The Development of a Genetic Method to Optimize the Flue Gas Desulfurization Process

Ievgen Fedorchenko, Andrii Oliinyk, Tetiana Fedoronchak, Tetiana Zaiko, Anastasiia Kharchenko

Department of Software Tools, National University "Zaporizhzhia Polytechnic", 64 Zhukovskoho str., Zaporizhzhia, Ukraine, 69063

Abstract

Sulfur dioxide is one of the most commonly found gases, which contaminates the air, damages human health and the environment. To decrease the damage, it is important to control the emissions on power stations, as the major part of sulfur dioxide in atmosphere is produced during electric energy generation on power plants. The present work describes flue gas desulfurization process optimizing strategy using data mining. The optimisation modified genetic method of flue gas desulfurization process based on artificial neural network was developed. It affords to represent the time series characteristics and factual efficiency influence on desulfurization and increase its precision of prediction. The vital difference between this developed genetic method and other similar methods is in using adaptive mutation, that uses the level of population development in working process. It means that less important genes will mutate in chromosome more probable than high suitability genes. It increases accuracy and their role in searching. The comparison exercise of developed method and other methods was done with the result that new method gives the smallest predictive error (in the amount of released $SO_{2)}$ and helps to decrease the time in prediction of efficiency of flue gas desulfurization. The results afford to use this method to increase efficiency in flue gas desulfurization process and to decrease SO_2 emissions into the atmosphere.

Keywords 1

flue gas desulfurization, sulfur dioxide, artificial neural network, genetic algorithm

1. Introduction

According to the Nature Geoscience data, NASA satellites detected 500 new sources of air contamination, about 40 of them produce dangerous sulfur dioxide [1]. This substance is considered to be one of the most threatening gases for the Earth atmosphere. The main sources of dioxide sulfur releases are power plants (which work on solid and liquid fuels) and other metallurgical plants. So, the control of SO_2 amount in flue gas after coal burning is the effective method of decreasing releases into the atmosphere [2].

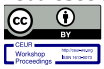
The SO_2 emissions can be reduced by setting equipment for desulfurization on new-built and old coal blocks and also by following corresponding requirements about desulfurization. To increase efficiency of desulfurization and decrease SO_2 emissions, it is important to optimize the desulfurization control system according to the industry needs [2].

The accurate estimation of relationship between process variables and factual sulfurization efficiency is the basis for optimization of desulfurization control system. At the present time, a lot of local equipment for analysis and monitoring flue gases can directly control SO_2 concentration in flue gases it the inlet and the outlet of desulfurization equipment and calculate its efficiency. However, this method is just a simple connection between desulfurization process results and doesn't show monitoring

MoMLeT+DS 2021: 3rd International Workshop on Modern Machine Learning Technologies and Data Science, June 5, 2021, Lviv-Shatsk, Ukraine

EMAIL: evg.fedorchenko@gmail.com (I. Fedorchenko); olejnikaa@gmail.com (A. Oliinyk); t.fedoronchak@gmail.com (T. Fedoronchak); nika270202@gmail.com (T. Zaiko); kharchenko.13.08@gmail.com (A. Kharchenko)

ORCID: 0000-0003-1605-8066 (I. Fedorchenko); 0000-0002-6740-6078(A. Oliinyk); 0000-0001-6238-1177 (T. Fedoronchak); 0000-0003-1800-8388 (T. Zaiko); 0000-0002-4313-555X (A. Kharchenko)



results. Simultaneously, monitoring equipment are affected with external factors and failures which are consequently occurred can lead to inaccurate measurement results. So, the desulfurization efficiency is affected with different factors like suspension pH value, flue gas outlet temperature, density of absorption tower suspension and mass concentration of SO_2 in the inlet of flue gases [2].

In recent years, a lot of methods were suggested, including experimental studies, mathematical models and machine learning methods for prediction the desulfurization efficiency. Among all methods, mathematical models and machine learning models sparked great scientific interest. However, the prediction of desulfurization efficiency is hard to model mathematically. Some researches simplify this system using assumptions, which introduce errors in predictions. Moreover, calculations, which are used in these mathematical models need great computing resources. In this context, machine learning models are useful tools for predictions [3].

One of the most perspective ways of solving this problem is based on using artificial neural networks and genetic algorithms, as the most progressive in relation to problem of prediction the flue gases desulfurization efficiency. Genetic algorithms relate to the class of searching methods, which iteratively improve the quality of solution options by recombination procedures and selections for surviving. Genetic algorithms are widely used in solving difficult multidimensional optimization problems due to general computation scheme, opportunities of parallel realisation and noise tolerance [4].

In the work, the modified genetic method of optimization of flue gas desulfurization process which is based on artificial neural network is suggested. It affords to represent the time series characteristics and factual efficiency influence on desulfurization and increase its precision of prediction.

2. Analysis of published data and problem definition

In the article [5] they propose a prediction model of the flue gas desulfurization efficiency, which is based on long short-term memory (LSTM) neural network. Authors examined 100MW power plan in China, considering main factors, which affect wet flue gas desulfurization of limestone and gypsum. The LSTM neural network is used for building a prediction model of desulfurization efficiency.

The main structure of prediction model based on LSTM consists of five functional modules: an input level, a hidden level, an output level, a network training module and a network prediction module. The output level is in charge of processing the time series output to match the input requirement. The hidden level uses LSTM cells to build one-layer cycle neural network. The output level gives the prediction results. During the LSTM neural network learning the backpropagation through time algorithm is used, which is an analogue to simple backpropagation algorithm.

The advantage of developed model (in comparison with other models) is that the model can represent the influence of time series characteristic on factual efficiency of desulfurization and increase its precision of prediction.

The disadvantage of developed model is that during the learning it is hard to interpret the results. It imposes limitations on the possible improvement of the model. Also, it is impossible to predict when the desulfurization dynamic will change (and if dynamic changes, the model will stop working).

In the work [6] they propose the prediction model of wet flue gas desulfurization of limestone and gypsum which is based on support vector machine method (SVM). The authors consider relation substance-gas, velocity of flue gas, volume of dioxide air, temperature, flue gas dust, concertation of sulfur dioxide outlet, suspension pH value of an absorption tower, relation between calcium and sulfur as independent variables and desulfurization efficiency as dependent variable. The idea of the given method is to build a hyperplane, which is considered as an area of solutions and to separate positive and negative examples from the initial set at most. The support vector machine is an approximative realisation of structural risk minimization method, which is based on the idea that the level of learning machine mistakes on the test set is represented as a sum of learning mistake and the other sum which is depends of a measured Vapnik-Chervonenkis index.

The advantage of the given method is that to solve the classification problem (in contrast to other methods) we can have a small amount of data.

The disadvantage of the given method is that we use not the all data for the classification but only a small part from the borders.

In the work [7] they propose an investigation on data mining and operating optimization for wet flue gas desulfurization systems. The authors suggest a complex criterion with using minimum expenditures as a main function for getting optimal work conditions for wet flue gas desulfurization systems. To increase an accuracy of data mining analysis of the main data they suggest an improved method of fuzzy clustering. This algorithm uses the results of k-means clustering method as a primary conditions and fuzzy clustering method as an analytic method. The procedure of the method has two steps: using k-means clustering method to get the initial number of clusters and their centers and to use fuzzy clustering to calculate the final results.

The advantages of the given method are a simple realisation, the lack of need of an initial layout and low requirements for computing resources.

The disadvantage the given method is its low resistance to outliers, which can create the average. We can use modified k-median clustering as an option to solve this problem. Also, another disadvantage is that the result of the work depends on the primary selection of centroids and, in a general case, this method doesn't give optimal results and can find only suboptimal solutions.

In the work [8] they propose an artificial intelligence-based emission reduction strategy for limestone forced oxidation flue gas desulfurization system. The authors created a neural network based on a multilayer perceptron (MLP) to study relations between controlled input variables and output variables. The given model demonstrated the ability to learn difficult kinetic of reaction between reagents. The input parameters of the model are: suspension pH value, inlet mass of SO_2, temperature, outlet NO_x temperature, percentage composition of oxygen, volume of air for oxidation, density of absorber suspension, oxygen volume, dust mass as a part of volume of flue gas expenditure. The output parameters of the model are: SO_2 mass, Hg, NO_x and outlet dust mass. The model consists of one hidden layer that have 27 neurons. The learning is based on backpropagation method. The learning is performed till it reaches the stop criterion. In other words, when it reaches the change of repeatability mistake on 0,0000001 or during the learning the maximum number of epochs will be achieved (a minor decreasing of mistake happens abroad the epochs).

The advantage of the given method is that it has a high rate of continuity which is realised with the synapse connections help. Also, during the learning of the given neural network (thanks to its inner composition) it sets the regularity between input and output data and generalizes obtained selection-based experience.

The disadvantage of the given method is that it uses a backpropagation method and, as a result, the learning process takes long time. The overgrowth of neural network can lead to its paralysis, when the synaptic weight stops changing not even reaching the end of the learning. The size of a step needs to be always controlled and it is necessary to keep network learning as it steadily forgets previous learning selections.

In the work [9] they propose a solution to optimization problem of wet flue gas desulfurization system. They developed a modified model of wet flue gas desulfurization to predict SO_2 emission. It combines a mathematical model and an artificial neural network. They suggested modified particle swarm optimization (PSO) with penalty method algorithm to increase the saving of desulfurization process, which is directed to minimize operational costs. The PSO method is based on calculation every iteration of centroid coordinates and their velocity. In other words, they solve the optimization problem of fitness function minimizations: the less the value of fitness function is on every iteration, the closer the result is to the optimal centroid of a swarm. Fitness function is defined as an average value of Euclidean distance between particle vector and a vector that define a centroid coordinate.

As a result, the operating parameters were optimized, including lime mortar pH, temperature, lime mortar density and a number of working circulator pumps. The results show that PSO with penalty function provide satisfying productivity of repeatability to decrease the operational costs with some limits.

The disadvantage of the given method is that it easily occurs in local search in multidimensional space and requires great amount of time to find the solution.

In the work [10] they propose prediction and optimization of a desulphurization system using a neural network and a genetic algorithm. The authors consider desulfurization coefficient and economic value as two objectives. They created a model with 10 inputs and 2 outputs. A modified genetic algorithm for building a model and optimizing the cost of desulfurization was created. The genetic algorithm has two steps. In the first step, the genetic algorithm finds suboptimal architecture of artificial

neural network (ANN). In the second step, the genetic algorithm defines suboptimal values of weight coefficient and ANN bias. The first step can be considered as a preparation stage before learning. This step helps to simplify a future task of defining weight and bias. The authors used the following approach: a chromosome was built with a number of neurons from a hidden layer, impulse, velocity of learning, type of activation function and ANN learning algorithm and a couple of other parameters. In connection with various variables it is impossible to use inversion operator in general case. A vector (which characterizes the input of a variable in a learning array) was included in the number of additional genes. Zeros and ones spread among loci randomly with a notation, that the number of zeros can't exceed the half of the number of output variables. The insertion of this type of genes in a chromosome was to identify a suboptimal set of variables. After identifying the ANN architecture, the neural learning starts. After that, the ANN forms the row of weight vectors and biases (which are set to n×1 size) and vector of mean squared error of learning. An initial population is created from a row of vectors (chromosomes), which sooner will evolve with using genetic algorithm.

The advantage of the given method is increased accuracy of classification using genetic algorithm-based approach. The genetic algorithm allows to select the best combination of the basic ANN features by practical consideration. This combination can vary not only in different tasks but also because of changing the static data variables. In this way, this approach allows to optimize desulfurization process.

The disadvantage of the given method is the need in great computing resources. So, the given genetic algorithm can be used only with the access to calculating resources.

The analysis of the works [5-10] affords to claim that the pursuance of the research about optimization and prediction of desulfurization flue gases from nitrous oxide is a live issue. Also, during the work it was figured out that the most efficient method of solving the problem is to use neural networks. This is because they have such advantages as an ability to adaptation in case of changes (that helps them to work even in a critical situation), self-learning, the speed of work (the compilation of prediction and planning can be faster than using exact algorithms), also a great resistance to data noise with an opportunity to use limitless number of independent variables. In the same time, regardless to obvious advantages, the main disadvantage is the difficulty in choosing an efficient learning method and creating the primal amount of data. Also, there is a difficulty in the right distribution the matrix weight on an initial step in the contexts of indeterminacy. We can use genetic method to solve this problem. The benefits of this method are the increased accuracy and decreased amount of time, which is spent on learning. It is efficient to learn some neural networks like multilayer perceptron, ANN with general regression, Kohonen networks and etc [11]. The usage of the evolution genetic algorithm while learning helps to decrease the number of learning cycles by using the parameters which have bigger weight.

So, to solve the problems, which occur in researched methods, it was decided to develop a modified genetic method for setting the weight coefficients of ANN to increase the efficiency of flue gas desulfurization process and to decrease the SO_2 emissions into the atmosphere.

3. The purpose and objectives of the study

The object of the study – desulfurization systems on coal power plants.

The subject of the study – the prediction methods of flue gas desulfurization efficiency

The purpose of the work – to develop a modified genetic method based on a neural network to solve the optimization problem of flue gas desulfurization

The research method – tradition models (decision tree, nearest neighbour algorithm, ant colony optimization algorithm), neural networks, combined methods (neural networks and genetic algorithms, neural networks and multi-agent systems)

4. The development of a genetic method to optimize flue gas desulfurization process

Nowadays, the most famous evolutionary method is a genetic algorithm of finding a global extremum of a multi-extrema function. The idea is in parallel processing of set with alternative

solutions, meanwhile the algorithm searches the most perspective of them. It says about the opportunity to use genetic algorithms in resolving optimization problems and making decisions. That's why it has been selected to set up the synaptic weight of the neural network [12-13].

In the beginning, an initial population is created in the given method by chromosomes. These chromosomes contain the information about values of network weigh coefficients from the created structure. Chromosome contains G genes, which store the information about weight values and biases of all neurons from the created network. In the developed method a valid coding is used. It is used to represent values of weight coefficients as chromosomes. The best way to code the weight (taking into account the topology we have) is a matrix coding. The matrix size is equivalent to ANN matrix topology. The elements of the matrix are weight coefficients of corresponding connection. The length of chromosome is calculated by the formula (1):

$$\lambda = Q_1(T+1) + \sum_{l=1}^{L} Q_l(Q_{l-1}+1), \quad (1)$$

where Q_l – the number of the neurons from the first layer; T – the number of elements in the learning sample; L – the number of neural network layers [14]

After forming, the rating of the initial population chromosomes starts. To do this, firstly we need to decode every chromosome from the population into the set of neural network weight coefficients. Then, the calculation of fitness-function value starts. This function rates the quality of the selected architecture by the value of neural network learning mistake using these formulae (2) and (3):

$$F_{opt} = \min(F), \quad (2)$$

$$F = \frac{1}{2C_t} \sum_{i=1}^{C_t} (O_i - R_i)^2, \quad (3)$$

where C_i –the power of learning pairs; O_i – the resulting value, which was received by output neuron in the network on an i^{th} step while learning; R_i – the required value of output neuron on an i^{th} step while learning [15].

As we can see, the fitness function will rate the mistake (this mistake defines how the resulting output differs from the needed) as the difference between the resulting output of the network and the needed one. So, as the smaller this mistake is the higher the value of fitness is. In other words, we need to set the matrix to the state, when the neural network learning mistake will be minimum.

According to the given results of the fitness function, a selection of individual starts. It is necessary to generate new solutions, which based on rank selection. To do this, the current population is sorted according to fitness function results. Every chromosome assigned with a number, which shows not its absolute value but its place in the sorted population rank. This approach lets to control selective pressure and to limit the number of offspring of one chromosome. The coefficients of selection pressure [16-17].

Linear ranking of all chromosome is calculated by the formula (4):

$$P_{sl}(Pos) = 2 - SP + 2 \cdot (SP - 1) \cdot \frac{Pos - 1}{C - 1}, \quad (4)$$

where Pos – the position of chromosome in population (Chromosome with the smallest value of the fitness function (F_{min}) has Pos =1, and chromosome with the largest value of the fitness function (F_{max}) has Pos=C value); C – the number of chromosome (individuals) in population; SP – the coefficient of selection pressure, which can be calculated by the formula (5):

$$SP = \frac{F_{\max}}{F_{avg}}, \quad (5)$$

where F_{max} – a chromosome with the largest value of the fitness function, F_{avg} – an average value the fitness function of all population [18-19].

The advantage of the ranking method is an opportunity to use it both for function maximisation and minimization. It also doesn't require scaling due to the problem of pre-timing repeatability, what is very topical for the roulette wheel method.

The crossover operator from the given method produces two offspring from two parental chromosomes. In other words, two new vectors are generated from the two vectors with valid numbers. SBX-crossover is used as a base. This crossover imitates the work of a binary crossover operator. If P^1=

(p_1^1, $p^1{}_2$,…,$p^1{}_\lambda$) and P^2=(p_1^2, $p^2{}_2$,…,$p^2{}_\lambda$) are chromosomes of two parents, then the genes of offspring's chromosomes are calculated with formulae (6) and (7):

$$c_j^1 = \frac{1}{2}\left((1-\omega)\cdot p_j^1 + (1+\omega)\, p_j^2\right), \tag{6}$$

$$c_j^2 = \frac{1}{2}\left((1+\omega)\cdot p_j^1 + (1-\omega)\, p_j^2\right), \tag{7}$$

where j=1,2, …, λ; λ – is the length of chromosome; p_j^1, p_j^2 – are the first and the second parents' genes; ω – the number, which is calculated by the formula (8):

$$\omega = \begin{cases} (2v)^{\frac{1}{b+1}}, v \le 0,5 \\ \left(\frac{1}{2(1-v)}\right)^{\frac{1}{b+1}}, v > 0,5 \end{cases}, \tag{8}$$

where ν – is a normally distributed random number, ν∈(0,1); b – is a value, that affects the probability of offspring appearance far from parents b∈[2,5]. During the research, it was found that low b values allow to generate offspring which are far located from parents, while increased b influences the offsrping to appear closer to parents' pairs [20-21].

The mutation operator in evolutionary algorithms shows the wide searching and capturing new areas of searching space. To set up this algorithm, a user has to additionally decide what the value for mutation should be used. With low mutation, the algorithm will rarely capture new areas of searching space and will more quickly converge to local extrema. With high mutation, the algorithm will more frequently explore different areas of searching space but it will no localize perspective areas of searching space. Accordingly, the selection of mutation operator is directly affecting the quality of algorithm work and results.

To solve this problem, we realized a method to calculate the possibility of mutation change. At first, method searches the level of population development on every iteration. In other words, the average fitness-function value of all individuals. After that, the result is compared to the fitness function value of every individuals. If fitness function value of an individual is lower than the average fitness function value of all individuals, then the possibility of mutation grows. The possibility of chromosome mutation is calculated by the formula (9):

$$\beta_{mut} = \begin{cases} 0.5\dfrac{F_{\max}(e) - F(P(e))}{F_{\max}(e) - F(e)}, if\ F(P(e)) \ge \bar{F}(e) \\ 0.5, if\ F(P(e)) < \bar{F}(e) \end{cases}, \tag{9}$$

where F_{max}(e) – is a fitness function maximum value on the current population, F(e) – is an average fitness function value of the current population [22].

If p_j gen can mutate, a new changed value p_j^{new} can be calculated by the formula (10):

$$p_j^{new} = p_j + b\left(o_j - p_j\right)\left(1 - \frac{e}{e_{\max}}\right)^{\xi}, \tag{10}$$

where b – is a random number in the interval [0,1]; o_j – randomly generated number from a set $\{p_{min}, p_{max}\}$, where p_{min}, i p_{max} – is a low and high border of a possible change of p_j value; e – is a number of current generation; e_{max} – is a maximum number of generations; ξ – is a refinement parameter, that is depends on repeatability type of international process. After reaching steady state, when the best of individuals hasn't changes during all generations, the value will halve. The result is, the searching area expands and it overcome local extrema traps [23-24].

Such adaptive mutation in genetic algorithm (evolution) realisation allows to stick to the necessary balance between two multiscale gene changes (mutations), as during first steps we have, in the main, multiscale changes (that provide wide searching area). Meanwhile, on the final steps (by reduced scale of mutation) the decision becomes more accurate.

It is necessary to use adaptive mutation for every individual, but it is not enough to prevent repeatability of population in local optimum. Let us assume, that i^{th} individual has a potential to mutate (the possibility is 97%). On the one hand, such a high birth mutation from i^{th} individual means that a chromosome can be born with random parameters. The parameters don't have any helpful information about previous evolution. On the other hand, i^{th} individual stays in the population because of the high fitness function value, displacing future generations with lower values. To avoid repeatability completely, let initiate selection which will thin population out and sift individual with a high mutation possibility by formula (11):

$$if\ \beta_{mut}\left(P^i\right) > \delta : removed\ P^i, \tag{11}$$

where β_{mut} – is a mutation possibility of P^i chromosome; δ – is a selection border, which shows the degree of mutating and sifts 'life-threating' individuals [25].

The developed method stops its work when reaches the maximum number of functional epochs. (User has inputted this number)

The proposed genetic method can completely explore searching area, avoid local extrema on genetic searching step and successfully use found 'good' decisions (in other words, it can improve results steadily using intermediate decisions).

5. The results of the algorithm

To build a prediction model of flue gas desulfurization efficiency we used experimental data of a coal boiler with 1000 MW power (from 07:00 24 of May 2020 to 06:00 31 of December 2020). The set has 5 attributes and 5330 examples:

- SO_2 inlet (mg/m^3);
- water consumption in an absorber (m^3/ hour);
- lime consumption (ton/hour);
- secondary reagent consumption (ton/hour);
- SO_2 outlet (mg/мm^3) [6].

To solve the problem, we selected Python-based IDE because it is easy to work with arrays and datasets there. We used NumPy library (Python package for scientific computing). To build a neural network model and to work with it, we chose Keras library [26] and Theano library [27].

The important requirement for data modelling is quality. If data have 'noise', seasonal component, outliers, gaps, then it will negatively effect on the accuracy of prediction and the quality of the models. Also, data (which are for learning datasets for ANN) should be normalized to decrease inaccuracy and to increase the quality of training.

Fresh data processing before transferring to the model have 6 steps [28]:

- to clean datasets with empty or unidentified fields
- to process gaps in predictor's data;
- to process abnormalities in predictor's data;
- to delete seasonal components from time series;
- to convers data to types used in calculations;
- data normalization.

While data gaps processing, empty values are replaced by a median value that is calculated by the formula (12):

$$M_e = X_{Me} + i_M \frac{\frac{\sum f}{2} + S_{Me-1}}{f_{Me}}, \tag{12}$$

where X_{Me} – is a low value of median interval; i_M – the median interval; S_{Me} – the sum of observation which accumulated before median interval; f_{Me} – the number of observations in median interval. In this way, we have the minimum statistical error which depends on rows values. The processing of abnormalities in data is a cleaning from abnormally high or low values. Time series can be cleaned from seasonal components by the decomposition method [29].

The normalized value of x is calculated by the formula (13):

$$z(x) = \frac{x - \min(x)}{\max(x) - \min(x)}. \quad (13)$$

The created network learnt for 100 epochs. While testing, we used the method of dividing the data sample into learning and testing samples in percentage 75/25%.

To rate the quality of prediction models we used Mean Absolute Error (MAE) and Mean Square Error (MSE) [30-31].

Figure 1 and Figure 2 shows the results of a fully-connected perceptron with one hidden layer.

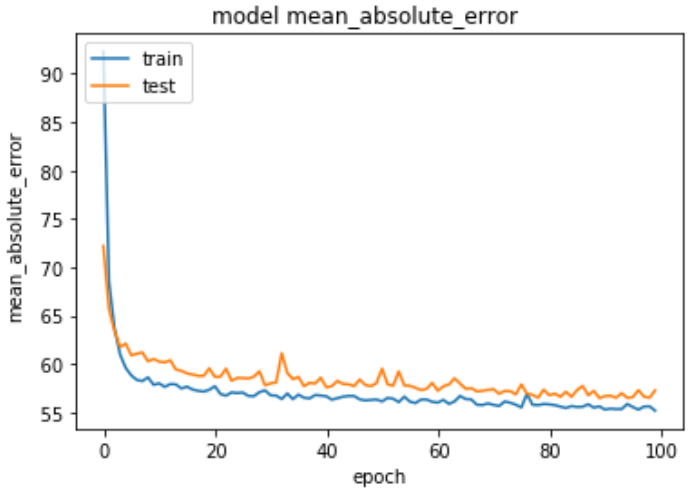


Figure 1: Network metrics value (MAE) for the fully-connected perceptron with one hidden layer

Figure 1 shows, that the best MAE value of this model is 59.85% and we can see that after 10th epochs the value remained constant during the whole learning process of the fully-connected perceptron with one hidden layer.

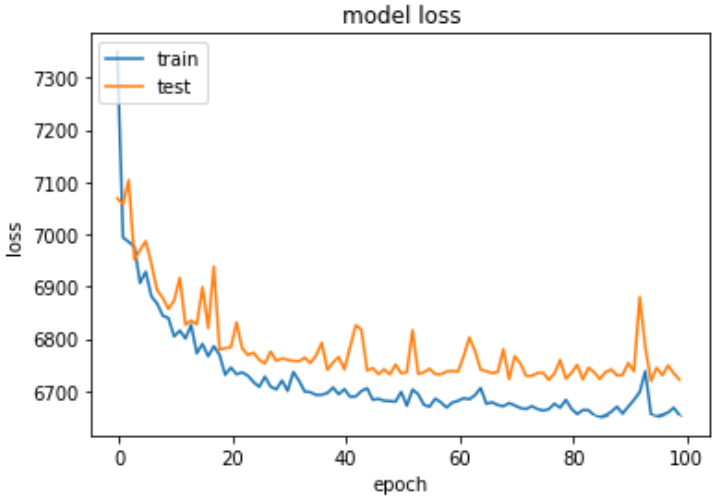


Figure 2: Network metrics value (MSE) for the fully-connected perceptron with one hidden layer

Figure 2 shows, that MSE value gradually declined and after 20th epoch it stayed almost the same and equals 6785.

Figure 3 and Figure 4 shows the work results of a fully-connected perceptron with two hidden layers.

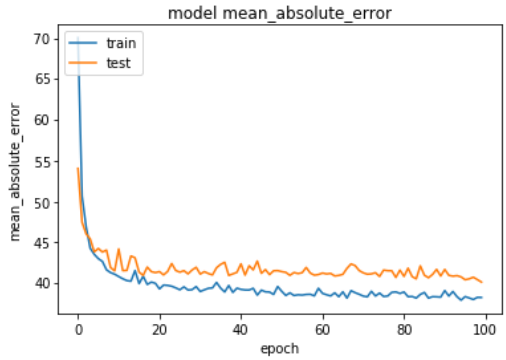


Figure 3: Network metrics value (MAE) for the fully-connected perceptron with two hidden layers

Figure 3 shows, that MAE value of the model equals 40.98 and, after 20th epoch, remained constant during the whole learning process of the fully-connected perceptron with two hidden layers.

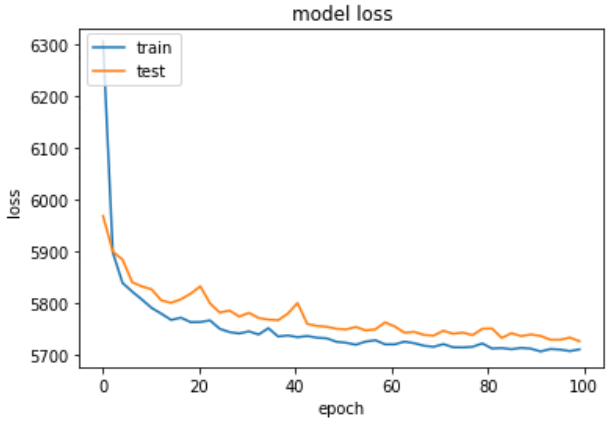


Figure 4: Network metrics value (MSE) for the fully-connected perceptron with two hidden layers

Figure 4 shows that value steadily go down between 10 and 15 epochs, reaches local minima and equals 5795. The further decreasing of the network mistakes says that minima mistake is local.

Figure 5 and Figure 6 shows work results of a fully-connected perceptron with two hidden layers and the genetic method with adaptive mutation.

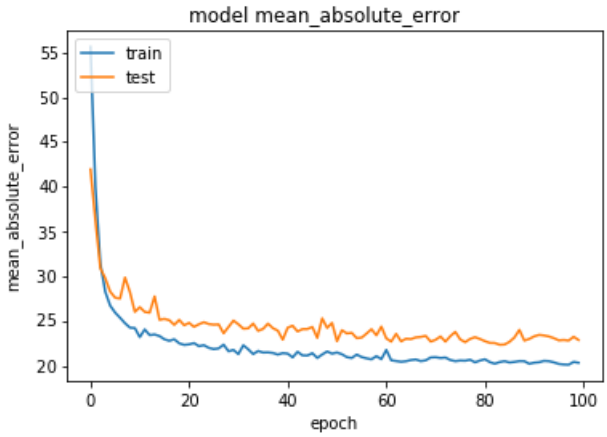


Figure 5: Network metrics value (MAE) for the fully-connected perceptron with two hidden layers and the genetic method with adaptive mutation

Figure 5 shows that MAE results for the fully-connected perceptron with two hidden layers and the genetic method with adaptive mutation equals 23.95 and, starting from 5th epoch stayed almost the same. So, we can say, that during the training the model reached its mistake minima and the network is considered to be trained.

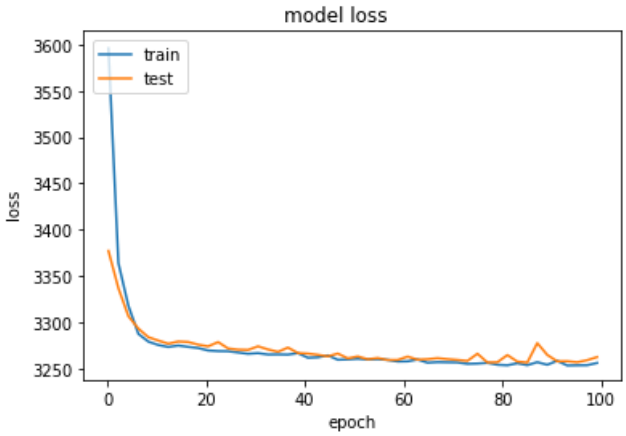


Figure 6: Network metrics value (MSE) for the fully-connected perceptron with two hidden layers and the genetic method with adaptive mutation

Figure 6 shows that value of the mistake slightly decreased. During 20th epoch it stayed almost the same and equals 3253.

There was a comparison exercise of developed models results with the methods as follows: linear regression, polynomial regression, logistic regression, nearest neighbour algorithm, random forest, ant colony optimization algorithm. The criterion of estimation were Mean Absolute Error (MAE), Mean Squared Error (MSE) and runtime (Table 1).

Table 1

Comparison exercise of developed method for prediction the flue gas desulfurization efficiency

Method	MAE	MSE	Runtime, seconds
Linear regression	79.24	9103	363
Polynomial regression	77.84	9031	401
Logistic regression	78.56	9041	374
Nearest neighbour algorithm	75.76	7421	428
Random forest	74.15	7201	469
Ant colony optimization algorithm	73.89	7523	393
Multilayer perceptron with one hidden layer	59.85	6785	864
Multilayer perceptron with two hidden layers	40.98	5795	964
Multilayer perceptron with two hidden layers and the developed genetic algorithm	24.95	3253	564

Table 1 shows, that the use of neural networks gives more accuracy than the use of regression models. For example, the average absolute mistake values of the multilayer perceptron with one or two hidden layers respectively equal 59.85 and 40.98. Meanwhile, the average absolute mistake values of linear and logistic regression respectively equal 79.24 and 78.56. So, neural networks can achieve more accuracy in prediction than regression models as they can better process a non-linear behaviour. But neural networks need more runtime to stop working. The runtime of the multilayer perceptron with one hidden layer equals 864 seconds while the runtime of the polynomial regression equals 401 seconds. To solve this problem, the modified genetic method with adaptive mutation was developed. It allows to increase the accuracy of prediction and to reduce the time needed for neural network learning. While testing, it showed the lowest MSE value, that equals 3253. It is 54% less than the random forest method

has and 43% more than the multilayer perceptron with two hidden layers has. Also, the runtime of the multilayer perceptron with two hidden layers and the genetic algorithm equals 564 seconds, that is 34% less than the multilayer perceptron with one hidden layer has and 41% more than the multilayer perceptron with two hidden layers has. According to the results, we can sum up that it is perspective to use evolutionary algorithms for ANN learning, as using them can reduce learning time and achieve more deeper minima of ANN learning mistake.

6. Conclusion

The work described the solution to optimization problem of flue gas desulfurization process strategy using heuristic methods. The modified genetic method for prediction of flue gas desulfurization based on the created neural network was developed. The vital difference between this developed genetic method and other similar methods is in using adaptive mutation, that uses the level of population development in working process. It means that less important genes will mutate in chromosome more probable than high suitability genes. The worse an individual is adapted the far it is located from the optimum which means it has more possibility to mutate, in which case it will move far away from the current non-optimal location. The proposed genetic method can completely explore a searching area, avoid local extrema on the genetic searching step and successfully use found 'good' decisions (in other words, it can improve results steadily using intermediate decisions). Also, it represents the influence of time series characteristic on factual efficiency of desulfurization and increases its precision of prediction. There was a comparison exercise of the proposed method and others. During the work, it was found that the developed method gives the smallest prediction mistake in prediction of SO_2 composition in the outlet after flue gas desulfurization (MAE equals 24.95 and MSE equals 3253, what is less comparing to other models built on different methods) and also has smaller runtime (learning time of the model equals 564 seconds). The practical use of the developed method can increase the efficiency of flue gas desulfurization process and decrease SO_2 emissions into the atmosphere.

7. References

[1] S. Liu, L. Sun, S. Zhu, J. Li, X. Chen, W. Zhong, Operation strategy optimization of desulfurization system based on data mining, Applied Mathematical Modelling 81 (2020) 144-158. doi: 10.1016/j.apm.2019.12.004.

[2] Z. Shao, F. Si, D. Kudenko, P. Wang, X. Tong, Predictive scheduling of wet flue gas desulfurization system based on reinforcement learning, Computers & Chemical Engineering 141 (2020) p. 107000. doi: 10.1016/j.compchemeng.2020.107000.

[3] J.A.J. Alsayaydeh, W.A.Y. Khang, W.A. Indra, V. Shkarupylo, J. Jayasundar, Development of smart dustbin by using apps, ARPN Journal of Engineering and Applied Sciences 14 (2019) 3703-3711.

[4] X. Li, Q. Liu, K. Wang, F. Wang, G. Cui, Y. Li, Multimodel Anomaly Identification and Control in Wet Limestone-Gypsum Flue Gas Desulphurization System, Complexity 2020 (2020) 1-17. doi: 10.1155/2020/6046729.

[5] J. Fu, H. Xiao, T. Wang, R. Zhang, L. Wang, X. Shi, Prediction Model of Desulfurization Efficiency of Coal-Fired Power Plants Based on Long Short-Term Memory Neural Network, in: 2019 International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), 2019, pp. 40–45. doi. 10.1109/iThings/GreenCom/CPSCom/SmartData.2019.00030

[6] D.Adams, D. Oh, D. Kim, C. Lee, M. Oh, Prediction of SOx–NOx emission from a coal-fired CFB power plant with machine learning: Plant data learned by deep neural network and least square support vector machine, Journal of Cleaner Production 270 (2020) p. 122310. doi: 10.1016/j.jclepro.2020.122310.

[7] Z. Qiao, X. Wang, H. Gu, Y. Tang, F. Si, C. Romero, X. Yao, An investigation on data mining and operating optimization for wet flue gas desulfurization systems, Fuel 258 (2019) p.116178. doi: 10.1016/j.fuel.2019.116178.

[8] J.A.J. Alsayaydeh, W.A.Y. Khang, W.A. Indra, J.B. Pusppanathan, V. Shkarupylo, A.K.M. Zakir Hossain, S. Saravanan, Development of vehicle door security using smart tag and fingerprint system, ARPN Journal of Engineering and Applied Sciences 9 (2019) 3108-3114. doi: 10.35940/ijeat.E7468.109119.
[9] G. Uddin et al., Artificial Intelligence-Based Emission Reduction Strategy for Limestone Forced Oxidation Flue Gas Desulfurization System, Journal of Energy Resources Technology 142 (2020) 1-16. doi: 10.1115/1.4046468.
[10] Y. Guo et al., Modeling and optimization of wet flue gas desulfurization system based on a hybrid modeling method, Journal of the Air & Waste Management Association 69 (2019) 565-575. doi: 10.1080/10962247.2018.1551252.
[11] J.A.J. Alsayaydeh, W.A.Y. Khang, A.K.M.Z Hossain, V. Shkarupylo, J. Pusppanathan, The experimental studies of the automatic control methods of magnetic separators performance by magnetic product, ARPN Journal of Engineering and Applied Sciences 15 (2020) 922-927.
[12] Z. Kong, Y. Zhang, X. Wang, Y. Xu, B. Jin, Prediction and optimization of a desulphurization system using CMAC neural network and genetic algorithm, Journal of Environmental Engineering and Landscape Management 28 (2020) 74-87. doi: 10.3846/jeelm.2020.12098.
[13] I. Fedorchenko, A. Oliinyk, A. Stepanenko, T. Zaiko, S. Shylo, A. Svyrydenko, Development of the modified methods to train a neural network to solve the task on recognition of road users, Eastern European Journal of Enterprise Technologies 98 (2019) 46–55. doi: 10.15587/1729-4061.2019.164789.
[14] E. Pardo, J. Blanco-Linares, D. Velázquez, F. Serradilla, Optimization of a Steam Reforming Plant Modeled with Artificial Neural Networks, Electronics 9 (2020) 1923. doi: 10.3390/electronics9111923.
[15] J.A. Alsayaydeh, M. Nj, S.N. Syed, A.W. Yoon, W.A. Indra, V. Shkarupylo, C. Pellipus, Homes appliances control using bluetooth, ARPN Journal of Engineering and Applied Sciences 14 (2019) 3344-3357.
[16] S. Stajkowski, D. Kumar, P. Samui, H. Bonakdari, B. Gharabaghi, Genetic-Algorithm-Optimized Sequential Model for Water Temperature Prediction, Sustainability 12 (2020) 5374. doi: 10.3390/su12135374.
[17] A. Oliinyk, S. Skrupsky, S. Subbotin, Experimental research and analysis of complexity of parallel method for production rules extraction, Automatic Control and Computer Sciences 52 (2018) 89-99. doi: 10.3103/S0146411618020062.
[18] Z. Erzurum Cicek, Z. Kamisli Ozturk, Optimizing the artificial neural network parameters using a biased random key genetic algorithm for time series forecasting, Applied Soft Computing 102 (2021) 107091. doi: 10.1016/j.asoc.2021.107091.
[19] A. Zaji, H. Bonakdari, H. Khameneh, S. Khodashenas, Application of optimized Artificial and Radial Basis neural networks by using modified Genetic Algorithm on discharge coefficient prediction of modified labyrinth side weir with two and four cycles, Measurement 52 (2020) 107291. doi: 10.1016/j.measurement.2019.107291.
[20] H. Cheng, J. Xie, Study on the application of recurrent fuzzy neural network in PH control system of absorption tower, in: 2017 Chinese Automation Congress (CAC). doi: 10.1109/cac.2017.8243850.
[21] A. Oliinyk, I. Fedorchenko, A. Stepanenko, M. Rud and D. Goncharenko, Combinatorial optimization problems solving based on evolutionary approach, in: 2019 15th International Conference on the Experience of Designing and Application of CAD Systems (CADSM), Polyana, Ukraine, 2019, pp. 41-45. doi: 10.1109 / CADSM.2019.8779290.
[22] X. Wang, L. Yang, X. Chen, J. Han, J. Feng, A Tensor Computation and Optimization Model for Cyber-Physical-Social Big Data, IEEE Transactions on Sustainable Computing 4 (2019) 326-339. doi: 10.1109/tsusc.2017.2777503.
[23] Z. Yang et al., Predicting particle collection performance of a wet electrostatic precipitator under varied conditions with artificial neural networks, Powder Technology 377 (2021) 632-639. doi: 10.1016/j.powtec.2020.09.027.
[24] I. Fedorchenko, A. Oliinyk, A. Stepanenko, T. Zaiko, S. Korniienko, N. Burtsev, Development of a genetic algorithm for placing power supply sources in a distributed electric network, Eastern-

European Journal of Enterprise Technologies 5 (2019) 6-16. doi: 10.15587 / 1729-4061.2019.180897.
[25] F. Wang et al., Application of genetic algorithm-back propagation for prediction of mercury speciation in combustion flue gas, Clean Technologies and Environmental Policy 18 (2016) 1211-1218. doi: 10.1007/s10098-016-1095-1.
[26] J.A.J. Alsayaydeh, W.A. Indra, W.A.Y. Khang, V. Shkarupylo, D.A.P.P. Jkatisan, Development of vehicle ignition using fingerprint, ARPN Journal of Engineering and Applied Sciences 14 (2019) 4045-4053.
[27] Z. Kong et al., Error prediction and structure determination for CMAC neural network based on the uniform design method, Expert Systems 38 (2020). doi: 10.1111/exsy.12614.
[28] A. Oliinyk, I. Fedorchenko, A. Stepanenko, A. Katschan, Y. Fedorchenko, A. Kharchenko, D. Goncharenko, Development of genetic methods for predicting the incidence of volumes of pollutant emissions in air, in: 2019 2nd International Workshop on Informatics and Data-Driven Medicine (IDDM), 2019, pp. 340-353. ISSN: 16130073.
[29] I. Fedorchenko, A. Oliinyk, O. Stepanenko, T. Zaiko, A. Svyrydenko, D. Goncharenko, Genetic method of image processing for motor vehicle recognition, in: 2nd International Workshop on Computer Modeling and Intelligent Systems, CMIS 2019, CEUR Workshop Proceedings, Zaporizhzhia, Ukraine, 2019, pp. 211-226. ISSN: 16130073.
[30] Q. Li, J. Wu, H. Wei, Reduction of elemental mercury in coal-fired boiler flue gas with computational intelligence approach, Energy 160 (2018) 753-762. doi: 10.1016/j.energy.2018.07.037.
[31] H. Jang, X. Shuli, S. So, Analysis the Compressive Strength of Flue Gas Desulfurization Gypsum Using Artificial Neural Network, Journal of Nanoscience and Nanotechnology 20 (2020) 485-490. doi: 10.1166/jnn.2020.17235.