Price Forecast Methodologies Comparison for Microgrid Control with Multi-Agent Systems

Marcos Eduardo Cruz Victorio, Behzad Kazemtabrizi, Mahmoud Shahbazi
Department of Engineering
Durham University
Durham, UK
marcos.e.cruz-victorio@durham.ac.uk

***Abstract*— Multi-Agent systems offer a way to control distributed generation in microgrids, reliability and cost minimisation capabilities can be improved by price forecast methodologies that can be deployed without the need of external control signals. This paper presents and compares two suitable electricity price forecast methodologies for use in distributed control of Microgrids' resources using Multi-Agents: Markov Chain Monte Carlo simulations with heuristic and numerical optimisation and price prediction with Non-linear Auto Regressive Artificial Neural Networks with different internal architectures. The methods are evaluated using MAPE and RMSE functions for the UK electricity market data. It was found that the proposed heuristic model has less error than the Neural Networks only when the price data contains outliers.**

Keywords — *AC Microgrid, Artificial Neural Network, Autoregression, Multi-Agent System, Price Forecast.*

I. Introduction

As the energy generation paradigm shifts from centralised to distributed, the control systems also shift to reflect this new nature. Control systems for Distributed Energy Resources (DER) are divided in levels depending on the control objectives and control speed.

For the case of the control within the microgrids, the control is divided in the primary control which directly regulates voltage and frequency of an individual DER, and the secondary control which coordinates the power schedule for the DERs and power flow from and to the grid, by sending the control references to the primary controllers, a tertiary level can be used to merge the optimisation of several microgrids and their interaction with the main grid. Such levels form a hierarchical framework for coordinating and controlling several Microgrids at once [1].

While the primary controller is realised by conventional PI controllers, there are centralised, decentralised and distributed secondary controllers [1]. Multi-Agent Systems (MAS) have been applied as an appropriate distributed secondary control method to coordinate DER units, to minimise power losses or maximise their economic benefit as they can combine elements of centralised and decentralised control, benefiting from separating control tasks in different agents, for asynchronous and parallel operation [2, 3].

This work is supported by the CONACYT-SENER energy sustainability programme with scholarship number 447877 and EU TESTBED Grant Agreement with number 734325.

Compared to larger power plants, the formulation of the microgrid cost minimisation problem can be simplified given the scale and dynamics of kilowatts of electric power. For example, cost minimisation models for grid-connected microgrids can neglect shutdown cost, minimum up and down time, and ramping rates, focusing instead on start up cost and operating cost for DER and the dynamic behaviour of the State of Charge (SOC) of the batteries [4]. To minimise costs from the battery, energy arbitrage has been proposed for the UK, as well as grid balancing for renewables [5]. This simplified sub problems can be solved by individual agents in a MAS control for optimal power flows [6].

To maximise the economic benefits of distributed generation and control, considering the starting cost for the DERs and limited storage capacity, price prediction is required to generate a proper power schedule to minimise the supply cost of cooperative DER owners. In previous MAS control systems, such as in [7], The MAS control depends on an external price prediction signal, which creates dependence of the agents to the source of such signal, to prevent this, the price prediction must also be distributed. There are two forecasting methods suitable for distributed online control, as they have low computation requirements for deployment.

The first family of methods for developing such forecast models are Auto-Regressive (AR) models, which are a forecasting tool and have been used to estimate the grid price from historical data [8]. Price prediction requires some assumptions based on the observed data: Price is considered to be normally distributed for a specific hour [9], and a high correlation exists between an hour price, and the prices for the same hour for the previous day and seven days ago as the demand tends to follow these patterns [10]. Monte Carlo Markov Chain (MCMC) simulations, are used to model the Probability Density Function (PDF), from which an AR model can be developed and further optimised with heuristic solvers [11]. MCMC models have been applied in renewable generation [11, 12], as well for modelling load uncertainties in microgrid optimisation [13]. The Metropolis-Hasting [14] is used for its easy implementation for the MCMC.

The second family of methods are Artificial Neural Networks (ANN), which are computationally expensive for training, but not for execution. Examples of ANNs that have been used for time series forecasting are found in [15–17].

For the case of intra-day price prediction, the entire day values can be considered as a single vector or wavelet [16] to model features that the ANN can learn.

The Non-Linear Auto Regressive network (NARNET) is suitable for distributed control where the control does not have exogenous data available, in order words, no external control input, which makes it a good candidate for MAS application in a microgrid, as it maintains control distribution. Applications of NARNETs include prediction of number of EVs in a city and albedo forecasting [18].

The basic architecture of the NARNET can be modified to add layers in series or in parallel [15]. Parallel networks require more time to train, however, for the Australian market they show better results than other architectures [17].

The rest of the paper is organised as follows: Section II describes the Multi-Agent System used for distributed microgrid control, section III details the two main forecast methods, section IV describes the implementation of price forecast to the microgrid control applied to the test case described in section V. The results of the forecast are analysed in section VI and the conclusion is in section VII.

II. Multi-Agent System

Microgrid control is normally divided in a hierarchy, with a fast primary control to regulate voltage and frequency of a DER, and a slower secondary control that commands the primary control by providing the power references to it.

The secondary control in this case is based on multi-agent systems that needs to be as distributed and independent as possible. This prevents single points of failure and maximise the flexibility and adaptability of the control system, such that the distributed generation sources can change in number and power. The primary controller of each resource consists in a PI control with an inner close loop for voltage references and a outer loop for power references as described in [19]. Each agent in the system is a small artificial intelligence that is capable of communicating with other agents, and carries out a set of specific tasks called behaviours.

A MAS system has been developed for microgrid control for supply cost minimisation as described in [20], with 3 agents developed for the microgrid control: The Energy Storage System (ESS) agent, the DER agent and the Grid agent. The ESS and DER agents generate the power references for the primary control based on the price signal sent by their respective grid agent.

An heuristic model was developed for electricity price forecasting as part of the Grid Agent behaviour with four months of data [20]. The proposed forecast methodologies are used to model the price forecast under more diverse scenarios by increasing the data size to two years.

III. Forecasting Methodologies

In this work two main price forecast methodologies are tested and compared for UK price forecast. The first one is an AR model built with the MCMC method [20],which is further optimised with numerical and heuristic solvers, referred as Weighted Average (WA) in this work. The second forecast method is based on NARNETs, analysing different architectures for the hidden layers and number of neurons. Both methodologies are directly comparable as they only use past prices for the price forecast.

A. Weighted Average Model

The MCMC method is used to estimate the parameters θ_f of a PDF P given a data set δ, in this case, estimate the mean $\overline{\pi}_x$ and the standard deviation σ_x, forming a chain Θ of j elements. Applying the Metropolis-Hasting algorithm, from an empirical starting point θ_0, and proposed parameters θ'_0, with the same value plus a random small deviation, selecting the next element as:

$$\Theta_{j+1} = \begin{cases} \theta'_j & Z < e^{(Log_{10}(a))} \\ \theta_j & otherwise \end{cases} \tag{1}$$

Where Z is a random variable with uniform distribution,$Z\mathcal{N}U$,and a is the acceptance ratio:

$$a = \frac{P(\delta/\theta')P(\theta')}{P(\delta/\theta)P(\theta)} \tag{2}$$

Where a is the acceptance, $P(\delta/\theta')$ and $P(\delta/\theta)$ are the likelihoods of θ' and θ, given the data δ, $P(\theta')$ and $P(\theta)$ are the priors on θ' and θ. The priors are the parameters evaluated in the normal distribution with respect to θ_0.

The chain continues to grow until the change in its average is below a given tolerance or maximum number of iterations, in this case 10^{-6} and 4000 to guarantee convergence, then the estimated parameter is calculated as:

$$[\overline{\pi}_x \quad \sigma_x]^T = \theta_f = \frac{1}{n_j}\sum \Theta_j, \forall j \in n_e \tag{3}$$

Where n_e is the number of elements j in the chain, and θ_f is the vector containing the parameters that describe the PDF of the electricity price of each price.

With the use of the MCMC, and considering the correlation of the day prices, a simple heuristic AR model is proposed as follows, putting a weight in each term for model optimisation:

$$\pi_i = w_1 p_{i-24} + w_2 p_{i-168} + w_3\overline{\pi}_a + w_4\overline{\pi}_b + w_5\overline{\pi}_c \tag{4}$$

Where p_{i-24} is the actual price for the previous day, p_{i-168} is the actual price for the previous week, $\overline{\pi_a}$ is the average price for the same day of the week, $\overline{\pi_b}$ the average price of the season, $\overline{\pi_c}$ is the average price of the entire data, and π_i is the forecast price at hour i. The averages $\overline{\pi}_x$ are estimated from the UK data set using the MCMC with Metropolis-Hastings method with the use of (3). Weights w_j are obtained by solving a least squares regression problem:

$$\min_W \sum (p_i - \pi_i(W))^2, \forall i \in m \tag{5a}$$

$$\text{s. t. } W \succcurlyeq 0 \tag{5b}$$

Where W is a vector containing the weights of (4) for the hours i in the set m. The problem is constrained to be semi-definite positive to reflect the positive correlation of

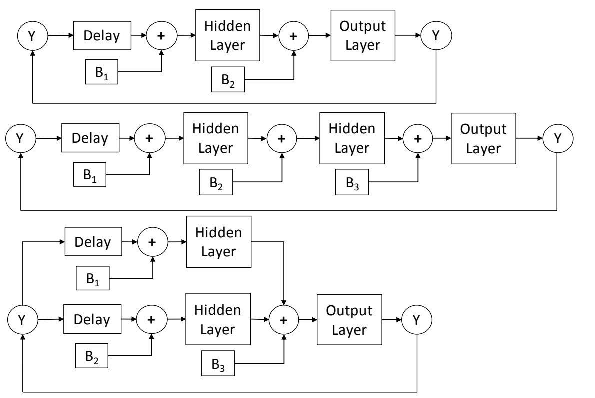

Fig. 1. Architectures used: a)Single layer, b) Series layers, c) Parallel layers

past prices with current prices and is solved using different methods: Interior point with the constraint, Quasi-Newton (QN) method without this constraint, and Genetic Algorithm for both cases.

The QN algorithm is based on gradient descent methods. At the start of the QN algorithm, a starting point x_0 and a H_0 as any symmetric positive definite matrix are chosen. With these, a direction d_k is calculated:

$$d_k = -H_k^{-}1 \cdot \nabla f(x_k) \tag{6}$$

An α_k is selected such that:

$$f(x_k + \alpha_k d_k) < f(x_k) \tag{7}$$

With the resulting α_k and d_k, the next iteration is calculated:

$$x_{k+1} = x_k + \alpha_k d_k \tag{8}$$

The difference from the Newton method is that the Hessian matrix update in its algorithm is estimated from its previous values instead of calculated analytically:

$$H_{k+1} = H_k + \frac{q_k q_k^T}{q_k^T s_k} - \frac{H_k s_k s_k^T H_k^T}{s_k^T H_k s_k} \tag{9}$$

Where:

$$s_k = x_{k+1} - x_k \tag{10}$$

$$q_k = \nabla f(x_{k+1}) - \nabla f(x_k) \tag{11}$$

The method finalises when the following is true for a given tolerance ϵ, otherwise, the algorithm repeats with the next iteration $k+1$:

$$|\nabla f(x_k)| < \epsilon \tag{12}$$

B. Non-linear Auto-Regressive Neural Network

The key difference between the NARNET and a conventional ANN is the use of the delay and feedback, to use the n value of a time series to obtain the $n+1$ value. The NARNET uses a combination of functions represented as blocks to form the network architecture. The three architectures analysed in this work are shown in Fig. 1. The main difference between the different architectures is the number and connection of the hidden layers inside the network. In all cases, the NARNET models a time dependant variable as a function of its past values, following the general equation [21]:

$$Y(t) = f(Y(t-1), ..., Y(t-d)) \tag{13}$$

Where d is the number of delays in the net, Y is vector of prices y for a time t, and f is the non linear model approximated by the ANN. The first block in a NARNET is the update of the delay vector D that contains the prices of each past day as its elements:

$$D(t) = [y_{t-1,1} \cdots y_{t-1,24} \cdots y_{t-d,1} \cdots y_{t-d,24}]^T \tag{14}$$

Where $y_{t,k}$ are the prices at hour k for a day t. As the NARNET operates, the values shift positions to the next time step, eliminating the oldest information first, while the vector D is completed with the network's feedback loop. The transfer function for each hidden layer is described by the logistic sigmoid function[22]:

$$\sigma(\omega_1 D) = \frac{1}{(1 + e^{-\omega_1 D})} \tag{15}$$

Where ω_1 is the matrix of weights, with a number of rows equal to the number of neurons in the layer. The open loop transfer function of a NARNET with a single hidden layer is:

$$f(Y(t)) = \omega_2 \sigma(\omega_1 D(t) + B_1) + B_2 \tag{16}$$

IV. IMPLEMENTATION

Regardless of the price forecast method, its execution needs to be compatible with the rest of the MAS control system, in that sense, the method has to be run from an agent and within 10 milliseconds, which is the set time an agent has to complete one cycle of their behaviours. To achieve this with the NARNET, the equivalent transfer function is obtained and applied as shown in Fig. 2.

The transfer function is programmed into the Grid agent, replacing the previous transfer function and the weights obtained from the training of the NARNET are transferred as CVS files to be read by the Grid agent, along with the

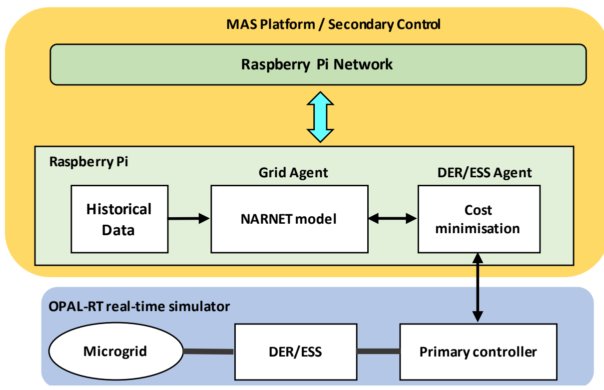

Fig. 2. Block diagram of the microgrid control system

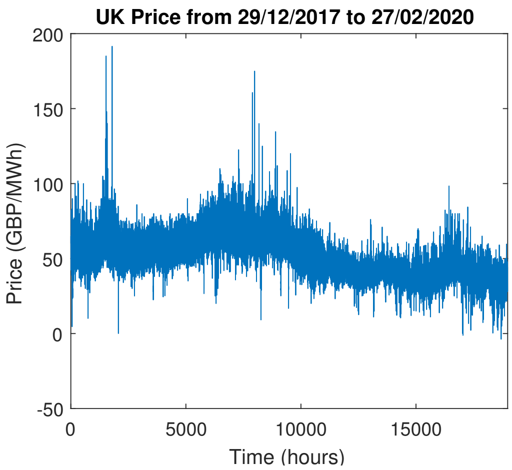

Fig. 3. UK price data

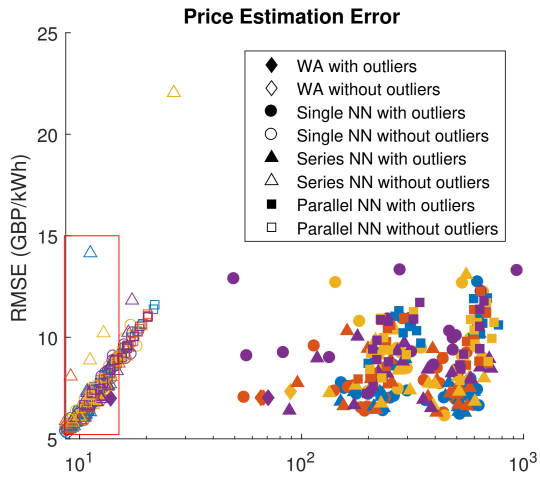

Fig. 4. Individual Results of each method and set of parameters

historical price data, to improve the simpler forecast method in [20], without affecting the response time of the system.

The MAS platform is realised in a network of one PC that connects to an OPAL-RT real-time simulator and two Raspberry Pi, each of them have an agent container, as well as a ring with the main container and the back-ups for further distribution of the control, as described in [20], where it was shown that the fault tolerance mechanism allows the microgrid control to remain operational even in the event of a computer being lost from the network.

V. Test Case

The UK price data used is from Noord Pool, for each hour over the period from the 29-12-2017 to 27-02-2020. During this time, price has sunk and spiked several times, and the monthly averages have decreased as seen in Fig. 3. The methods are tested for four data scenarios: a) the entire data set, b) the entire data set without the 20 most significant outlier days, c) weekdays only and d) weekdays only without the outliers. Each scenario was run 10 times for each configuration of the tested methods. For the case of the WA, the forecast is optimised with each solver. For the NARNET the forecast is tested with varying number of neurons in the hidden layers between 5 or 10 and the amount of delay in each of the architectures shown in Fig. 1 for one week of delay and two weeks.

The forecasts obtained are evaluated calculating the Mean Absolute Percentage Error (MAPE) and the Root Mean Squared Error (RMSE). Two evaluations functions are required to better dimension the accuracy of the methods, as the data contains values close to zero [16, 23]:

$$MAPE = \sum \frac{|p_i - \pi_i|}{p_i}, \forall i \in m \tag{17}$$

$$RMSE = \sqrt{\sum (p_i - \pi_i)^2}, \forall i \in m \tag{18}$$

Where p_i is the hour real price and π_i is the hour price from (4).

VI. Precision Analysis

After running each case 10 times for each set of parameters, the total MAPE and RMSE is calculated. The results obtained are plotted in Fig. 4. The results of each set of parameters show that the NARNET can outperform the WA when the outliers are not included. When the entire dataset is taken into account, the MAPE increases in two orders of magnitude.

To focus on the best solutions, a zoom is done to Fig. 4, marked as a red rectangle in Fig. 4 and the solvers re coloured as seen in Fig. 5. It can be seen that in terms of individual runs, the best solution is provided by the single layer with 5 neurons with a week of delay for the NARNET, and the QN solver for the WA method. The best WA solutions have weights $W = [0.490.320.1800]$ and $W = [0.480.340.960.03 - 0.82]$ for the constraint and unconstrained case. Several trials of the training are needed to achieve the best results for the NARNET, as the RMSE and MAPE output depends on the division of the data in 70% for training, 15 % for testing and 15 % validation sub sets, which are chosen at random.

The errors over the data set for each of these two best cases without the outliers and weekends for each method are shown in the Fig. 6 and Fig. 7. The individual errors are shown as blue circles deviating from the actual price in orange. While they show a similar level of correlation between the target and output, the WA deviates more at the higher prices, increasing the final RMSE and MAPE scores. In both cases the error increases for the lower values, however, this may be caused by the prices in the data set that tend to zero.

The summary of error distributions is shown in Fig. 8, which compares the QN solver for the WA and the NARNET using all the data and without the outliers. It can be seen that for both methods and data sets the errors have a normal distribution, and the difference in score depends on the number and deviation of large errors shown as blue circles.

Finally, Table I shows the average MAPE and RMSE of the combined results of the different parameters of each method and the best individual score. It can be seen that in general, the WA method is better as it has less variations in each run compared to the NARNET, but the latter has the best

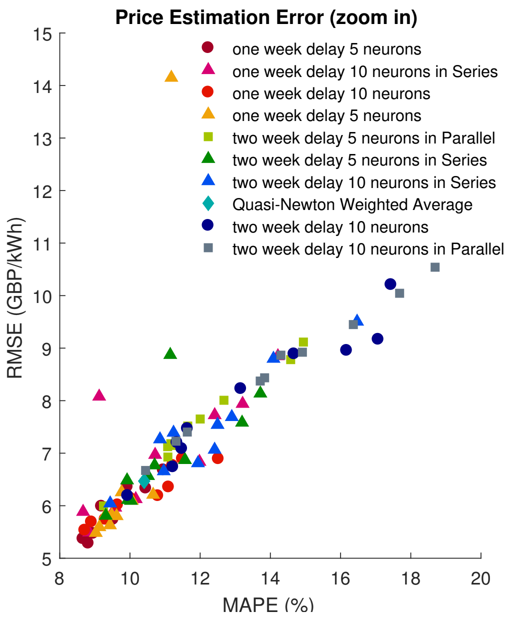

Fig. 5. Error zoom in

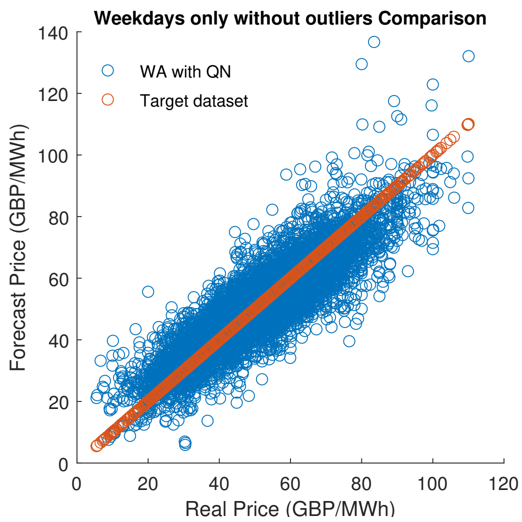

Fig. 6. Error distribution of WA regression Model solved with QN

individual score, with the neural network with a single layer of 5 neurons and one week of delay achieving a MAPE of 8.67 % and a RMSE of 5.88 GBP/KWh.

VII. Conclusion

Two main methodologies were tested for price forecast of the UK's electricity market. It was found that the NARNET achieves the highest accuracy if the outliers are removed from the data, while the WA method is better otherwise.

For the case of the UK price data from 29/12/2017 to 27/02/2020, the NARNETs with one hidden layer and with 2 hidden layers in series with one week of delay achieved the best MAPE score at 8.67 %, with 5.37 GBP/kWh and 5.88 GBP/KWh RMSE respectively.

The WA solution shows that the previous day has the highest weight if each weight must be positive, the season is the most significant term without this constraint.

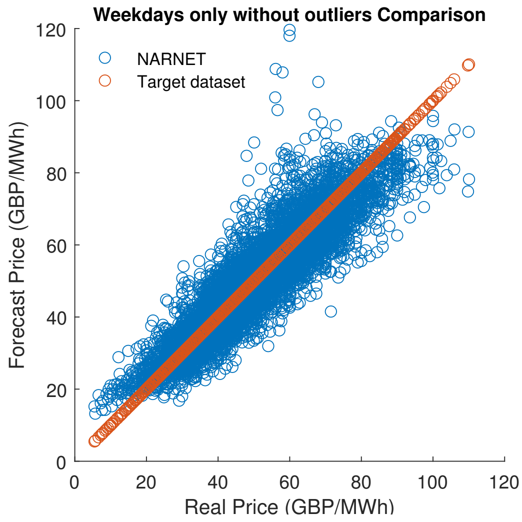

Fig. 7. Error distribution of the NARNET

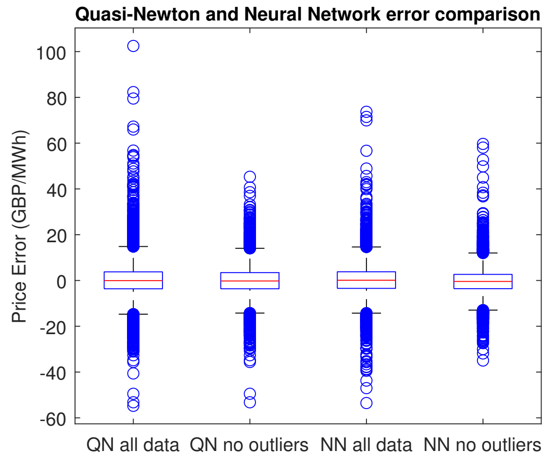

Fig. 8. Error boxplot of best settings for each method

The transfer function of the single layer NARNET was replicated in an agent and implemented as part of the MAS control for the secondary control of the microgrid, to improve the price forecast's accuracy to minimise supply cost.

References

[1] S. K. Sahoo, A. K. Sinha, and N. K. Kishore, "Control techniques in ac, dc, and hybrid ac–dc microgrid: A review," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 6, no. 2, pp. 738–759, 2018.

[2] C. Dou and B. Liu, "Multi-agent based hierarchical hybrid control for smart microgrid," *IEEE Transactions on Smart Grid*, vol. 4, no. 2, pp. 771–778, 2013, ISSN: 1949-3061. DOI: 10.1109/TSG.2012.2230197.

TABLE I
Error results of each method

Method	Average Score		Best Score	
	MAPE	RMSE	MAPE	RMSE
WA Regression model	28.20904	7.11	10.41	6.47
Single Layer NARNET	189.44	8.00	8.67	5.37
Series NARNET	187.93.96	7.75	8.67	5.88
Parallel NARNET	223.98	9.41	9.25	6.00

[3] A. Bidram, A. Davoudi, F. L. Lewis, and Z. Qu, “Secondary control of microgrids based on distributed cooperative control of multi-agent systems,” *IET Generation, Transmission Distribution*, vol. 7, no. 8, pp. 822–831, 2013, ISSN: 1751-8695. DOI: 10.1049/iet-gtd.2012.0576.

[4] G. Liu, Y. Xu, and K. Tomsovic, “Bidding strategy for microgrid in day-ahead market based on hybrid stochastic/robust optimization,” *IEEE Transactions on Smart Grid*, vol. 7, no. 1, pp. 227–237, 2016, ISSN: 1949-3061. DOI: 10.1109/TSG.2015.2476669.

[5] B. Gundogdu, D. Gladwin, and D. Stone, “Battery energy management strategies for uk firm frequency response services and energy arbitrage,” *The Journal of Engineering*, vol. 2019, no. 17, pp. 4152–4157, 2019, ISSN: 2051-3305. DOI: 10.1049/joe.2018.8226.

[6] M. Aragüés-Peñalba, T. L. Nguyen, R. Caire, A. Sumper, S. Galceran-Arellano, and Q.-T. Tran, “General form of consensus optimization for distributed opf in hvac-vsc-hvdc systems,” *International Journal of Electrical Power & Energy Systems*, p. 106 049, 2020, ISSN: 0142-0615. DOI: https://doi.org/10.1016/j.ijepes.2020.106049. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0142061519337639.

[7] V. Bui, A. Hussain, and H. Kim, “A multiagent-based hierarchical energy management strategy for multi-microgrids considering adjustable power and demand response,” *IEEE Transactions on Smart Grid*, vol. 9, no. 2, pp. 1323–1333, 2018.

[8] H. Golmohamadi and R. Keypour, “Stochastic optimization for retailers with distributed wind generation considering demand response,” *Journal of Modern Power Systems and Clean Energy*, vol. 6, no. 4, pp. 733–748, 2018.

[9] A. Botterud, Z. Zhou, J. Wang, R. J. Bessa, H. Keko, J. Sumaili, and V. Miranda, “Wind power trading under uncertainty in lmp markets,” *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 894–903, 2012.

[10] A. Mohsenian-Rad and A. Leon-Garcia, “Optimal residential load control with price prediction in real-time electricity pricing environments,” *IEEE Transactions on Smart Grid*, vol. 1, no. 2, pp. 120–133, 2010, ISSN: 1949-3061. DOI: 10.1109/TSG.2010.2055903.

[11] X. Yang, Y. Yang, Y. Liu, and Z. Deng, “A reliability assessment approach for electric power systems considering wind power uncertainty,” *IEEE Access*, vol. 8, pp. 12 467–12 478, 2020, ISSN: 2169-3536. DOI: 10.1109/ACCESS.2020.2966275.

[12] I. Kim, “Markov chain monte carlo and acceptance–rejection algorithms for synthesising short-term variations in the generation output of the photovoltaic system,” *IET Renewable Power Generation*, vol. 11, no. 6, pp. 878–888, 2017, ISSN: 1752-1424. DOI: 10.1049/iet-rpg.2016.0976.

[13] A. Y. Alanis, “Electricity prices forecasting using artificial neural networks,” *IEEE Latin America Transactions*, vol. 16, no. 1, pp. 105–111, 2018.

[14] F. Llorente, L. Martino, and D. Delgado, “Parallel metropolis–hastings coupler,” *IEEE Signal Processing Letters*, vol. 26, no. 6, pp. 953–957, 2019, ISSN: 1558-2361. DOI: 10.1109/LSP.2019.2913470.

[15] C. Wan, Z. Xu, Y. Wang, Z. Y. Dong, and K. P. Wong, “A hybrid approach for probabilistic forecasting of electricity price,” *IEEE Transactions on Smart Grid*, vol. 5, no. 1, pp. 463–470, 2014, ISSN: 1949-3061. DOI: 10.1109/TSG.2013.2274465.

[16] L. Wu and M. Shahidehpour, “A hybrid model for day-ahead price forecasting,” *IEEE Transactions on Power Systems*, vol. 25, no. 3, pp. 1519–1530, 2010, ISSN: 1558-0679. DOI: 10.1109/TPWRS.2009.2039948.

[17] H. Mosbah and M. El-hawary, “Hourly electricity price forecasting for the next month using multilayer neural network,” *Canadian Journal of Electrical and Computer Engineering*, vol. 39, no. 4, pp. 283–291, 2016.

[18] G. Zhang, H. Zhou, C. Wang, H. Xue, J. Wang, and H. Wan, “Forecasting time series albedo using narnet based on eemd decomposition,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 5, pp. 3544–3557, 2020.

[19] M. E. Cruz Victorio, B. Kazemtabrizi, and M. Shahbazi, “Real-time cost optimisation for power management in microgrids using multi-agent control,” in *2019 9th International Conference on Power and Energy Systems (ICPES)*, 2019, pp. 1–6.

[20] M. E. Cruz Victorio, B. Kazemtabrizi, and M. Shahbazi, “Distributed real-time power management in microgrids using multi-agent control with provisions for fault-tolerance,” in *2020 IEEE 29th International Symposium on Industrial Electronics (ISIE)*, 2020, pp. 108–113.

[21] M. Jawad, S. M. Ali, B. Khan, C. A. Mehmood, U. Farid, Z. Ullah, S. Usman, A. Fayyaz, J. Jadoon, N. Tareen, A. Basit, M. A. Rustam, and I. Sami, “Genetic algorithm-based non-linear auto-regressive with exogenous inputs neural network short-term and medium-term uncertainty modelling and prediction for electrical load and wind speed,” *The Journal of Engineering*, vol. 2018, no. 8, pp. 721–729, 2018.

[22] K. L. Priddy and P. E. Keller, in *Artificial Neural Networks: An Introduction*. SPIE, 2005, ch. 1, pp. 1–12. DOI: 10.1117/3.633187.

[23] X. Chen, Z. Y. Dong, K. Meng, Y. Xu, K. P. Wong, and H. W. Ngan, “Electricity price forecasting with extreme learning machine and bootstrapping,” *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 2055–2062, 2012, ISSN: 1558-0679. DOI: 10.1109/TPWRS.2012.2190627.