rules incentivise free market participants to balance the system, making electricity prices more stable. This ensures against extreme prices during peak demand times, guarantees lower electricity bills for households, and protects UK consumers.

In Great Britain, physical power can be traded in multiple markets. Here, we consider only 30-minute contracts in the *intraday continuous market*. Contracts are for same day delivery, and *market close* (the last time that contracts can be traded) is 15 minutes before a settlement period begins. In the following sections, we attempt to build a model for predicting energy imbalances, and we use that prediction to simulate algorithmic trading in the intraday continuous market.

3. Data

ELEXON provide publicly available data relating to the GB electricity transmission system through the *Balancing Mechanism Reporting Service* (BMRS), a platform used extensively by market participants to help make trading decisions. BMRS data can be viewed online and can be accessed programmatically through the BMRS APIs. We downloaded BMRS data relating to the generation, demand, and balancing mechanism for every half-hour settlement period between January 2019 and February 2020, inclusive. All files were processed programmatically, using a virtual machine on

Table 2. Demand Forecasting: Neural Network Inputs

ID	Feature	Period	Loss (MW)
1	Month	t	
2	Day of Week	t	
3	Settlement Period	t	12,386
6	Interconnector FPN	t	8,244
10	Temperature Forecast	t	8,081
8	Wind Forecast	t	6,483
9	Intraday Solar Forecast	t	5,873
13	Rolling System Demand	t-2	1,405
11	Day-ahead Demand Forecast	t	640
12	Day Demand Forecast	t-1	627
12	Day Demand Forecast	t-2	428

Output: *Predicted Demand*(t + 2); evaluated against *Rolling System Demand*(t + 2)
Final test loss after network hyperparameter tuning: 396 MW

Google Cloud. A single data table was created with 20352 rows, each corresponding to a settlement period, t, with columns containing the 16 features listed in Table 1. Data was split 80:20 for training and testing, respectively.

4. Forecasting Demand

We develop a model for forecasting total transmission system demand (henceforth, simply "demand") for each settlement period. By plotting demand data (not shown), daily trends become clear. Demand is lowest during the night and starts rising around 6 am; dips in the middle of the day before increasing to a peak around 6-7pm; then steadily decreases. Winter demand is significantly higher than summer demand for every settlement period. The transmission system demand is also affected by embedded generation: more generation from solar panels and wind turbines in the local distribution networks means lower demand from the transmission system. Finally, the transmission system demand is affected by the electricity flow through interconnectors. If electricity is flowing out of the country, this will increase the demand from the UK transmission system, acting as extra demand from abroad.

We used Google's TensorFlow to build a feedforward neural network to predict demand at $t+2$ (i.e., one hour ahead), which we trained using the adaptive moment estimation (Adam) optimiser. For initial exploration of input features, we trained a series of networks for 50 epochs, with loss function minimising absolute mean error between network output and the ground truth value of *Rolling System Demand*(t + 2). Table 2 presents the final set of eleven features selected as network inputs. Ordered from top to bottom, we see that test loss decreases as additional features are added, for example: networks with only three input features (month, day of week, and settlement period) produce a demand forecast with high test loss of 12,386 MW; the addition of interconnector FPN as a fourth feature reduces test loss to 8,244 MW; and networks with all eleven input features produce the lowest test loss of 428 MW.

To further improve performance, we next tuned the network hyperparameters, resulting in a best performing network with architecture containing: an input layer of 11 nodes (each corresponding to the input features listed in Table 2); two hidden layers of 50 nodes each; and an output layer of 1 node. Training was performed over 1000 epochs (a greater number of epochs demonstrated overfitting). We report a final training loss of 367 MW and testing loss of 396 MW.

4.1. Live Testing of Demand Forecast

When backtesting the performance of forecasts over time-series data, it is easy to incorporate look-ahead bias such that data that would not have been available at the time of the forecast is accidentally used as input into the model. Such bias can greatly inflate performance, as the model is effectively looking into the future. To guarantee no look-ahead bias, the best approach is to live test the system in real time, on current data, so that future information does not yet exist. To this end, we implemented our demand forecast model into a real-time system that downloads the latest data from the BMRS every 30 minutes, calculates the input variables (Table 2), and then prints the neural network's one hour ahead demand forecast. This is the same process that would be used in a live forecasting system.

We ran the software on Google Cloud from 09 March 2020 to 22 March 2020. Approximately 600 forecasts were made over the course of two weeks. Live forecasts were compared to the actual metered demand during each period, giving a mean absolute error (MAE) of 378 MW (equivalent to mean *relative* error of 1.2%). This value is similar to (and slightly better than) the MAE recorded during back-testing, therefore demonstrating that the system has no look-ahead bias.

To put this forecasting error of 378 MW into context, we compare it against results reported by Amira Technologies, a data science consultancy dedicated to the GB power sector. Their website states: "Amira Technologies produce market leading electricity demand and renewable generation forecasts for the GB power sector". In their report for Autumn 2019, they present a graph showing that Amira's transmission system demand forecast performance for one hour ahead has MAE 364 ± 10 (Amira Technologies, 2019). For the same period, they also report National Grid's BMRS one hour ahead forecast performance as having MAE 470 ± 10. While the comparison is not exact (Amira's reported errors are for forecasts during the last quarter of 2019; while our errors are from forecasts during March 2020), results indicate that our neural network demand forecast has performance better than National Grid, and similar performance to a leading commercial provider.

Outperforming National Grid's forecast is perhaps not as surprising as it may at first appear, since Na-

Table 3. NIV Forecasting: Neural Network Inputs

ID	Feature	Period	Loss (MWh)
1	Month	t	
2	Day of Week	t	
3	Settlement Period	t	
4	Final Physical Notification (FPN)	t	
11	Day-ahead Demand Forecast	t	274.90
*	(NeuralNet) *PredictedDemand*	t+2	249.15
14	Net Imbalance Volume (NIV)	t-2	217.34
7	Generator Trips	t	214.01
5	Wind FPN	t	
8	Wind Forecast	t	211.96

Output: *Predicted NIV*($t+2$); evaluated against *Net Imbalance Volume*($t+2$)
* *Predicted Demand*($t+2$) is generated at time t by network described in Section 4
Final test loss after network hyperparameter tuning: 210.75 MWh

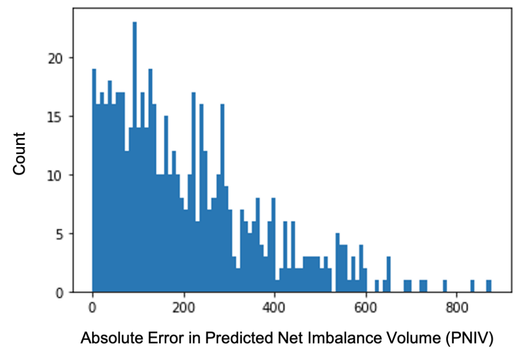

Figure 2. Distribution of NIV forecast errors during live testing.

tional Grid make extensive use of regression models for forecasting (Blaavand et al., 2017). These linear models are constructed using historical values for demand and various explanatory factors, such as weather variables. Model errors may arise because of changes in the general energy landscape (in particular, systemic increases in wind and photovoltaic energy generation witnessed in recent years), which might mean that the historical data on which the models are based is no longer representative of the current electricity system (Blaavand et al., 2017). To address these issues in forecast errors, National Grid is currently undertaking a multi-year project to improve forecasting ability by applying advanced statistical and machine learning modelling techniques, to be completed in March 2022 (National Grid ESO, 2020). We should therefore not be too surprised that a neural network model is able to improve upon a suite of linear regression models. In the next sections we demonstrate that forecasting with accuracy greater than National Grid can form a consistently profitable trading strategy.

5. Forecasting Net Imbalance Volume

In this section, we develop a neural network model for forecasting net imbalance volume (NIV) at $t+2$ (i.e., one hour ahead). The NIV is the net sum of all the balancing actions taken by National Grid during a settlement period. To predict the NIV, it is necessary to predict the difference between the demand and the independent generation from power plants (the generation that occurs without instruction from National Grid). This provides an estimate of the amount of balancing actions that National Grid needs to take to keep the system at a balanced frequency of 50 Hz.

We used Google's TensorFlow to build a feedforward neural network to predict net imbalance volume. Input features were explored using networks containing one input node corresponding to each input feature, two hidden layers of 20 nodes each, and one output node corresponding to forecast net imbalance volume at $t+2$. Networks were trained for 50 epochs, with loss function minimising absolute mean error between output and the ground truth value *Net Imbalance Volume*($t+2$). Table 3 presents the final set of ten features selected as network inputs. Ordered from top to bottom, we see that test loss decreases as additional features are added, for example: 5 input features (month, day of week, settlement period, FPN, and day-ahead demand forecast) have test loss 274.90 MWh; 6 input features (the previous five inputs, plus *Predicted Demand*($t+2$), generated at time t by the optimised neural network described in Section 4), reduces test loss to 249.15 MWh; while using all 10 input features listed in Table 3 results in the lowest test loss of 211.96 MWh.

To further improve performance, we next tuned network hyperparameters, resulting in a best performing network with architecture containing: an input layer of 10 nodes (corresponding to the features presented in Table 3); two hidden layers of 50 nodes each; and an output layer of one node. Training was performed over 200 epochs (with more epochs, overfitting was observed). We report a final training loss of 179.75 MWh and testing loss of 210.75 MWh.

5.1. Live Testing of NIV Forecast

To avoid the charge of look-ahead bias, we again live tested the NIV forecast between 09 March and 22 March 2020, using the same procedure described in Section 4.1. The network produced approximately 600 live NIV forecasts, giving a mean absolute error of 207.54 MWh (equivalent to mean *relative* error of 10.03%), with the majority of forecasts having error value below the mean (see error distribution presented in Figure 2).

The performance presented is difficult to contextualise since, as far as the authors are aware, there are no published one hour ahead NIV forecasts with which to directly compare. An influential literature review by Weron (2014) on electricity price forecasting makes no mention of imbalance volume predictions. The nearest related work is by Garcia and Kirschen (2006), who forecast one month ahead and one week ahead NIV predictions for the purpose of trading in the *forwards* *market*. The best mean error of their one week ahead forecast is 514 MWh; however a direct comparison is unfair, since the forecast horizon of one week is much longer than our one hour horizon.

6. Trading the Electricity Market

Speculative traders in the electricity market buy and sell power without ever consuming or generating it. Their profit comes from the difference between the market price at which they buy (or sell) power and the imbalance price at which they sell (or buy it back).

For *short* periods—where there is a *shortage* of generation—in the vast majority of cases the imbalance price is *higher* than the average spot price; and for *long* periods—where there is an *excess* of generation—the imbalance price is *lower* than the average spot price. Using BMRS data for 2019, we observed that almost 96% of short periods had an imbalance price higher than the market index price; and over 91% of long periods had an imbalance price lower than the market index price. Therefore, in general, power traders aim to buy power at spot price if they predict a short period (in order to sell it back for a profit at imbalance price); and if they predict a long period, they aim to sell power at spot price (in order to buy it back at a profit for less at imbalance price).

Here, we introduce a minimal trading algorithm that uses *predicted net imbalance volume* (PNIV) to trade:

If *PNIV* > 0 **then** *Buy*; **else if** *PNIV* < 0 **then** *Sell*; **else** *Do Nothing*.

The trading algorithm is deliberately simple, to ensure that trading profits directly correlate with the quality of PNIV. If a short period is forecast (i.e., $PNIV > 0$) the trader buys at spot price and sells back at imbalance price. If a long period is forecast (i.e., $PNIV < 0$) the trader sells at spot price and buys back at imbalance price. If the forecast is balanced (i.e., $PNIV = 0$) the trader does not trade during the period.

We do not have access to the prices in the spot market. Spot prices are determined by the bids and offers that market participants continuously submit to the live market and are not published anywhere. The only way to access them is to have direct access to the market itself, which would require us to set up a company, complete an admission process, sign an agreement with the European Commodity Clearing House (ECC), and deposit trading capital into the account. This is not possible. Therefore, we evaluate the performance of the trading algorithm by assuming that the trader bought/sold power at the market index price, which is the average price of all half-hourly traded contracts during a settlement period. This is a reasonable assumption: on average we expect that we can buy/sell for the average price during the period. Over a large number of settlement periods, we expect this to be true.

We backtested the algorithm for all settlement periods during the calendar year 2019. In total, the algorithm traded 16,294 times (i.e., once every settlement period); with 70.8% of trades returning a profit (i.e., the decision to *Buy* or *Sell* was correct 70.8% of the time). *Buy* decisions were correct 70.4% of the time and *Sell* decisions were correct 71.0% of the time, suggesting that the PNIV forecast is unbiased, with equal accuracies in both directions. During 2019, approximately 45% of periods were short, while approximately 55% of periods were long. The trading algorithm chose to *Buy* 42% of the time and *Sell* 58% of the time , demonstrating that the PNIV forecast direction closely aligns with the real market imbalance.

To simplify profit calculations, we assume that each trade has size 2 MW (i.e., 1 MWh), which is a very small trade size relative to the market, and therefore guarantees that a trade of this size can be executed. By the end of 2019, the trading algorithm generated a simulated profit of £73,407. Given that contracts cost in the region of £50/MWh, this profit demonstrates exceptional potential returns on minimal capital investment.

6.1. Live Trading the Electricity Market

As before, to ensure no look-ahead bias, we tested the trading algorithm live between 09 March 2020 and 22 March 2020, making trading decisions every 30 minutes, before the market close for each settlement period. During this period, the algorithm made 599 trades, with 431 resulting in profit and 168 resulting in loss; giving an overall accuracy of 72%. *Buy* decisions were correct 73.6% of the time and *Sell* decisions were correct 70.6% of the time. These directional accuracies are relatively close, but may suggest that there is a small bias towards *Buy*, indicating that the PNIV may slightly underestimate the true NIV.

To put this trading accuracy into context, let us consider a random trader that issues *Buy* decisions 45% of the time and *Sell* decisions 55% of the time (i.e., in the same ratio as the short/long periods exhibited in the market during 2019). Then, the chance of this random trader making profit in at least 400 of 599 trading periods is less than 1 in 100 million. This clearly suggests that profiting 431 times out of 599 trades is extremely unlikely to have occurred by chance.

During two weeks of live trading, assuming 1 MWh is traded at market index price for each settlement period, the trading algorithm generated a simulated profit of £2,858. The Sharpe ratio (SR) is a common measure used to evaluate the risk-adjusted performance of an investment. It is defined as the difference between the returns of the investment R_p and the risk-free return R_f (we assume a 3% annual risk-free rate of return), divided by the standard deviation of the investment returns $\sigma(p)$, i.e., $SR = (R_p - R_f)/\sigma(p)$. For live trading, we report a daily Sharpe ratio of 26.1, demonstrating an exceptionally high risk-adjusted return on investment.

7. Discussion & Conclusions

We have used publicly available open-source data to train a feedforward neural network to predict total transmission system demand in Great Britain's electricity network (Section 4). This one hour ahead prediction is more accurate than the National Grid's own forecast, and similar to the forecasting accuracy of Amira Technologies, one of the industry's best known commercial consultancies. To guarantee that the prediction accuracy is robust and not affected by look-ahead bias, we generated live forecasts during two weeks in early March 2020, and demonstrated that accuracy is unchanged (Section 4.1).

We used our demand prediction as input to a second feedforward neural network to predict net imbalance volume (NIV) in the electricity system (Section 5). Using live testing, once again, we demonstrated that the model has high prediction accuracy (Section 5.1). Since there are no metrics for direct comparative evaluation of the performance of our NIV prediction, we used our prediction as a signal for a simple automated trading strategy (Section 6) which *Buys* a contract when predicted $NIV > 0$ (i.e., when there is a predicted *shortage* of electricity), in the anticipation that prices will rise, and *Sells* a contract when predicted $NIV < 0$ (i.e., when there is a predicted *excess* of electricity), in the anticipation that prices will fall. During live testing, the trading algorithm traded 599 times and was correct (i.e., made a profit) 431 times, giving a success rate of 72% (Section 6.1). This trading performance demonstrates the quality of the NIV prediction, and has the potential to translate to significant and reliable risk-adjusted returns, if traded in the real electricity markets.

The attentive reader may have observed that the live prediction and trading periods we have used ended on 22 March 2020. On 23 March 2020, the UK went into "lockdown" due to the COVID-19 pandemic, with the majority of the population (other than select "key workers") forced to stay at home apart from one form of daily exercise. Across the country, schools, shops, factories, and offices were either closed or put under severe restrictions. As a result of this unprecedented and unforeseeable event, electricity demand in the British transmission system was significantly impacted. In particular, transmission system demand was significantly reduced during weekdays, making every day look like a weekend (Wilson et al., 2020). For this reason, we ignore testing results from 23 March onwards. However, we have since retrained the neural networks for demand prediction and NIV prediction during lockdown and results are consistent with those presented here; as are the results of trading.

In future, we aim to extend this work by optimising the model retraining period and exploring forecasting time horizons greater than one hour ahead. We will also explore alternative deep learning neural network models, such as LSTM/RNN, which are designed for time series data. Finally, we plan to explore more sophisticated algorithmic trading strategies, including risk hedging and portfolio optimisation across multiple forecast horizons.

Acknowledgements

John Cartlidge is sponsored by Refinitiv.

References

Amira Technologies (2019). Forecast performance report: Autumn 2019. Online https://www.amira-technologies.co.uk/pdf/forecast_report_latest.pdf.

Blaavand, J., Warrington, R., and Leese, R. (2017). Demand forecasting and proposed incentives. Technical report, Smith Institute. Online https://www.ofgem.gov.uk/ofgem-publications/111128.

EPEX (2019). New monthly record on EPEX intraday: Volumes on British, French, and German continuous markets climb to new record high. Online: http://static.epexspot.com/document/41140/2019-11-2005_EPEXSPOT_PowerTradingResults2010-2019.pdf.

Garcia, M. P. and Kirschen, D. S. (2006). Forecasting system imbalance volumes in competitive electricity markets. *IEEE Transactions on Power Systems*, 21(1):240–248.

National Grid ESO (2020). Platform for energy forecasting (PEF) strategic project roadmap update June 2020. Online https://data.nationalgrideso.com/plans-reports-analysis/platform-for-energy-forecasting-pef-strategic-project-roadmap.

Ofgem (2002). New electricity trading arrangements (NETA) - one year review. Online: https://www.ofgem.gov.uk/ofgem-publications/76199/1109-factsheet110224july.pdf.

Ofgem (2015). Balancing and settlement code (BSC) P305: Electricity balancing significant code review developments. Online: https://www.ofgem.gov.uk/sites/default/files/docs/2015/04/p305d_1.pdf.

Ofgem (2018). Analysis of the first phase of the electricity balancing significant code review. Online https://www.ofgem.gov.uk/system/files/docs/2018/08/analysis_of_the_first_phase_of_the_electricity_balancing_significant_code_review_as_final_version_publication.pdf.

Pozzetti, L. (2020). Algorithmic trading in UK power markets using neural networks. BSc thesis, Department of Computer Science, University of Bristol.

Tribe, E. (2017). P305 post implementation review. ELEXON Online: https://www.elexon.co.uk/wp-content/uploads/2017/02/P305-Post-Implementation-Review.pdf.

Weron, R. (2014). Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting*, 30(4):1030–1081.

Wilson, G., Godfrey, N., Sharma, S., and Bassett, T.

(2020). UK's COVID-19 restrictions bring about historical lows in energy demand. Channel News Asia. Online: https://www.channelnewsasia.com/news/commentary/covid-19-coronavirus-uk-britain-electricity-grid-fuel-demand-12598220.