Predicting Power Deviation in the Turkish Power Market Based on Adaptive Factor Impacts

Denizhan Guven, Avni Ozozen, Gülgün Kayakutlu, M. Özgür Kayalica

Predicting Power Deviation in the Turkish Power Market based on Adaptive Factor Impacts

Denizhan Guven[1[0000-0003-1605-7661]], Avni Ozozen[2[0000-0001-7093-670X]],

Gulgun Kayakutlu[3[0000-0001-8548-6377]] and M. Ozgur Kayalica[3[0000-0001-9828-7385]]

[1] Eurasia Institute of Earth Sciences, Istanbul Technical University, Istanbul 34467, Turkey
[2] Axpo Turkey, Besiktas, Istanbul 34340, Turkey
[3] Energy Institute, Istanbul Technical University, Istanbul 34467, Turkey

guvende@itu.edu.tr

Abstract. Energy market models are generally focused on energy balancing using the optimum energy mix. In countries where the energy markets are not fully liberalised, the State Regulators reflect any cost of being off-balance on the utility companies and this affects the consumers as well. The right short term prediction of the market trends is beneficial both to optimise the physical energy flow and commercial revenue balance for suppliers and utility companies. This study is aimed to predict the sign trends in the power market by selecting the influencing factors adaptive to the conditions of the day ahead, 10 hours, 5 hours, 2 hours and 1 hour before the electricity balance is active. There are numerous factors consisting of weather conditions, resource costs, operation costs, renewable energy conditions, regulations, etc. with a considerable impact on the predictions. The contribution of this paper is to choose the factors with the highest impacts using the Genetic Algorithm (GA) with Akaike Information Criteria (AIC), which are then used as input of a Recursive Neural Network (RNN) model for forecasting the deviation trends. The proposed hybrid method does not only reduce the prediction errors but also avoid dependency on expert knowledge. Hence this paper will allow both the market regulator and the suppliers to take precautions based on a confident prediction.

Keywords: Energy Market Balancing, Auxiliary Power Market Modeling, Turkish Power Market, Adaptive Prediction, Genetic Algorithm and Recursive Neural Networks.

1 Introduction

The structure of the monopolistic power industry controlled by the government has been changed by the development of deregulation and the emergence of competitive markets since the beginning of the 1990s. Since then, electricity as an indispensable commodity has been traded based on the market rules with spot and derivative contracts [1]. On the other hand, as electricity storage for long durations is not economically feasible at present, it is very crucial to set a balance between generation and consumption for the sake of the electricity grid. Ensuring a complete balance between them is hardly accomplishable because of the uncertainness in both electricity supply and demand. That is why avoiding perfect grid imbalance is not possible. To inhibit grid unstableness, these imbalances should be neutralized by the Transmission System Operator (TSO) [2].

With the transition through cleaner electricity generation, the structure and requirements for supply security are consistently evolving. Since the share of Renewable Energy Sources (RES) in electricity generation increases gradually, the prediction of electricity generation becomes more difficult. This situation leads to make Ancillary Services (AS), such as voltage and frequency control (system services) more critical to ensure a consistent electricity network [3]. System services also called balancing, assures that the active power insertion and withdrawals are met in real-time to keep the frequency constant [4].

In this respect, Turkey is one of the leading countries determining its direction towards cleaner electricity generation from RES. Although electricity generation is dominated by fossil-based power plants, the share of electricity from RES has increased from 26.4% to 40% in the last decade (author's calculation based on EXIST [5] and TETC [6]). In this period, the RES capacity of Turkey has jumped from 15.5 GW to 49.5 GW [7]. On the other hand, this situation may cause electricity system imbalances due to the instability of electricity generation from RES.

With the balancing and settlement regulation published in November 2004, the Day-Ahead Balancing system was adopted to facilitate real-time balancing and improve system security in Turkey [8]. Since then, Balancing and settlement regulation have undergone various changes as a result of the studies carried out in line with the demands of the market participants to ensure the healthy execution of real-time balancing in the electricity market and to fully respond to the needs of the market participants.

Balancing Power Market (BPM) is operated by the Turkish Electricity Transmission Company (TETC). Although a market with balanced production and consumption amounts is presented to the system operator (National Load Dispatch Centre-NLDC) with the Day-Ahead Market, deviations may occur in real-time. For instance, if a power plant is out of order due to a malfunction or the electricity generation from RES does not meet the foresight or a large consumption facility starts to operate/stop working suddenly, it disrupts the system balance (See Figure 1). In this case, NLDC tries to achieve the system balance by using the offers submitted to the BPM to ensure the balance.

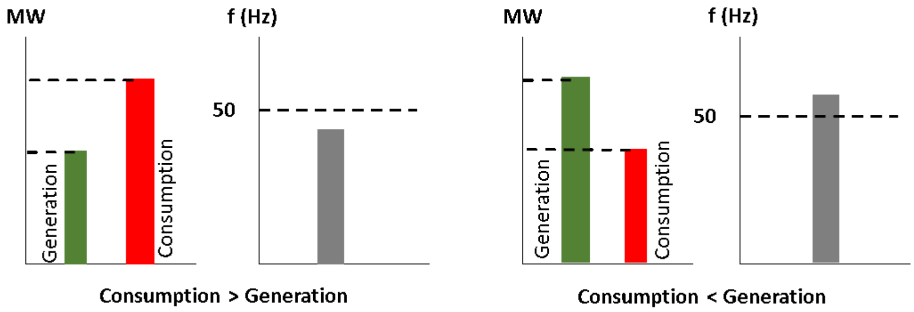

Fig. 1. Frequency imbalances

Therefore, this study aims to present a new methodology combining the Genetic Algorithm (GA) model and the Recurrent Neural Network (RNN) model to predict grid imbalance volumes in Turkey. GA is applied to select the best regressors, and RNN is utilized to predict the grid imbalances for different forecast horizons. Although there are numerous balancing price forecasting studies in the literature, only a few studies are focusing on volume forecasting. In this context, this study contributes to the literature by filling two gaps. Firstly, this study will be the first study that combines these methods to forecast grid imbalances. Secondly, as far as is known, there is not a single study for forecasting imbalance volume specific to the Turkish electricity balancing market. As a country with very high grid imbalance volumes, this study is very crucial for Turkey.

This study is organized as follows. An overview of grid imbalance studies is given in the next section. The third section presents a review of the methodology used in this study. In the fourth section, the proposed model is run with appropriate data specific to the Turkish case. Finally, the last section concludes with the discussion on estimation results.

2 Literature

The number of studies for the electricity market imbalance is progressively increasing in the literature. These studies can be grouped under three different types, namely, fundamental, price forecasting, and volume forecasting. Within the scope of fundamental works, while some studies aim to form a framework for how the balancing market works, some studies analyse the impact of Renewable Energy Sources (RES) on the balancing market structure. In this context, Ortner & Totschnig [9] aimed to evaluate the suitability of electricity balancing markets in Europe based on the different scenarios for 2030. A large-scale market model was established to obtain a perception about market shares and revenues of day-ahead, intra-day and balancing markets. The result of the model exhibited that balancing volume will grow in the mid-term, but the financial magnitude of the balancing market will range in a few percent of day-ahead market volumes. In another example, Lago et al. [2] discussed whether the grid, which is controlled by Transmission System Operator (TSO), is not efficient with respect to grid imbalances. It was stated that market players may help to compensate grid imbalances under the supervision of TSO. As a result of this study, it was proven that the new market structure may contribute between 10-20 percent of the total balancing energy needed and reduce the balancing costs.

Unlike previous studies, Borggrefe & Neuhoff [10] analysed the ongoing designs of intraday and balancing markets in North America and Europe. They evaluated these market structures whether they are capable of sufficiently dealing with the reliability of wind. To evaluate market structures, they defined six criteria, and it was presented that none of the power markets in Europe can completely meet all these criteria. On the other hand, some electricity markets in North America, such as New York ISO, Texas and California satisfy all criteria. Similar to this study, Goodarzi et al. [11] assessed the impact of both wind and solar energy forecasts errors on electricity market imbalance volumes and spot electricity prices in Germany. Ordinary Least Squares (OLS) regression, quantile regression and Auto-Regressive Moving Averages (ARMA) methods were performed to determine these relationships. This study showed that higher wind and solar forecast errors cause an increase in the absolute values of imbalance volumes, and therefore this situation may lead to higher spot prices. More recently, Sirin & Yilmaz [12] examined the impact of RES on the Turkish balancing market prices utilizing quantile regression and investigated how imbalances are altered based on the renewable energy technology with Ordered Logistic Regression. As a result of this analysis, it was found that an increase in renewable energy production suggests lower prices, but superior positive imbalances for the electricity grid.

In addition to studies explaining the fundamentals of balancing electricity markets, the bulk of the literature focuses on price forecasting on the balancing market. Kölmek & Navruz [13] attempted to model the balancing market price using Artificial Neural Networks (ANN) for Turkey. After establishing the best proper configuration, the results were compared with a model built with an auto-regressive integrated moving average (ARIMA). It was also stated that the ANN model performs better than the ARIMA model with an average error rate of 14.15%. Similar to the previous study, Mordasiewicz [14] utilized three econometric methods, including Moving Weighted Average (MWA), OLS, and Autoregressive Moving Average with Exogenous input (ARMAX) to forecast balancing price in Poland. ARMAX method was found to be the best method for this purpose. As one of the latest studies, Lucas et al. [15] used three machine learning models, namely Gradient Boosting (GB), Random Forest (RF) and XGBoost to forecast balancing market prices for the UK. XGBoost method was detected as the most successful model with a Mean Absolute Error (MAE) of 7.89 £/MWh among others.

Although there are many other fundamental and price forecasting studies in the literature as aforementioned, this paper focuses on volume forecasting in balancing markets. Several studies are focusing on the imbalance volume forecasting which is the main scope of this study. However, the literature on this subject is quite limited. Within this scope, Garcia & Kirschen [16] proposed a methodology combining both classical and data mining techniques to forecast electricity grid imbalance volume for England and Wales. It was stated that the proposed model shows great potential, but the accuracy of forecasts is found to be still low. For the Czech Republic, Kratocvil [17] developed an imbalance volume interval forecasting model which also involves a variety of exogenous variables. Depending on different intervals, accuracies of imbalance forecasts were found to be ranged between 60 and 70 percent. In a similar study, Contreras [18] built up a random forest model to forecast electricity grid imbalance volume and utilized a genetic algorithm for applying the bidding plan minimizing the imbalance costs in Spain. As a result of this study, the random forest technique was found to be completely applicable for forecasting electricity system imbalance volume. Parallel with the aforementioned studies, Ferreira [19] applied both Boosted Decision Tree regression model and Decision Forest regression machine learning techniques to forecast balancing power volume and balancing power price for Nordic countries. The forecasting model showed high accuracy (87%) at determining the direction of the total imbalance.

The model also succeeded to detect the peaks in the magnitude of the total imbalance. More recently, Salem et al. [20] presented a quantile regression forests model to forecast two-hour-ahead imbalances in the form of prediction intervals for Norway. This established model was found to have great potential towards a fully automated decision support tool with a longer prediction horizon.

Despite the number of studies surveyed above, the imbalance volume forecasting in Turkey has not been sufficiently analysed. This study attempts to fill this gap by selecting the best variables with the Genetic Algorithm (GA) method and forecasting imbalance volume with the Recurrent Neural Network (RNN) method. More importantly, to the authors' knowledge, this is the first study combining GA and RNN methods to forecast electricity grid imbalance volume.

Table 1. A brief summary of Literature

Reference	Type	Region/ Country	Concept
[16]	VF	England & Wales	Proposing a new imbalance volume forecasting model by combining classical and data mining techniques
[14]	PF	Poland	Forecasting balancing market price using MWA, OLS, and ARMAX
[10]	F	Europe & N. America	Evaluating the impact of wind intermittency both on the current intraday and balancing markets
[21]	PF	Europe	Analyzing the impact of the imbalance pricing mechanism on market behavior with Agent-based modelling
[22]	F	Spain	Proposing a new scheme to minimize the use of ancillary services
[1]	PF	Review	Explaining the complexity, strengths, weaknesses, opportunities and of price forecasting tools
[23]	F	Germany	Analyzing the interactions between balancing system and variable renewable energy sources
[24]	PF	-	Benchmarking time series based balancing market price forecasting models
[13]	PF	Turkey	Forecasting balancing market price using ANN
[17]	VF	Czech Republic	Developing an imbalance volume interval forecasting model which also involves a variety of exogenous variables
[18]	VF	Spain	Proposing a random forest model to forecast electricity grid imbalance volume
[19]	VF	Nordic Countries	Applying Boosted Decision Tree regression and Decision Forest regression models to forecast balancing power volume and balancing power price
[25]	F	Germany	Presenting an empirical examination of the function of intraday trading in the balancing system

[9]	F	Europe	Evaluate the suitability of electricity balancing markets based on the different scenarios for 2030
[11]	F	Germany	Assessing the impact of wind and solar energy forecasts errors on electricity market imbalance volumes and spot electricity prices in Germany using OLS, quantile regression and ARMA
[26]	PF	Hong Kong	Proposing a genetic algorithm based on dynamic pricing
[27]	PF	Germany	Optimizing balancing power bidding strategy with a MILP model considering the price forecasts and spot market prices
[20]	VF	Norway	Presenting a quantile regression forests model to forecast two-hours-ahead imbalances in the form of prediction intervals
[28]	F	China	Examining the impact of imbalance settlement design and utilizing an effective evaluation method
[3]	F	Swiss	Focusing on balancing market design and opportunity cost
[15]	PF	UK	Forecasting balancing market price using GB, RF and XGBoost
[29]	F	NL	Investigating the impact of market design on strategic bidding behavior
[2]	F	NL	Examining the effect of market players on the balancing market
[12]	F	Turkey	Examining the impact of RES on the balancing market prices, and investigating how imbalances are altered based on renewable energy technology

(The abbreviations in the type section: VF-Volume Forecasting, PF-Price Forecasting, and F-Fundamental)

3 Methodology

This study selects the best features to be used in the forecasting model by using the Genetic Algorithm (GA). The Genetic Algorithm concept is derived from biology and its philosophy depends on Darwin's theory of survival of the fittest. Genetic algorithms are a technique based on both natural genetics and computer science. The terms used to describe genetic algorithms are a mixture of terms used in these two sciences. Thus, the terms used in genetics and their responses in GA are shown in Table 2.

Table 2. Terms of GA and their descriptions (Adapted from Vural [30])

Term	Response in GA	Description
Gene (Bit)	Character feature	It is the element that determines the character of an individual, and therefore its value.
Chromosome (String)	Individual	It is the string formed by the combination of genes. Chromosomes are formed by a certain coding system. They show candidate solutions.
Genotype	Gene structure of the individual	It is a candidate solution that involves certain groups of genes within the chromosomes.
Phenotype	Decoded gen structure	It is an alternative solution set. It is the determination of the original value by decoding the genotype.
Allele	Feature value	It is the value set of genes.
Locus	Position of character	It is the position (place) of the gene on the chromosome.
Population	Candidate solutions community	It is a collection of solutions formed by a combination of chromosomes. It is usually kept constant throughout the algorithm.
Fitness Function	Objective function	It is the value that shows the survival status of the individual. Higher values indicate that the individual is more likely to survive.
Selection	Selection	It is the selection of living things that can survive from the population.

The fundamental characteristic of a GA is the manipulation of a population whose individuals are characterized by having a chromosome [31]. A chromosome can be also formed as strings of binary bits (0 or 1). Every bit has a certain position within the string describing the chromosome to which it belongs. The fitness function (F) provides the connection between the GA and the problem. The F builds a matching from the chromosomes to some group of real numbers. The higher F indicates a better level of adaptation of the individual [32]. This procedure is generative in virtue of three main operators namely, mutation, crossover and reproduction.

The reproduction operator is a mechanism that chooses the best-fit individual strings in accordance with some selection operators and it is responsible for selecting the members which are authorized to reproduce during the ongoing generation. The selection of these members is dependent on the basis of their fitness values and the best-fit individuals are handed down to the next generations.

The crossover mechanism makes the exchange of chromosomes between mated parents possible. Mated parents create a child with a chromosome series that is a mixture of the parent's chromosomes (See Figure 1a). The mutation operator is a background operator which makes random and self-generated changes in different chromosomes. In Figure 1b, four different mutation types are shown. According to Haldenbilen and

Ceylan [32], this operator has a very important role in GAs by replacing the genes that are lost from the population during the selection progress so that they can be tried in a new generation or bring the genes that are not present in the initial population.

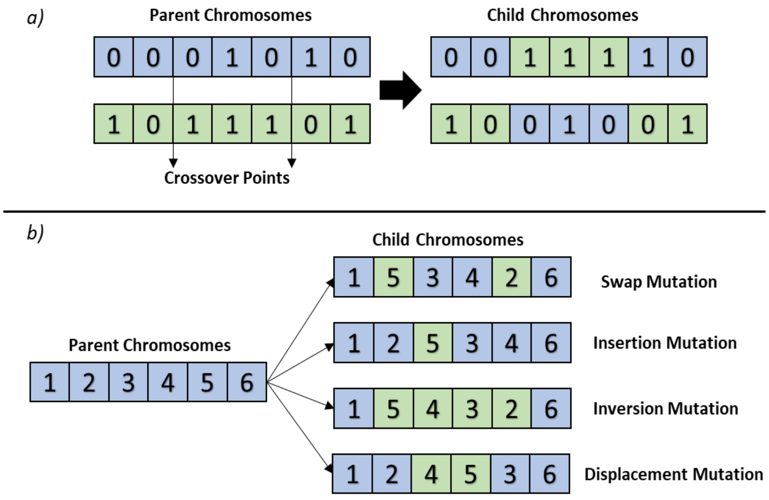

Fig. 2. a) Two-point crossover example b) Mutation examples

The process of GA is shown in Figure x. This cycle continues until the stopping criteria are provided. The GA process starts with initial genetic coding. In this very first step of the process, a phenotype is transformed into a genotype. A population is described as a collection of chromosomes and a chromosome as a collection of genes. When a set of solutions are created, it is necessary to determine how good these solutions are. For this, in the second step, the objective function must be defined. One advantage of the GA is that it brings ease of use to problems where it is difficult to formulate the objective function. The third step of the process is to create an initial population. With this step, individuals (parents) to be used in the algorithm are created. The choice to be made here is the number of individuals in the starting population. The starting population number for the best solution is different for each problem type. The population number is generally maintained across generations. The fourth step is the evaluation of fitness which is the function that finds out how good the chromosomes are. This function forms the brain of GA, and it is the only part that works specifically for the problem in GA. The fitness function turns the chromosomes into the parameters of the problem, in a way decoding them.

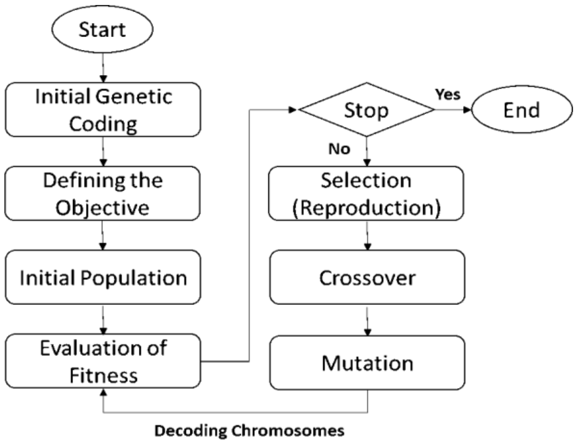

Fig. 3. Steps of the Genetic Algorithm

The next three steps are the generative procedures; selection, crossover, and mutation. After these steps, the last part of the GA process is defining the stopping criteria. The criteria required for the completion of the cycles of the Genetic Algorithm (termination of generations) can be defined differently. Some of the criteria are [33]:

1. After a certain number of generations is finished, the cycles are stopped and it is checked who is most suitable over the generations.
2. As a result of evolution, when the improvement in fitness function falls below a certain value, the cycles are terminated.
3. By giving a certain period for evolution, evolution can be stopped when this time is over.
4. By giving a certain value for the fitness function, the evolution can be completed in the first cycle exceeding this value.

After the best features are selected with GA, the forecasting model was built on the Recurrent Neural Network (RNN) methodology in this study. RNN's are one of the most popular Neural Network (NN) architectures for forecasting problems. The most widely-used RNN architectures are the Elman RNN (ERNN), the Long Short Term Memory (LSTM), and the Gated Recurrent Unit (GRU) [34]. Among the various RNN architectures, only ERNN architecture will be discussed in this study.

ERNN is basic neural networks with a hidden state. However, it does not contain an advanced gating mechanism [35]. ERNN architecture can be mathematically formulated as

$$h_t = f(W_i x_t + W h_{t-1} + b_h) \tag{1}$$

$$z_t = g(W_0 h_t + b_0) \tag{2}$$

where $h_t \in \mathbb{R}^d$, $x_t \in \mathbb{R}^m$, and $z_t \in \mathbb{R}^d$ represent the hidden state of the RNN cell, input and output of the cell at time $_t$, respectively, while d and m indicate the size (dimension) of the cell and input, respectively. W_i stands for the hidden weights of input, W is the recurrent weight matrix of the hidden layer and b_h is the bias vector for the hidden state. $W_0 \in \mathbb{R}^{d \times d}$ and $b_0 \in \mathbb{R}^{d}$ are the weight matrix and the bias vector of the cell output [34, 35]. The structure of ERNN is shown in Figure 4.

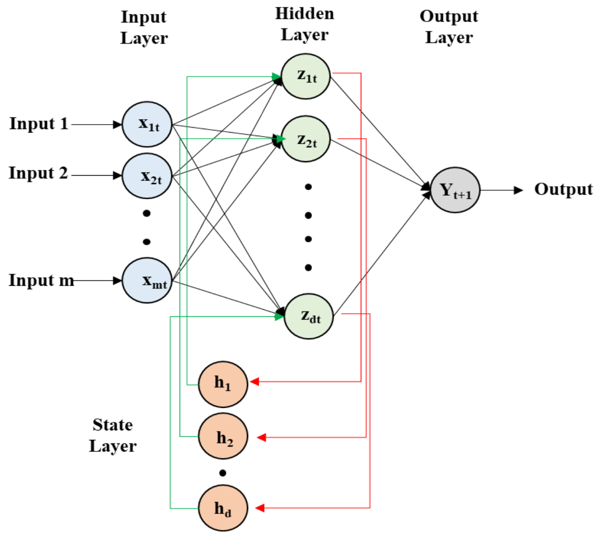

Fig. 4. Structure of ERNN

RNNs have some advantages, such as nonlinear prediction abilities, quicker convergence, and more accurate mapping capability. In RNN architecture, the hidden layer outputs are permitted to feedback onto themselves over a buffer layer, called the recurrent layer. This feedback provides ERNN with some capability, including learning, recognizing, and generating temporal patterns. Every hidden neuron has a connection with only one recurrent layer neuron. Thus, the recurrent layer virtually comprises a duplicate of the state of the hidden layer one trice before. The number of recurrent neurons is equal to the number of hidden neurons [36]. The current hidden state is contingent upon both the hidden state of the prior time step and the current input. The feedback cycle in the RNN architecture linking its current state to the next state supports this dependency. These connections have a crucial role to take into account past information when updating the current cell state [34].

4 Imbalances in State Regulated Power Markets: Case of Turkey

Since renewable energies take a growing share in the energy resource portfolio, short-term forecasts considering the effect of intermittent resources on both the price and imbalances of the power market as in Jiang et al. [37], Qiu et al. [38] and Gligoric et al. [39]. These researches either try to reduce the prediction errors for the day ahead analysis or introducing stochasticity and fuzziness. However, the wind and solar energies are very much dependent on meteorological conditions and that is why they change by the hour and with the impact of several factors at a time. There are certainly grid or turbine outage predictions to realize the power market balances [40]. Market imbalances are more important in partly liberated markets since the cost will be effective on both the suppliers and the consumers. This is shown in the spread differences drawn for Germany and Turkey, shown in Figure 5.

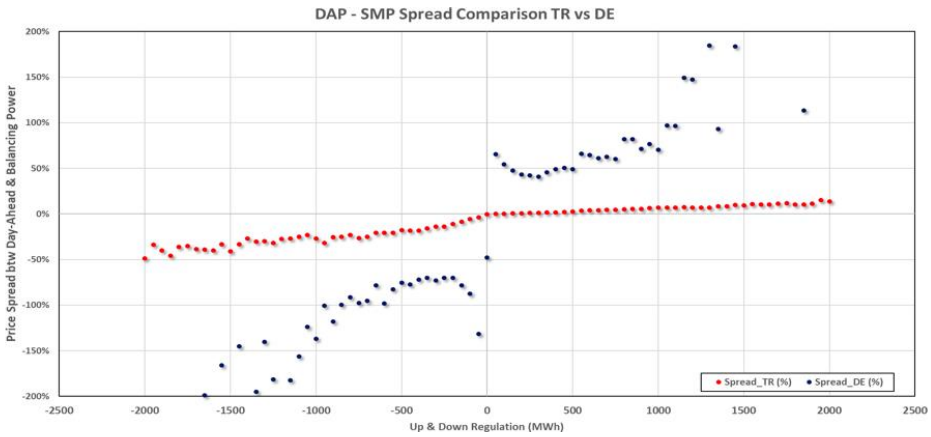

Fig. 5. Comparison of Germany and Turkey

After having reviewed the imbalance studies and observing the power markets we can conclude with the following list of items that cause the imbalances in the power markets:

- **Power markets fluctuate a lot in the short term**. There are unavoidable prediction errors whether the analysis is made for the day ahead conditions or the hour ahead. The prediction errors are very sensitive to the demand because of industrial and commercial demand variations. However, the meteorological conditions vary in even seconds to affect renewable energy resources. Intermittent energy flow due to the operational conditions of the supply unit or transmission infrastructure are also unavoidable. Prediction errors cause the market regulator to apply penalties to the suppliers and utility companies.
- **Supplier behavior in planning the available capacity.** Opportunity costs of upward or downward regulations are important for the suppliers. If the suppliers see some profit in tertiary balancing or a chance to avoid penalties on imbalances, they might give day ahead commitments accordingly. Unexpected commercial behavior of the suppliers might cause imbalances in supplier unit outages. There are even outage problems based on suppliers running the operations without any plans.
- **Discrepancies between market operations and physical energy flow.** It is a fact that the power markets operate time-dependent which means the actions are taken in minutes, hours or days. That means the market operations are run and regulated with

a discrete approach, whereas the resource flow, i.e the physical energy flow is continuous. The stochasticity in the continuous flow should be considered when any forecast of wind or solar resources or plant outages applied, though the load dispatch schedules are planned by combinatorial optimisation.

- **Unexpected Grid Outages may occur.** Grid operators try to balance the power market in terms of frequencies (50Hz) and the TSO ensures a balance between the production and electricity consumption 24 hours a day 7 days a week. Any trend for the negative outage will cause blackouts. Suppliers may put the balance in danger through unplanned plant outages, TSO asks for higher or lower production. The regulator's behavior to smooth the prices may also affect consumer behavior. The regulator can have special contracts with the big consumers or ask for the use of cheaper tariffs.

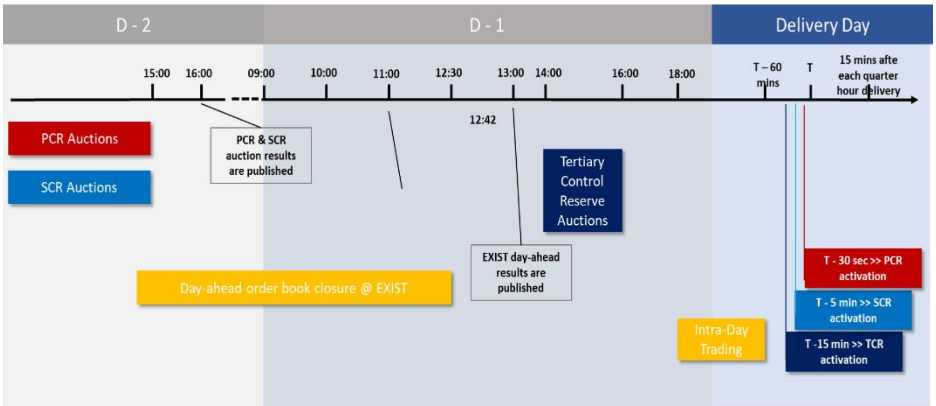

Fig. 6. Timeline of Power Trading in Turkey (Finalized generation/consumption schedules are nominated to corresponding TSO by the BRP on D-1 at 16:00 each day. The nominations can be updated every 30 minutes after the corresponding intra-day product's gate closure)

All the above-mentioned imbalances can only be balanced through the TSO interventions by purchasing or reducing regulations. Balancing the imbalances have a considerable cost for the regulator which cannot be reflected onto the end consumer. In some countries, it is socialised and the generation imbalances penalties are given for the suppliers, whereas in emerging countries like Turkey, the penalty is shared among the supplier and the utility company, which has some influence on the end consumers. The ideal market should be the one where commercial regulations are designed to motivate the mitigation of imbalances. As long as the ideal market does not exist we should have more reliable predictions for the benefit of all the role players in the market.

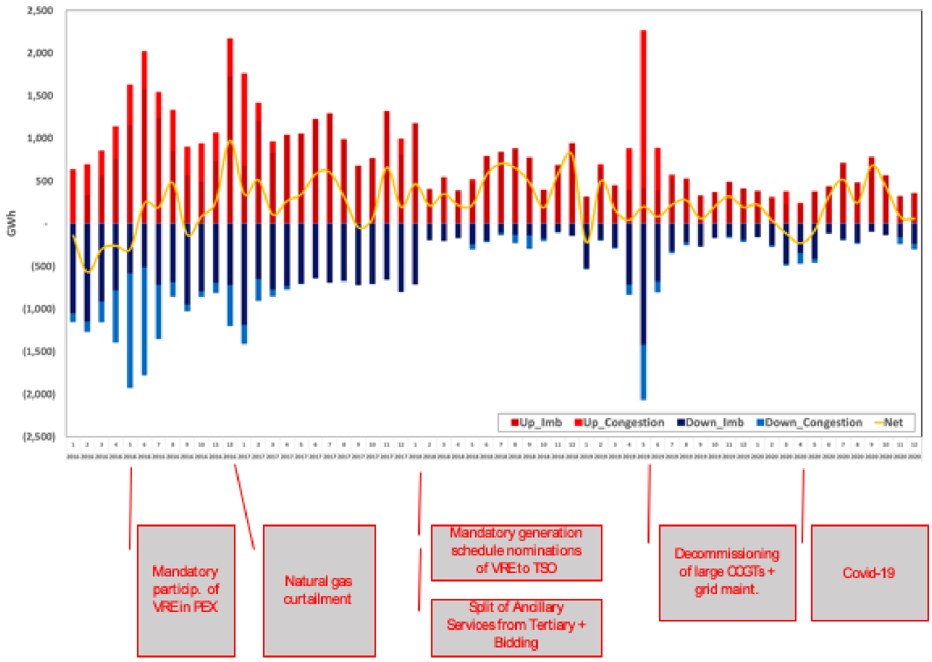

Fig. 7. A Sample of TSO regulations in the past years in Turkey (based on TETC data)

Better predictions will support the trading optimisation, which is not only considered by the suppliers but has beneficial impacts on TSO and the utility companies as well. Though recently there are new technologies like smart meters, load shedding and electricity storage, it is wise to avoid the costs before investing in the new technologies. That is mainly because the prediction of imbalances will help the optimisation of physical energy flow. If all the resources of imbalances could be transferred into a mathematical formula to represent the imbalances the operation costs would be immediately reduced (Fig. 8).

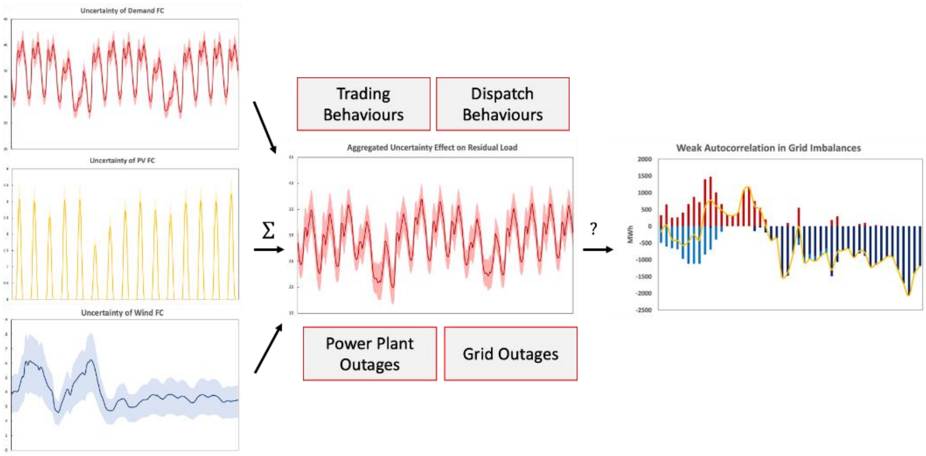

Fig. 8. Is there a transfer function to represent imbalances?

5 Imbalance Prediction Model Proposed

5.1 Data

Prediction of the trend in imbalances (upwards/downwards) demand, supply, resource data are used considering the financial values and variances as well. The meteorological conditions and the day and time effects are also taken into account in addition to price as seen in Table 3 (Appendix).

Demand, wind strength and solar irradiation history are taken as one day and one week previous to the day of prediction to see the trends. Both the financial Cost and variances are considered while taking the return on revenue values for the same history.

Supply is represented by the availability of dispatchable units, reserve margins and the residual loads. These data are also taken to represent one day and one week ahead situations. The financial costs and variances of supply items are also input.

As to price data both market prices and ancillary service prices are taken as input. Opportunity costs are the original input of the proposed model.

All the above are taken for each hour of 24 hours of each day of the week, differentiating the holidays and workdays. When this multiplier is considered for each hour we have 33*2*7*24=11,088 regressors giving the imbalance trend as an output. Each of the data is a time series of at least in time series of a few years.

The time-series data is preprocessed in order to fulfil the missing data and to normalise in binary form to have consistency.

5.2 Phase 1: Regressor Selection

In order to reduce the prediction process time and improve the results, the number of regressors is selected to have the highest impacts for the considered hour of the day. These change a lot by the weekdays and weekend or by holidays. Our GA model uses Akaike Information Criteria (AIC)for the fitness function to be optimised. This is a linear regression model which minimises the least squared error of learning which penalises the high number of factors as it tries to avoid both local optima trap and data overfit.

$$\text{AIC} = -2(\text{log-likelihood}) + 2\text{K} \tag{3}$$

where K is the number of factors and log-likelihood is the fitness measure.

An initial population of 700 data is randomly generated, of which 15% of best performing chromosomes are selected. The cross-over is applied as 80% for fast learning. A mutation rate of 10 % will be used to avoid fast local convergence. In this algorithm, the stopping criteria are defined as no change in 100 consecutive runs.

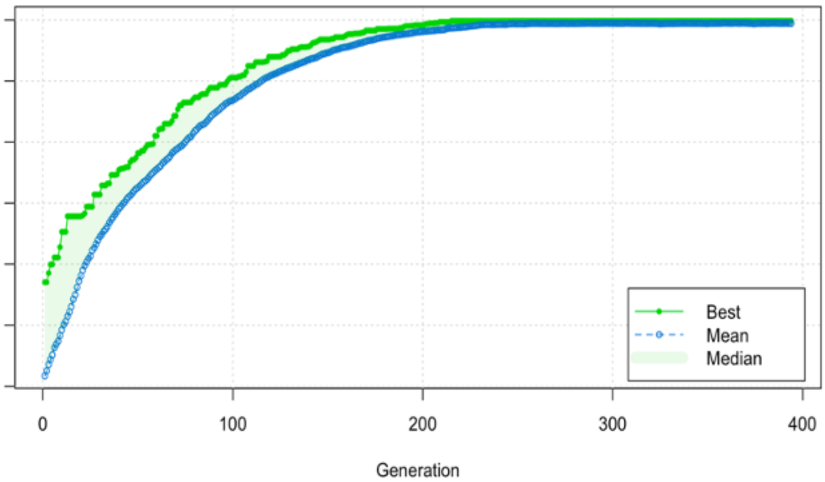

Fig. 9. Regressor Selection using GA

5.3 Phase 2: Imbalance Trend Prediction

The output of Phase 1 is used as input of Phase 2. Hence, regressors with the highest impact are chosen and used as an input for a recursive neural network (RNN) to process classification to predict the next hour (day) imbalance to be upwards or downwards. Elman RNN is modelled to benefit the memory recording.

The transfer function is chosen as *tanh* because it realizes the overfitting and corrects fast in classification application. It is used both between the input and the hidden layers as well as between the hidden layer and the output with a learning rate of 10% to allow fast learning. The number of hidden nodes are tried as similar to the number of input nodes, half of the number of input notes, double the number of input nodes and two times the average of the number of input and output nodes. Best results are achieved by the last one. The prediction errors are calculated using Mean Absolute Percentage Error (MAPE). The formulation of MAPE is given as

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right| \qquad (4)$$

5.4 Results and Discussions

The model is applied to perform a prediction for the day ahead (12 to 36 hours), 10 hours ahead, five hours ahead, two hours ahead and one hour ahead. The results are compared with the naive flipping coin, linear regression, GA and long term average which are generally accepted methods applied by the suppliers, which are given in Figure 10.

Flipping a coin can easily be defeated by long term averages. The linear regression model (OLS) uses the benefits of modelling with certain parameters. The most relevant parameters are chosen by GA and the difference can be recognised. Signals in real-time are only observed in the proposed model.

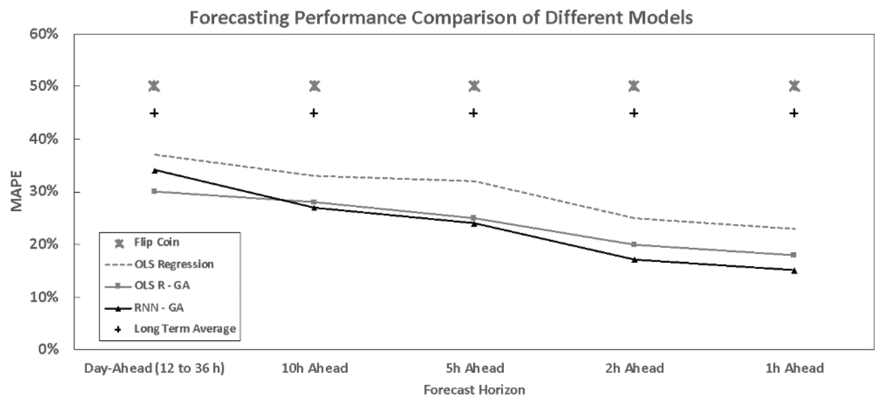

Fig. 10. MAPE comparisons of predictions in different time horizons

These results can be achieved by a dedicated expert to whom it would cause a lot of time. The proposed model processes automatically using the machine learning method and it spends just a few seconds of computer time. In a complex world of energy markets, a number of input factors are non-stop changing which would need the model to be renewed if an adaptive model like the proposed is not used.

Moreover, using the GA with random populations, cross over and mutations avoid the shade of the historical conditions but adapt to the current situation by choosing the regressors at the last hour applied. Using the RNN method helps that by improving what is kept in the memory cell.

The achievements of this study imply that the proposed model will be most beneficial in intraday markets, where prediction up to 1-2 hours are more critical. The benefits are equally important for the commercial players and the regulator who will perform the balance faster with the suppliers being well prepared. The cost of active energy, capacity loans and planning costs will be reduced in the long term.

The suppliers can also increase the revenues by avoiding the capacity loans and having more knowledge about the business conditions in the market.

6 Summary and Conclusion

This study aims to forecast grid imbalance volumes by combining the Genetic Algorithm and Recurrent Neural Networks methods. The Genetic Algorithm method is used to select the best regressors, whereas, Recurrent Neural Network is applied to forecast grid imbalances for different forecast horizons using these selected regressors. After running the established model with the appropriate data, the results are compared with that of Ordinary Least Squares Regression in the presence and absence of Genetic Algorithm models.

The proposed model will be a leader in adaptive prediction in the power markets where real-time is difficult to handle due to deterministic and discrete regulations. The adaptive approach gives a new angle by changing the input based on changes in the market as close as one hour.

Achievements of the proposed model will contribute to all the role players in countries with partly liberated power markets. In those countries, the regulator reflects all the costs on the suppliers and the utility companies, that is why even the end consumers will take some share of the cost of imbalances

In future studies, more variety of the features can be analysed. We suggest that the features are selected using the swarm optimisation so that regressors will be selected not only by impact performance but also being uncorrelated with the other features. stochasticity and continuity in the factors are to be handled by using a stochastic Deep Learning method, which can also add more precision in predictions.

Appendix

Table 3. Potential Inputs

Day-Ahead (12 to 36h)	10h Ahead	5h Ahead	2h Ahead	1h Ahead	Pre-Process	Capturing
D-1 Demand					Normalized btw min-max	Participants' behaviours
D-7 Demand					Normalized btw min-max	Participants' behaviours
D Demand FC	Update	Update	Update	Update	Normalized btw min-max	Forecast error of participants
D Demand FC Variance	Update	Update	Update	Update	Normalized btw min-max	Forecast error of participants
D-1 Wind					Normalized btw min-max	Participants' behaviours
D-7 Wind					Normalized btw min-max	Participants' behaviours
D Wind FC	Update	Update	Update	Update	Normalized btw min-max	Forecast error of participants
D Wind FC Variance	Update	Update	Update	Update	Normalized btw min-max	Forecast error of participants
D-1 RoR					Normalized btw min-max	Participants' behaviours
D-7 RoR					Normalized btw min-max	Participants' behaviours
D RoR FC	Update	Update	Update	Update	Normalized btw min-max	Forecast error of participants
D RoR FC Variance	Update	Update	Update	Update	Normalized btw min-max	Forecast error of participants
D-1 PV					Normalized btw min-max	Participants' behaviours
D-7 PV					Normalized btw min-max	Participants' behaviours
D PV FC	Update	Update	Update	Update	Normalized btw min-max	Forecast error of participants
D PV FC Variance	Update	Update	Update	Update	Normalized btw min-max	Forecast error of participants
D-1 Residual Load					Normalized btw min-max	Participants' behaviours
D-7 Residual Load					Normalized btw min-max	Participants' behaviours
D Residual Load FC	Update	Update	Update	Update	Normalized btw min-max	Forecast error of participants
D Residual Load FC Variance	Update	Update	Update	Update	Normalized btw min-max	Forecast error of participants
D-1 Nat. Avg. Temp.					Normalized btw min-max	Participants' behaviours
D-7 Nat. Avg. Temp.					Normalized btw min-max	Participants' behaviours

D Nat. Avg. Temp. FC	Update	Update	Update	Update	Normalized btw min-max	Forecast error of participants
D Nat. Avg. Temp. FC Variance	Update	Update	Update	Update	Normalized btw min-max	Forecast error of participants
D-1 Availability of Dispatchable Units					Normalized btw min-max	Balancing providing capacity
D-7 Availability of Dispatchable Units	Update	Update	Update	Update	Normalized btw min-max	Balancing providing capacity
D Availability of Dispatchable Units FC	Update	Update	Update	Update	Normalized btw min-max	Balancing providing capacity
D-1 Reserve Margin					Normalized btw min-max	Grid tightness
D-7 Reserve Margin	Update	Update	Update	Update	Normalized btw min-max	Grid tightness
D Reserve Margin FC	Update	Update	Update	Update	Normalized btw min-max	Grid tightness
Weekday Type					Binary	Pattern recognition
Holiday or Not					Binary	Pattern recognition
Hour of the Day					Binary	Pattern recognition
Ancillary Service Prices					Normalized btw min-max	Generator behaviours
Opportunity Costs of Providing Ancillary Services					Normalized btw min-max	Cost efficiency of system
Day-Ahead Price FC						Participants' behaviours
	Day-Ahead Price	Day-Ahead Price	Day-Ahead Price	Day-Ahead Price		Participants' behaviours
			Past Imb.	Past Imb.		Latest output data

References

1. Weron, R.: Electricity price forecasting: A review of the state-of-the-art with a look into the future. International journal of forecasting *30*(4), 10 (2014).
2. Lago, J., Poplavskaya, K., Suryanarayana, G., De Schutter, B.: A market framework for grid balancing support through imbalances trading. Renewable and Sustainable Energy Reviews, 110467 (2020).
3. Schillinger, M.: Balancing-market design and opportunity cost: The Swiss case. Utilities Policy 64, 101045 (2020).
4. Nobel, F. A.: On balancing market design. Technische Universiteit Eindhoven: Eindhoven, The Netherlands (2016).
5. [Data] EXIST- Energy Exchange Istanbul: Real Time Generation: 01.01.2020-31.12.2020. https://seffaflik.epias.com.tr/transparency/uretim/gerceklesen-uretim/gercek-zamanli-uretim.xhtml, last accessed 2021/01/29.
6. [Data] TETC- Turkish Electricity Transmission Company: Electricity Generation and Transmission Statistics. https://www.teias.gov.tr/tr-TR/turkiye-elektrik-uretim-iletim-istatistikleri, last accessed 2021/01/31.
7. [Data] TETC- Turkish Electricity Transmission Company: Installed Power Plant Capacity Report: December 2020. https://webapi.teias.gov.tr/file/2bd6a902-7b75-4226-9c54-55ae9afea22f?download, last accessed 2021/02/11.
8. Cetintas, H., Bicil, I. M.: Elektrik Piyasalarında Yeniden Yapılanma ve Türkiye Elektrik Piyasasında Yapısal Dönüşüm. Optimum Ekonomi ve Yönetim Bilimleri Dergisi 2(2), 1-15 (2015).
9. Ortner, A., Totschnig, G.: The future relevance of electricity balancing markets in Europe-A 2030 case study. Energy Strategy Reviews 24, 111-120 (2019).
10. Borggrefe, F., Neuhoff, K.: Balancing and intraday market design: Options for wind integration. DIW Berlin Discussion Paper No. 1162, Available at SSRN: https://ssrn.com/abstract=1945724 or http://dx.doi.org/10.2139/ssrn.1945724 (2011).
11. Goodarzi, S., Perera, H. N., Bunn, D.: The impact of renewable energy forecast errors on imbalance volumes and electricity spot prices. Energy Policy 134, 110827 (2019).
12. Sirin, S. M., Yilmaz, B. N.: The impact of variable renewable energy technologies on electricity markets: An analysis of the Turkish balancing market. Energy Policy 151, 112093 (2021).
13. Kölmek, M. A., Navruz, I.: Forecasting the day-ahead price in electricity balancing and settlement market of Turkey by using artificial neural networks. Turkish Journal of Electrical Engineering & Computer Sciences 23(3), 841-852 (2015).
14. Mordasiewicz, Ł.: Price forecasting in the balancing mechanism. Rynek Energii (3), 94 (2011).
15. Lucas, A., Pegios, K., Kotsakis, E., Clarke, D.: Price Forecasting for the Balancing Energy Market Using Machine-Learning Regression. Energies 13(20), 5420 (2020).
16. Garcia, M. P., Kirschen, D. S.: Forecasting system imbalance volumes in competitive electricity markets. IEEE Transactions on Power Systems 21(1), 240-248 (2006).
17. Kratochvíl, Š.: System Imbalance Forecast (Doctoral dissertation). Czech Technical University, Prague, Czech Republic (2016).

18. Contreras, C.: System imbalance forecasting and short-term bidding strategy to minimize imbalance costs of transacting in the Spanish electricity market (Doctoral dissertation). Universidad Pontificia Comillas, Madrid, Spain (2016).
19. Pires Ferreira, P.: Volume and Price in The Nordic Balancing Power Market (Master's thesis). Norwegian University of Science and Technology, Trondheim, Norway (2016).
20. Salem, T. S., Kathuria, K., Ramampiaro, H., Langseth, H.: Forecasting intra-hour imbalances in electric power systems. In: Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 33, No. 01), pp. 9595-9600. (2019)
21. van der Veen, R. A., Abbasy, A., Hakvoort, R. A.: Agent-based analysis of the impact of the imbalance pricing mechanism on market behavior in electricity balancing markets. Energy Economics 34(4), 874-881 (2012).
22. Bueno-Lorenzo, M., Moreno, M. Á., Usaola, J.: Analysis of the imbalance price scheme in the Spanish electricity market: A wind power test case. Energy policy 62, 1010-1019 (2013).
23. Hirth, L., Ziegenhagen, I.: Balancing power and variable renewables: Three links. Renewable and Sustainable Energy Reviews 50, 1035-1051 (2015).
24. Klæboe, G., Eriksrud, A. L., Fleten, S. E.: Benchmarking time series based forecasting models for electricity balancing market prices. Energy Systems 6(1), 43-61 (2015).
25. Koch, C., Hirth, L.: Short-term electricity trading for system balancing: An empirical analysis of the role of intraday trading in balancing Germany's electricity system. Renewable and Sustainable Energy Reviews 113, 109275 (2019).
26. Huang, P., Xu, T., Sun, Y.: A genetic algorithm based dynamic pricing for improving bidirectional interactions with reduced power imbalance. Energy and Buildings 199, 275-286 (2019).
27. Schäfer, P., Westerholt, H. G., Schweidtmann, A. M., Ilieva, S., Mitsos, A.: Model-based bidding strategies on the primary balancing market for energy-intense processes. Computers & Chemical Engineering 120, 4-14 (2019).
28. Wu, Z., Zhou, M., Zhang, T., Li, G., Zhang, Y., Liu, X.: Imbalance settlement evaluation for China's balancing market design via an agent-based model with a multiple criteria decision analysis method. Energy Policy 139, 111297 (2020).
29. Poplavskaya, K., Lago, J., De Vries, L.: Effect of market design on strategic bidding behavior: Model-based analysis of European electricity balancing markets. Applied Energy 270, 115130 (2020).
30. Vural, M.: Genetik algoritma yöntemi ile toplu üretim planlama (Doctoral dissertation). Istanbul Technical University, Istanbul, Turkey (2005).
31. Ozturk, H.K., Canyurt, O.E., Hepbasli, A., Utlu, Z.: Residential-commercial energy input estimation based on genetic algorithm approaches: an application of Turkey. Energy and Buildings 36, 175-183 (2004).
32. Haldenbilen, S., Ceylan, H.: Genetic Algorithm Approach to Estimate Transport Energy Demand in Turkey. Energy Policy 33, 89–98 (2005).
33. Man, K. F., Tang, K. S., Kwong, S.: Genetic algorithms: concepts and designs. Springer Science & Business Media, London (2001).
34. Hewamalage, H., Bergmeir, C., Bandara, K.: Recurrent neural networks for time series forecasting: Current status and future directions. International Journal of Forecasting 37(1), 388-427 (2021).
35. Achanta, S., Gangashetty, S. V.: Deep Elman recurrent neural networks for statistical parametric speech synthesis. Speech Communication 93, 31-42 (2017).
36. Wang, J., Wang, J., Fang, W., Niu, H.: Financial time series prediction using elman recurrent random neural networks. Computational intelligence and neuroscience (2016). DOI: 10.1155/2016/4742515

37. Jiang, Y., Chen, M., You, S.: A Unified Trading Model Based on Robust Optimization for Day-Ahead and Real-Time Markets with Wind Power Integration. Energies 10(4), 554 (2017).
38. Qiu, J. Zhao, J., Wang, D., Zheng, Y.: Two-Stage Coordinated Operational Strategy for Distributed Energy Resources Considering Wind Power Curtailment Penalty Cost. Energies 10(7), 965 (2017).
39. Gligoric, Z., Savic, S.S., Grujic, A., Negovanovic, M., Music, O.: Short-Term Electricity Price Forecasting Model Using Interval-Valued Autoregressive Process. Energies 11(7), 1911 (2018).
40. Sun, P., Li, J., Chen, J., Lei, X.: A Short-Term Outage Model of Wind Turbines with Doubly Fed Induction Generators Based on Supervisory Control and Data Acquisition. Energies 9(11), 882 (2016).