EKSPLOATACJA I NIEZAWODNOŚĆ MAINTENANCE AND RELIABILITY

Eksploatacja i Niezawodnosc – Maintenance and Reliability

Volume 24 (2022), Issue 1

journal homepage: http://www.ein.org.pl

Article citation info:
Chen X, Przystupa K, Ye Z, Chen F, Wang C, Liu J, Gao R, Wei M, Kochan O. Forecasting short-term electric load using extreme learning machine with improved tree seed algorithm based on Lévy flight. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2022; 24 (1): 153–162, http://doi.org/10.17531/ein.2022.1.17.

Forecasting short-term electric load using extreme learning machine with improved tree seed algorithm based on Lévy flight

Indexed by: Web of Science Group

Xuan Chen [a], Krzysztof Przystupa [b,*], Zhiwei Ye [a], Feng Chen [a], Chunzhi Wang [a], Jinhang Liu [a], Rong Gao [a], Ming Wei [c,d], Orest Kochan [a,d]

[a] Hubei University of Technology, School of Computer Science, Wuhan, Hubei, 430068, PR China
[b] Lublin University of Technology, Department of Automation, ul. Nadbystrzycka 36, 20-618 Lublin, Poland
[c] Wuhan Fiberhome Technical Services Co., Ltd., Wuhan, Hubei, 430074, PR China
[d] Wuhan FiberHome Telecommunication Technologies Co., Ltd., Wuhan, Hubei, 430074, PR China
[e] Lviv Polytechnic National University, Bandery 12, 79013 Lviv, Ukraine

Highlights

- Forecasting Short-term Electric Load using Extreme Learning Machine is considered.
- Improved Tree Seed Algorithm based on Lévy flight is proposed.
- This method has better convergence, prediction accuracy and stability than similar ones.

Abstract

In recent years, forecasting has received increasing attention since it provides an important basis for the effective operation of power systems. In this paper, a hybrid method, composed of kernel principal component analysis (KPCA), tree seed algorithm based on Lévy flight (LTSA) and extreme learning machine (ELM), is proposed for short-term load forecasting. Specifically, the randomly generated weights and biases of ELM have a significant impact on the stability of prediction results. Therefore, in order to solve this problem, LTSA is utilized to obtain the optimal parameters before the prediction process is executed by ELM, which is called LTSA-ELM. Meanwhile, the input data is extracted by KPCA considering the sparseness of the electric load data and used as the input of LTSA-ELM model. The proposed method is tested on the data from European network on intelligent technologies (EUNITE) and experimental results demonstrate the superiority of the proposed approaches compared to the other methods involved in the paper.

Keywords

short-term electric load forecast; extreme learning machine; Lévy flight; tree-seed algorithm; Kernel principal component analysis.

This is an open access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/)

1. Introduction

High quality and uninterrupted power supply are indispensable today [25, 40], because many processes and systems are integrated into a global network using Internet of Things (IoT) concept [15, 48]. They need a reliable power supply for proper operation [1, 42]. Improper operation, failures and accidents in the power supply network not only directly affect the operation of devices and systems, but also can indirectly affect the modes of their operation [34, 49]. In general, reliability in technology and industry is very important [41], as a consequence, a lot of attention is paid to the development of methods of preventing equipment failure [14], its diagnostics and early failure detection [12, 22, 51].

However, the development of science and technology needs more research, especially in the estimation of equipment lifespan, the forecast of different modes of its operation [27, 28, 35] and process control methods [7, 24]. Regression analysis is traditionally used for prediction [26, 60], however, the methods of artificial intelligence have been actively used recently in this area [13, 44]. This paper is devoted to the use of such methods.

Short-term electric load forecasting generally refers to forecasts within a year and in units of months, weeks, days or hours, including ultra-short-term load forecasts with a period of one hour or even a few minutes. The accuracy of forecasting is directly related to the normal use of electricity by customers. The improvement of electric load forecasting technology not only improves the safety and reliability of grid operation but also reduces the cost of power generation and maximizes the benefits [39]. Major methods for short-term electric load forecasting are broadly classified into traditional forecasting methods and neural network methods [18, 32]. Traditional forecast-

(*) Corresponding author.
E-mail addresses: X. Chen (ORCID:0000-0003-1343-7132): xuan927@hbut.edu.cn, K. Przystupa (ORCID: 0000-0003-4361-2763): k.przystupa@pollub.pl, Z. Ye (ORCID: 0000-0001-6668-4634): hgcsyzw@mail.hbut.edu.cn, F. Chen (ORCID: 0000-0001-9517-1301): chenfeng4179@fiberhome.com, C. Wang (ORCID: 0000-0002-6742-3644): chunzhiwang@mail.hbut.edu.cn, J. Liu (ORCID:0000-0002-5509-0846): liujinhang@hbut.edu.cn, R. Gao (ORCID:0000-0001-7935-7173): gaorong@hbut.edu.cn, M. Wei (ORCID: 0000-0003-4695-5657): mwei@fiberhome.com, O. Kochan (ORCID: 0000-0002-3164-3821): orest.v.kochan@lpnu.ua

ing methods include regression analysis [33] and expert systems [37]. These methods are difficult to build effective mathematical models and fit highly nonlinear multi-factor electric loads. Neural network methods contain regression trees [57], grey prediction [65], support vector machines [38] and artificial neural networks. Artificial neural network has satisfied fault tolerance rate, nonlinear mapping ability, adaptive learning ability and efficiency compared with other forecasting methods. As a result, artificial neural network is widely used in the field of electric load forecasting for solving complex nonlinear practical problems [29].

Extreme learning machine (ELM) has been widely used in the field of electric load forecasting as one of artificial neural networks [4, 9]. Chen [3] proposed a hybrid intelligent optimization algorithm to optimize the parameters of ELM. The method achieves high accuracy for short-term electric load forecasting. ELM is feed-forward neural network with a single hidden layer [20] and has the advantages of simple structure, fast training speed, fewer adjustment parameters, strong generalization and nonlinear approximation ability. However, there are some defects of ELM in practical applications [54]. Impact on prediction results as the setting of relevant parameters in ELM includes three aspects according to the literature [21]. Firstly, the choice of activation function in different instances influences the prediction results. The second is the predefined number of hidden layer nodes in ELM increases the subjectivity of the process of determining the hidden layer nodes in the network. Finally, the randomly defined input weight matrix and hidden layer thresholds in ELM lead to the failure of partially hidden layer nodes, less specific sample learning and unstable performance.

Some evolutionary algorithms have been used to solve these problems. For instance, particle swarm optimization algorithm (PSO) [61], grey wolf optimizer algorithm (GWO) [10], moth flame optimization algorithm (MFO) [56], differential evolution algorithm (DE) [19], cuckoo search algorithm (CS) [63] and harmony search algorithm (HS) [43] were used to optimize ELM. For parameter optimization of ELM, Wei [52] presented a prediction method based on MFO optimizing the parameters of random forest and ELM to forecast CO2 emissions. Yang [58] proposed a differential evolution-based feature selection and parameter optimization for ELM in tool wear estimation. Wang [55] proposed a hybrid model based on CS algorithm to forecast solar radiation. These works achieved good results. Nonetheless, some evolutionary algorithms have complicatedly operators and constantly adjusted parameters, which results in ineffective convergence to the global optimum. Therefore, trying more new algorithms to deal with this problem is necessary. The tree seed algorithm (TSA) [5] is one of the newest heuristic algorithms. Compared with other heuristic algorithms, TSA is easy to implement, has fewer tuning parameters and takes less time to compute. This method was successfully used to solve different engineering optimization problems. Ali [8] adopted TSA to solve the optimal power flow problem in large-scale power systems. Zhao [64] utilized residual vectors and TSA to identify the structural damage. Muneeswaran [36] developed a performance evaluation method for the radial basis function neural network based on TSA. However, TSA is affected by the search trend (ST) which leads to the update falling into local optimum. TSA with the Lévy flight is proposed for balancing the global and local search capabilities to obtain a better prediction effect [2].

In the paper, LTSA is introduced to optimize initialization parameters to improve the prediction performance of ELM. This method reduces the training time of network and improves the stability and accuracy. The method of ELM combined with LTSA (LTSA-ELM) is used to predict the electric load data processed by kernel principal component analysis (KPCA). The experimental results show that the hybrid method of KPCA, LTSA and ELM (KPCA-LTSA-ELM) proposed in this paper achieves better prediction results.

The rest of the paper is structured as follows. The basic principle of TSA and ELM is illustrated in section 2. The proposed KPCA-LTSA-ELM is explained in section 3. The experimental results and discussion are demonstrated in section 4. Finally, the paper is concluded in section 5.

2. Preliminary

2.1. Extreme learning machine

ELM is a three-layer network composed of the input layer, hidden layer and output layer. Each layer is connected by neurons. The structure diagram of ELM is depicted in Figure 1.

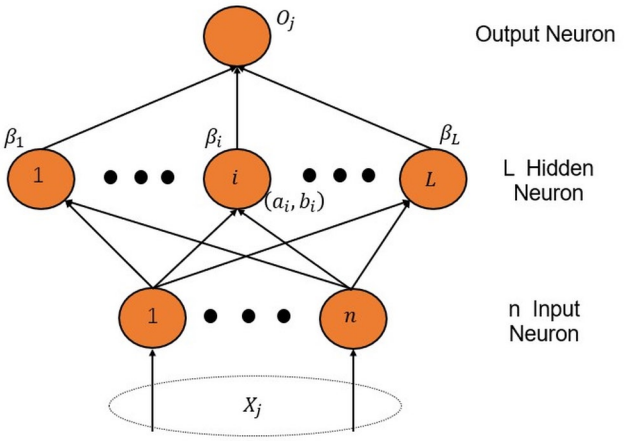

Fig. 1. The structure diagram of ELM

Let us assume the amount of neurons in the hidden layer is L. The standard form of the model is expressed in equations (1-3):

$$t_i = \sum_{j=1}^{L} \beta_j g(w_j x_i + b_j)(i = 1,2,...,N) \tag{1}$$

where L is a number of hidden layer, N is a number of training samples, $g(x)$ is the activation function, w_j is the input weight, β_j is the output weight, b_j is the bias of the j_{th} hidden layer unit, and $w_j x_i$ is the inner product of w_j and x_i.

When the output error is the minimum, it can be calculated by equation (2):

$$\sum_{i}^{N} \|t_i - y_i\| = 0 \tag{2}$$

Thus, there exists β_j, w_j and b_j such that:

$$y_i = \sum_{j=1}^{L} \beta_j g(w_j x_i + b_j)(i = 1,2,...,N) \tag{3}$$

The matrix form of the model is expressed in equations (4-7):

$$H\beta = Y \tag{4}$$

where H is the output matrix of the hidden layer of the neural network, β is the weight output matrix, Y is the target output matrix:

$$H = \begin{bmatrix} g(\omega_1 \cdot x_1 + b_1) & g(\omega_2 \cdot x_1 + b_2) & \cdots & g(\omega_L \cdot x_1 + b_L) \\ g(\omega_1 \cdot x_2 + b_1) & g(\omega_2 \cdot x_2 + b_2) & \cdots & g(\omega_L \cdot x_2 + b_L) \\ \vdots & \vdots & \ddots & \vdots \\ g(\omega_1 \cdot x_N + b_1) & g(\omega_2 \cdot x_N + b_2) & \cdots & g(\omega_L \cdot x_N + b_L) \end{bmatrix}_{N \times L} \tag{5}$$

$$\beta=\left[\beta_1^{\mathrm{T}}\beta_2^{\mathrm{T}}\cdots\beta_{\mathrm{L}}^{\mathrm{T}}\right]_{L\times m} \tag{6}$$

$$Y=\left[Y_1^T Y_2^T\cdots Y_L^T\right]_{N\times m} \tag{7}$$

To forecast a single hidden layer neural network, it can be defined as follows:

$$\left\|H\left(\hat{w}_j,\hat{b}_j\right)\hat{\beta}_j-Y\right\|=\min_{w,b,\beta}\left\|H\left(w_j,b_j\right)\beta_j-Y\right\| \tag{8}$$

where $j=1,2,...,L$, this is equivalent to minimizing the loss function. It can be defined in equations (9-10):

$$E=\sum_{i=1}^{N}(\sum_{j=1}^{L}\beta_j g(w_j x_i+b_j)-y_i)^2 \tag{9}$$

$$\hat{\beta}=H^{+}Y \tag{10}$$

where H^{+} is the Moore-Penrose generalized inverse matrix of the output matrix of the hidden layer H. When the activation function is infinitely differentiable, the connection weight w_j of the input layer to the hidden layer and the threshold b_j of the hidden layer can be randomly assigned before training and remain unchanged during the training. Then the output matrix of the hidden layer H is determined, that is, the connection weight matrix $\hat{\beta}$ of the hidden layer to the output layer can be determined by equation (10).

2.2. Tree seed algorithm

TSA is a new intelligent optimizer proposed by M.S. Kiran in 2015 to solve the continuous optimization problems. In nature, trees spread their seeds to the surfaces. If these surfaces are considered as the search space for the optimization problem, the location of the tree and seeds are the possible solutions to the optimization problem. Therefore, the search for seeds location is an important step in TSA to solve the optimization problem. In the primary TSA, each tree simulates the solution of the optimization problem in the search space. The fitness value of the tree is usually calculated by the objective function or the optimization problem.

Before the algorithm starts searching, equation (11) is used to generate the latest initial tree positions for subsequent iterations of the search. The initial tree positions are the feasible solutions of TSA:

$$T_{i,j}=Low_j+r_{i,j}(High_j-Low_j) \tag{11}$$

where $T_{i,j}$ represents the corresponding value of the j_{th} dimension of the i_{th} tree randomly generated in search space. Low_j and $High_j$ are the lower and upper bounds of the j_{th} dimension, respectively.

The trees generate new seed locations and the number of seeds depends on the size of the population, therefore the number of seeds can exceed one. In the analysis of the influence of controlled variables on the performance of TSA, the minimum quantity of tree seeds is 10% of the population size and the maximum quantity is 25% of the population size. The amount of seeds produced is completely random in TSA.

Then equation (12) is used to optimize the feasible solution for the population obtained in the first batch to select the trees with strong capability for seed production and the optimal location:

$$Best=\min\left\{f\left(\vec{T}_i\right)\right\}i=1,2,...,N \tag{12}$$

where N is the population size of trees in TSA.

The selected trees will continue to update their positions and produce new seeds. There are two search modes for seeds, one of which focuses on the global search and the other on the local one.

Two ideal conditions are usually assumed for the search:

(1) The first update of the new seed location is determined by the position of the tree and the position of the optimal tree in the tree population. This search enhances the local search capability of the algorithm. It can be updated by equation (13).

(2) The second update of the new seed location is determined by two randomly selected trees with different locations. It can be updated by equation (14):

$$S_{i,j}=T_{i,j}+\alpha_{i,j}\times\left(Best_j-T_{r,j}\right) \tag{13}$$

$$S_{i,j}=T_{i,j}+\alpha_{i,j}\times\left(T_{i,j}-T_{r,j}\right) \tag{14}$$

where $S_{i,j}$ is the j_{th} dimension of the i_{th} seed which generated by the i_{th} tree, $T_{i,j}$ is the j_{th} dimension of the i_{th} tree, $Best_j$ is the best tree currently obtained, $T_{r,j}$ is the j_{th} dimension of the r_{th} tree randomly selected from the population, and $\alpha_{i,j}$ is a scale factor randomly generated within the range of [-1, 1].

The choice of the specific update mode of the seed is regulated by the search tendency (ST) parameter within the scope of [0,1]. A larger ST value provides powerful local search capability and faster convergence speed. A smaller ST value results in slower convergence but strong global search capabilities. According to previous experiments [23], when the value of ST is 0.1, most functions can get the optimal solution. If the ST generated randomly within the range of [0, 1] is less than 0.1, then equation (13) is selected to update each dimension of the seed produced by each tree, otherwise equation (14) is used. The procedure of TSA is described in Algorithm 1.

Algorithm 1 Procedure of TSA

Input: The parameters and the termination condition of TSA.
Output: The best solution obtained by TSA.
Step1: The initialization of the algorithm
Randomly generate tree locations in the D-dimensional search space.
Evaluate the tree location by the fitness function.
Select the best solution.
Step2: Search with seeds
FOR all trees
Decide the number of seeds produced for this tree.
FOR all seeds
FOR all dimensions
IF (rand < ST)
Update this dimension using equation **(13)**
ELSE
Update this dimension using equation **(14)**
END IF
END FOR
END FOR
Select the best seed and compare it with the tree.
If the seed location is better than the tree location, the seed substitutes for this tree.
END FOR
Step3: Selection of the best solution
Select the best solution of the population.
If the new best solution is better than the previous best solution, new best solution substitutes for the previous best solution.
Step4: Testing termination condition

If the termination condition is not met, go to **Step 2.**
If the termination condition is met, output the best solution.

3. The proposed method

3.1. Kernel principal component analysis

In order to ease the training of the model to learn the correct load variation law for short-term electric load forecasting, we perform feature extraction on data processed by LTSA-ELM. These re-extracted indicators can reflect the information of the original data as much as possible, simplifying the network structure and improving the training rate of the network.

Since the relationship between the short-term electric load indicators is usually non-linear, the use of linear principal component analysis (PCA) [50, 62] tends to cause the contribution rate of each principal component to be too dispersed to find a component with comprehensive capabilities. KPCA [45] circumvents the unascertained nonlinear transformation in nonlinear principal component analysis (NLPCA) [6] by using the kernel techniques. Therefore, the principal components can be obtained in a more concentrated manner, and the evaluation results are more consistent with objective facts. KPCA is applied to improve the input of LTSA-ELM, which can effectively reduce the input dimension while retaining most of the original information. Therefore, we can predict and analyze load data in the actual power grid and improve the efficiency and precision of the forecast.

In the process of the kernel principal component analysis, the analysis results are related to the choice of kernel function. The proper selection of kernel functions and parameters can effectively improve the overall performance of KPCA. There are two common kernel functions:

The Polynomial kernel can be expressed by equation (15):

$$K(x,y)=\left[s(x,y)+c\right]^{d} \tag{15}$$

The Gaussian kernel can be expressed by equation (16):

$$K(x,y)=\exp(-\frac{\left\|x-y\right\|}{2e^{2}}) \tag{16}$$

In the paper, these two functions are selected for calculations. The kernel parameter in the polynomial kernel is $\mathrm{var}=sign=2$, and the kernel parameter in the Gaussian kernel is $\mathrm{var}=281, sign=1$. The selection of kernel functions is explained in section 4.3.

3.2. Lévy flight distribution introduced to TSA

Lévy flight (LF) [30] is a random walk mode between short-range search and occasional long-range search. Similarly, researchers found that the Lévy flight can also improve the performance of nature-inspired algorithms [46, 47].

The LF distribution generates new solutions by randomly selecting short or long steps. At present, the LF distribution is widely utilized in many fields for improving the exploration ability, because it can increase the variety of species and expand the search range. For example, CS algorithm uses the LF for updating position [59], bat algorithm uses the LF strategy to mimic the predation behavior of bats instead of the speed and position of the original algorithm [31], PSO uses the LF to update particle position after iterating multiple times [16] etc.

The position of LF is updated by equations (17-19):

$$\begin{gathered} x_i^{(t+1)}=x_i^{(t)}+\alpha\oplus Levy(\lambda) \\ i=1,2,\ldots,n \end{gathered} \tag{17}$$

$$Levy(\lambda)\sim u=t^{-\lambda} \tag{18}$$
$$1<\lambda\le 3$$

$$Levy(\lambda)\sim\frac{\phi*\mu}{\left|\upsilon\right|^{\frac{1}{\beta}}} \tag{19}$$

where, μ,υ follow the standard normal distribution β=1.5 . The mathematical definition formula of ϕ can be expressed by equation (20):

$$\phi=\left(\frac{\Gamma(1+\beta)*\sin(\pi\times\frac{\beta}{2})}{\Gamma\left(\left(\frac{1+\beta}{2}\right)*\beta*2^{\frac{(\beta-1)}{2}}\right)}\right)^{\frac{1}{\beta}} \tag{20}$$

where Γ is the standard Gamma function.

To improve the global and local search capabilities of TSA, LTSA combines the advantages of the LF distribution with TSA for overcoming the problem of trapping into a local optimum. In the case of rand<ST, the LF random walk strategy is introduced when the seed is updated. The seed update rule is changed from equation (13) to equation (21):

$$S_{i,j}=T_{i,j}+\alpha_{i,j}\oplus Levy\left(\lambda\right)\times\left(Best_j-T_{r,j}\right) \tag{21}$$

where Levy(λ) represents a random search vector whose jump step obeys the Lévy distribution, $\lambda\left(1<\lambda\le 3\right)$ is a scale parameter.

The choice of the fitness function directly affects the rate of convergence of the tree algorithm and whether the best solution can be found. In the evolutionary search, the algorithm calculates the individual fitness values according to the fitness function, and fitness value is used to evaluate the pros and cons of the individual tree.

The tree individual $\vec{T}_i$ in the population corresponds to the arrangement of the weights and thresholds of ELM network.

The value of the fitness function can be calculated by equation (22):

$$fitness(\vec{T}_i)=E(\vec{T}_i)=\frac{1}{2}\sum_{k=1}^{m}\left(y_k-o_k\right)^2 \tag{22}$$

where: o_k is the actual output of the output layer neurons, y_k is the expected output of the output layer neurons, and m is the number of output layer neurons. According to the error function, the weight and threshold between each neuron are continuously adjusted to train the neural network to achieve the optimal solution.

The procedure of LTSA-ELM is described in Algorithm 2.

Algorithm2 Procedure of LTSA-ELM

Input: The best solution obtained by LTSA
Output: The predicted results obtained by LTSA-ELM
Step 1: Set the best solution
The best solution is the connection weight for the input layer and hidden one and the neuron threshold of the hidden layer.
Step 2: Select an activation function to calculate the output matrix H of the hidden layer neurons.

Step 3: Calculate a connection weight matrix $\hat{\beta}$ of the hidden layer to the output layer.
Step 4: Predict the results by the output weight matrix and activation function.

The flowchart of KPCA-LTSA-ELM model is shown in Figure 2.

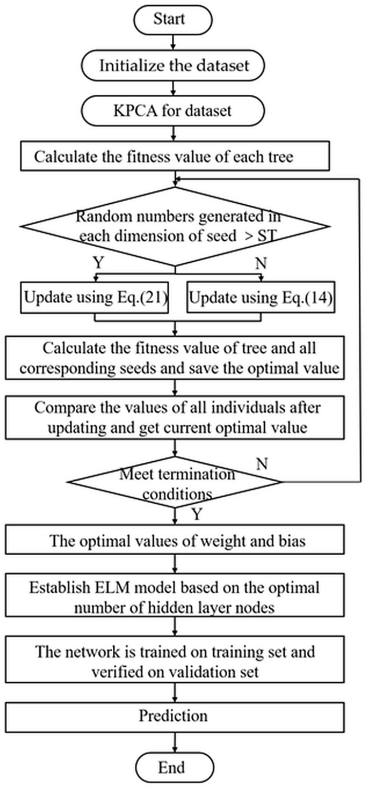

Fig. 2. The flowchart of KPCA-LTSA-ELM model

4. Experiments and Discussions

The experiment verifies the model proposed in the paper from three aspects. First, the stability and accuracy of LTSA-ELM prediction were verified by experiments on six UCI public datasets using five common methods (section 4.1). Second, KPCA-LTSA-ELM is applied to data provided by EUNITE (section 4.2). Third, the relevant parameters of LTSA-ELM are selected. Then PCA, KPCA based on the Polynomial kernel and KPCA based on the Gaussian kernel are applied to evaluate eleven characteristic attributes in data provided by EUNITE. The experimental results indicate that KPCA based on the Gaussian kernel should be used preferentially in dimensionality reduction of short-term electric load data (section 4.3).

Table 1. Parameters of different algorithms

Parameters	Value
Population size	100
Dimension	Number of bands
Hidden layer nodes	10
Number of runs for each technique	30
c1,c2 Acceleration coefficient in the PSO algorithm	c1 = 2.1, c2 = 1.6
A Convergence factor in the GWO algorithm	2
B Parameter in the MFO algorithm	1
The ST Search trend in TSA algorithm	0.1
The Nvar Parameter in LTSA algorithm	1
The β Parameter in LTSA algorithm	1.5

Table 2. Statistics of the data set

Data set	Instances	Attributes	Forecast number
Combined Cycle Power Plant(CCPP)	9568	4	1
Airfoil Self-Noise(ASN)	1503	5	1
Concrete Compressive Strength(CCS)	1030	8	1
Yacht Hydrodynamics(YH)	308	6	1
Wine quality-red(WR)	1599	11	1
Wine quality-white(WW)	4898	11	1

4.1. Experiments of LTSA-ELM on UCI datasets

To verify the effectiveness of LTSA-ELM, the proposed method is compared with some commonly used algorithms: PSO, GWO, MFO, and parameters of these algorithms are shown in Table 1. In this paper, six UCI standard datasets are used to test the model, as shown in Table 2. All experimental data sets are mapped to the [-1, 1] by the maximum and minimum normalization method [17]. Using 5-fold cross-validation technique [11], each data set is divided into five parts, four of which are selected as training sets and one as a testing set. The ELM adopts S-type activation function, the selection of activation functions is explained in section 4.3. The results of the prediction accuracy (ACC) and the mean square error (MSE) on six UCI public datasets are shown in Table 3.

Table 3. Prediction results of five methods on the UCI datasets

Data set	ELM		PSO-ELM		GWO-ELM		MFO-ELM		TSA-ELM		LTSA-ELM	
	ACC /%	MSE	ACC /%	MSE	ACC/%	MSE	ACC /%	MSE	ACC /%	MSE	ACC /%	MSE
ASN	98.77	1.51e-16	99.15	7.23e-17	99.15	7.28e-17	99.15	7.29e-17	99.15	7.22e-17	**99.45**	7.30e-18
CCS	98.99	1.03e-12	99.01	9.72e-13	99.02	9.54e-13	99.02	9.52e-13	99.02	9.66e-13	**99.87**	9.52e-13
YH	91.61	7.03e-7	98.56	2.08e-8	98.51	2.21e-8	98.35	2.73e-8	98.56	2.07e-8	**99.51**	2.02e-8
WR	97.50	7.22e-5	99.22	6.03e-9	98.40	2.56e-8	98.89	1.22e-8	98.92	1.17e-8	**99.77**	5.12e-10
WW	98.50	2.47e-9	99.43	3.28e-11	99.46	2.91e-11	99.41	3.44e-11	99.40	3.61e-11	**99.96**	4.12e-12
CCPP	99.13	7.48e-19	99.17	6.80e-19	99.65	1.20e-17	99.88	1.38e-18	99.85	7.12e-19	**99.99**	1.29e-19

As can be seen from Table 3, LTSA-ELM outperforms other competing algorithms on all test data sets. The accuracy of LTSA-ELM on the six data sets has reached more than 99%. In addition, LTSA-ELM achieves better accuracy especially on YH which is higher than ELM without LTSA about 8.6%. Moreover, the comparison between the results of TSA-ELM and LTSA-ELM demonstrates that the combination of Lévy flight and TSA further improve the probability of getting the optimal parameters. Meanwhile, LTSA-ELM has better stability compared with other algorithms as it gets the minimum results on MSE for most datasets. It is proved that the predefined parameters optimized by LTSA effectively improve the performance of ELM.

4.2. Experiments of KPCA-LTSA-ELM for electric load datasets

The proposed method in this paper is applied to predict short-term electric load based on EUNITE competition data [53].

KPCA is used to perform the principal component analysis of sample data. The electric load data extracted by KPCA is used as the input sample of LTSA-ELM model. Among them, 6880 samples are randomly selected as the training data and 1719 samples are used as the test data. In order to compare the prediction advantages of each model, the mean absolute error (MAE), the average absolute percentage error (MAPE), the predicted root mean square error (MSE), the coefficient of determination (R^2), the test accuracy rate (accuracy) and the test time (T) are used as evaluation indicators. Then, the evaluation values corresponding to the respective prediction models are calculated. The results are shown in Table 4 and Figures 3 and 4.

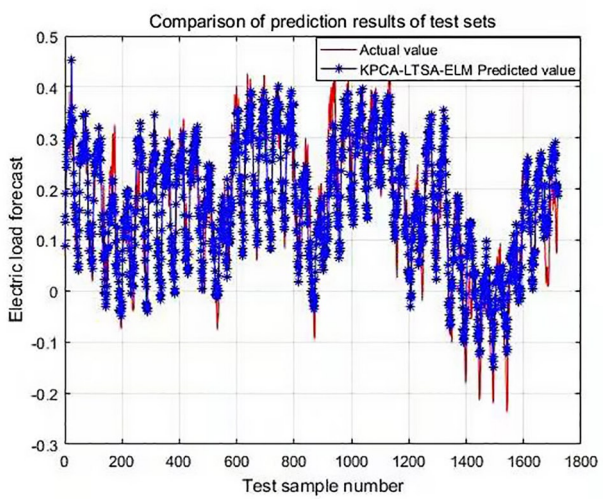

Fig. 3. Electric load prediction of KPCA-LTSA-ELM

As can be seen from Table 4, the predicted values of KPCA-LTSA-ELM has the highest prediction accuracy. The mean square error of prediction is the minimum compared with the other models. Compared with LTSA-ELM, the prediction accuracy of KPCA-LTSA-ELM is improved by 1.84%, the mean square error is reduced by 0.00051,

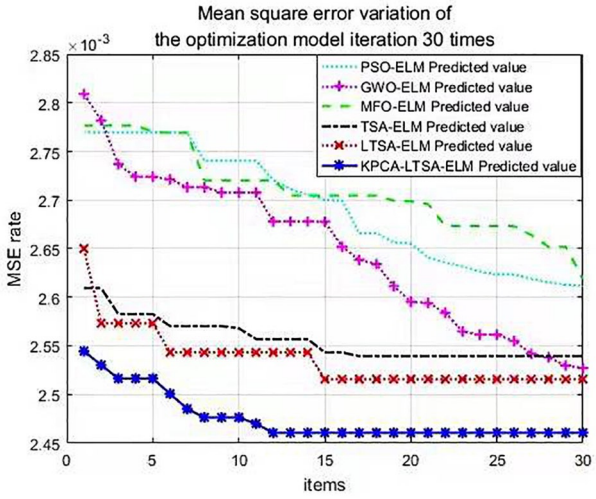

Fig. 4. The root mean square error (MSE) curve of each model versus the number of iteration

and the test determination coefficient is increased by 0.01855. This shows that the sample data processed based on KPCA eliminates the related redundancy between influencing factors. From the stability of the model, KPCA-LTSA-ELM outperforms ELM, PSO-ELM, GWO-ELM, MFO-ELM, TSA-ELM, LTSA-ELM and SVM models. In terms of model convergence, the convergence time of LTSA-ELM and KPCA-LTSA-ELM are shorter and more efficient than most other models.

As shown in Figure 3, the red line represents the true value, and the blue line represents the predicted value of KPCA-LTSA-ELM. The predicted value curve and the true value curve can fit well. This indicates that KPCA-LTSA-ELM has a high prediction accuracy. According to Figure 4, the mean square error predicted by each model decreases as the number of iterations increases. The convergence curve shows that KPCA-LTSA-ELM has smaller training errors and faster convergence speed than other models. The model can fully exploit the internal implicit laws of the prominent samples and effectively interpret the nonlinear relationship between short-term electrical load and other influencing factors. Therefore, this model can be applied to the field of short-term electric load forecasting to provide a theoretical guarantee for users to use normal electricity and reduce generation costs.

4.3. Sensitivity analysis of parameters

4.3.1. Determination of hidden nodes and the type of activation function for ELM

In order to determine the optimal activation function and the number of nodes in the hidden layer, the trial and error method is used in this paper. The number of hidden layer nodes is determined as 13 based on the empirical formula.

Table 4. Analysis of results of each prediction model

Model category	MAE	MAPE	MSE	R^2	Accuracy(%)	T(s)
ELM	0.04448	1.17311	0.00334	0.7902	92.20	0.0468
PSO-ELM	0.04128	0.87613	0.00253	0.8468	95.08	0.0468
GWO-ELM	0.04092	0.89789	0.00272	0.83549	92.90	0.1248
MFO-ELM	0.04113	0.94533	0.00295	0.82159	93.75	0.0312
TSA-ELM	0.03995	0.84463	0.00274	0.83412	93.76	0.0312
LTSA-ELM	0.03892	0.81437	0.00263	0.84077	94.76	0.0312
KPCA-LTSA-ELM	0.03800	0.76717	0.00212	0.85932	96.60	0.0312
SVM	0.05572	1.27575	0.00506	0.69640	92.88	5.7849

ELM has three kinds of activation functions: S-type (Sigmoid type), sine function (Sine type), and the hard-threshold type transfer function (Hardlim type). The generalization ability of each activation function has different prediction effects in different instances. The expressions of these three activation functions correspond to equations (23-25) respectively as below:

$$f(x_1) = \frac{1}{1+e^{-x_1}} \tag{23}$$

$$f(x_2) = \sin(x_2) \tag{24}$$

$$f(x_3) = \begin{cases} 1 & x_3 > 0 \\ -1 & x_3 < 0 \end{cases} \tag{25}$$

The prediction result of the above three different activation functions is calculated to determine the optimal activation function. The results are shown in Table 5:

Table 5. Prediction results of different activation functions

Type of activation function	Number of hidden layer nodes	Mean square error (MSE%)
Sigmoid type	13	0.21398
Sine type	13	0.21848
Hardlim type	13	0.22056

The Sigmoid activation function has the best prediction effect, followed by the Hardlim and Sine activation functions. Therefore, the activation function of ELM can be determined as the Sigmoid function.

Then, the number of hidden layer nodes is calculated with a test interval of [3, 13] to determine the optimal number of nodes in the hidden layer. In the paper, the same data sample set is selected. Under the same conditions in all aspects, the quantity of neurons in the hidden layer is determined by comparing the mean square error. Multiple network trainings are conducted on different hidden layer neurons, and the mean square error is shown in Table 6.

From the comparison of the distributions in Table 6, it can be seen that the mean square error tends to be minimized when the number of neurons in the hidden layer is 13. Therefore, the number of nodes in the hidden layer is set as 13.

Table 6. Mean Square Errors for different number of neurons in the hidden layer

Number of hidden layer neurons	Number of training	Mean square error (MSE%)	Decision coefficient (R^2)
3	50	0.27461	0.83014
4	50	0.26966	0.84129
5	50	0.25291	0.85116
6	50	0.24351	0.85668
7	50	0.23656	0.86077
8	50	0.22557	0.86195
9	50	0.21156	0.86385
10	50	0.20934	0.86724
11	50	0.20861	0.87075
12	50	0.20852	0.87143
13	50	0.20845	0.87679

4.3.2. Feature extraction based on KPCA

The eleven characteristic attributes in the short-term electric load data of Europe are processed by PCA, KPCA based on the polynomial kernel and KPCA based on the Gaussian kernel extract the first six principal components.

The eigenvalue (Eig), the Variance contribution rate (Vcr), and the cumulative contribution rate (Ccr) are calculated. In this paper, the cumulative contribution rate of principal component is greater than 96% as the standard. When using KPCA (Gaussian kernel) to select features, the cumulative contribution rate of the variance of the first six principal components reaches 96.05%, which can replace the original eleven indicators. The experimental results are shown in Table 7:

As can be seen from the above table, the methods using PCA and KPCA reduce the features of training samples inputs, and retains most of information. However, there are some differences in the effectiveness of these three methods. The short-term electric load feature extraction method based on PCA has the advantages of dimension reduction and feature extraction. But its first principal component contribution rate is only 29.23%, which is 15.96% lower than KPCA based on linear kernel function and 14.3% lower than KPCA based on Gaussian kernel function. KPCA expands the research range of data characteristics from linear to nonlinear, therefore it can reduce the dimensionality and obtain better performance than PCA. KPCA based on linear kernel function has the similar principal component contribution rate as KPCA based on Gaussian kernel function. However, KPCA based on Gaussian kernel function extracts fewer features than the linear kernel function extraction, consequently, it is better to choose the Gaussian radial kernel function.

Table 7. Comparison table of three analytical methods

Number	PCA			KPCA (Linear kernel)			KPCA(Gaussian kernel)		
	Eig	Vcr(%)	Ccr(%)	Eig	Vcr(%)	Ccr(%)	Eig	Vcr(%)	Ccr(%)
1	3.2150	29.23	29.23	3702.5	45.19	45.19	0.0034	43.53	43.53
2	2.4834	22.58	51.80	2351.2	28.70	73.89	0.0023	29.98	73.51
3	1.9021	17.29	69.10	628.9	7.68	81.57	6.65e-04	8.62	82.13
4	1.0194	9.27	78.36	473.6	5.78	87.35	4.67e-04	6.06	88.19
5	0.9343	8.49	86.86	373.3	4.56	91.90	3.60e-04	4.67	92.85
6	0.5030	4.57	91.43	259.5	3.17	95.08	2.46e-04	3.19	96.05
7	0.3899	3.54	94.97	122.6	1.50	96.57	1.26e-04	1.63	97.68
8	0.2386	2.17	97.14	76.6	0.94	97.51	7.65e-05	0.99	98.67
9	0.1531	1.39	98.53	58.9	0.72	98.23	4.43e-05	0.57	99.25
10	0.0971	0.88	99.42	38.1	0.47	98.69	3.78e-05	0.49	99.25
11	0.0641	0.58	100	27.2	0.33	99.03	2.04e-05	0.26	99.74

5. Conclusion

In general, a method for short-term electric load forecasting based on ELM and the improved TSA is put forward. The performance of LTSA-ELM is evaluated on six different dimensional datasets. It is demonstrated that LTSA is more suitable for learning the optimal parameters of ELM in terms of convergence speed, prediction accuracy and stability among experimental results. Moreover, KPCA is used to extract the features in the experiment for forecasting short-term electric load in Europe. Experimental results of the test load data show that KPCA-LTSA-ELM successfully forecasts the required load at a certain time of the day. In conclusion, the combination of ELM and LTSA achieves higher accuracy in less time and maintains a better balance between prediction efficiency and accuracy. In recent years, the scale of electricity consumption continually increases and the power system faces greater challenges. This research only focuses on short-term electric load forecasting, whereas middle-term and long-term load forecasting is more valuable to the energy market than short-term electric load forecasting. We will explore these issues in future research.

Acknowledgement

This research was funded by the National Natural Science Foundation of China under Grant No.61502155, 61772180, and the technological innovation project of Hubei Province 2019 No.2019AAA04, also funded by Fujian Provincial Key Laboratory of Data Intensive Computing and Key Laboratory of Intelligent Computing and Information Processing, Fujian No.BD201801.

This work was financed in the framework of the project Lublin University of Technology - Regional Excellence Initiative, funded by the Polish Ministry of Science and Higher Education (contract no. 030/RID/2018/19).

References

1. Čepin M. Evaluation of the importance factors of the power plants within the power system reliability evaluation. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(4): 631-637, https://doi.org/10.17531/ein.2019.4.12.
2. Barshandeh S, Haghzadeh M. A new hybrid chaotic atom search optimization based on tree-seed algorithm and Levy flight for solving optimization problems. Engineering with Computers 2020: 1-44, https://doi.org/10.1007/s00366-020-00994-0.
3. Chen F. Research on Short-term Power Load Forecasting Based on Hybrid Intelligent Optimization Algorithm. Master thesis, Hubei University of Technology, 2020.
4. Chen Y, Kloft M, Yang Y et al. Mixed kernel based extreme learning machine for electric load forecasting. Neurocomputing 2018; 312: 90-106, https://doi.org/10.1016/j.neucom.2018.05.068.
5. Dash R, Dash PK, Bisoi R. A self adaptive differential harmony search based optimized extreme learning machine for financial time series prediction. Swarm and Evolutionary Computation 2014; 19: 25-42, https://doi.org/10.1016/j.swevo.2014.07.003.
6. Dong H, Li M, Zhang S et al. Short-term power load forecasting based on kernel principal component analysis and extreme learning machine. J. Electron. Meas. Instrum 2018; 32: 188-193.
7. Dukalski P, Będkowski B, Parczewski K et al. Dynamics of the vehicle rear suspension system with electric motors mounted in wheels. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(1): 125-136, https://doi.org/10.17531/ein.2019.1.14.
8. El-Fergany AA, Hasanien HM. Tree-Seed Algorithm for Solving Optimal Power Flow Problem in Large-Scale Power Systems Incorporating Validations and Comparisons. Applied Soft Computing 2018; 64: 307-316, https://doi.org/10.1016/j.asoc.2017.12.026.
9. Ertugrul ÖF. Forecasting electricity load by a novel recurrent extreme learning machines approach. International Journal of Electrical Power & Energy Systems 2016; 78: 429-435, https://doi.org/10.1016/j.ijepes.2015.12.006.
10. Faris H, Mirjalili S, Aljarah I. Automatic selection of hidden neurons and weights in neural networks using grey wolf optimizer based on a hybrid encoding scheme. International Journal of Machine Learning and Cybernetics 2019; 10(10): 2901-2920, https://doi.org/10.1007/s13042-018-00913-2.
11. Fushiki T. Estimation of prediction error by using K-fold cross-validation. Statistics and Computing 2011; 21(2): 137-146, https://doi.org/10.1007/s11222-009-9153-8.
12. Glowacz A, Tadeusiewicz R, Legutko S et al. Fault diagnosis of angle grinders and electric impact drills using acoustic signals. Applied Acoustics 2021; 179: 108070, https://doi.org/10.1016/j.apacoust.2021.108070.
13. Glowacz A. Fault diagnosis of electric impact drills using thermal imaging. Measurement 2021; 171: 108815, https://doi.org/10.1016/j.measurement.2020.108815.
14. Glowacz A. Ventilation diagnosis of angle grinder using thermal imaging. Sensors 2021; 21(8): 2853, https://doi.org/10.3390/s21082853.
15. Greengard S. The internet of things. MIT press: 2021, https://doi.org/10.7551/mitpress/13937.001.0001.
16. Haklı H, Uğuz H. A novel particle swarm optimization algorithm with Levy flight. Applied Soft Computing 2014; 23: 333-345, https://doi.org/10.1016/j.asoc.2014.06.034.
17. Han J, Pei J, Kamber M. Data mining: concepts and techniques. Elsevier: 2011.
18. Houimli R, Zmami M, Ben-Salha O. Short-term electric load forecasting in Tunisia using artificial neural networks. Energy Systems 2020; 11(2): 357-375, https://doi.org/10.1007/s12667-019-00324-4.
19. Hu J, Wu M, Chen X et al. A multilevel prediction model of carbon efficiency based on the differential evolution algorithm for the iron ore sintering process. IEEE Transactions on Industrial Electronics 2018; 65(11): 8778-8787, https://doi.org/10.1109/TIE.2018.2811371.
20. Huang G, Zhu Q, Siew CK. Extreme learning machine: Theory and applications. Neurocomputing 2006; 70(1): 489-501, https://doi.org/10.1016/j.neucom.2005.12.126.
21. Kang F, Liu J, Li J, Li S. Concrete dam deformation prediction model for health monitoring based on extreme learning machine. Structural Control and Health Monitoring 2017; 24(10): e1997, https://doi.org/10.1002/stc.1997.
22. Kapłonek W, Nadolny K, Królczyk GM. The use of focus-variation microscopy for the assessment of active surfaces of a new generation of coated abrasive tools. Meas. Sci. Rev 2016; 16(2): 42-53, https://doi.org/10.1515/msr-2016-0007.
23. Kiran MS. TSA: Tree-seed algorithm for continuous optimization. Expert Systems with Applications 2015; 42(19): 6686-6698, https://doi.org/10.1016/j.eswa.2015.04.055.
24. Kochan O, Sapojnyk H, Kochan R. Temperature field control method based on neural network. 2013 IEEE 7th International Conference on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS), IEEE: 2013; 1: 21-24, https://doi.org/10.1109/IDAACS.2013.6662632.
25. Kozieł J, Przystupa K. Using the FTA method to analyze the quality of an uninterruptible power supply unitreparation UPS. Przeglad

Elektrotechniczny 2019; 95(3): 37-40, https://doi.org/10.15199/48.2019.01.20.
26. Kozłowski E, Mazurkiewicz D, Sęp J, Żabiński T. The Use of Principal Component Analysis and Logistic Regression for Cutter State Identification. International Conference Innovation in Engineering, Springer: 2021: 396-405, https://doi.org/10.1007/978-3-030-78170-5_34.
27. Kozłowski E, Mazurkiewicz D, Żabiński T et al. Assessment model of cutting tool condition for real-time supervision system. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2019; 21(4): 679-685, https://doi.org/10.17531/ein.2019.4.18.
28. Krolczyk G, Gajek M, Legutko S. Predicting the tool life in the dry machining of duplex stainless steel. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2013; 15: 62-65.
29. Kyriakides E, Polycarpou M. Short Term Electric Load Forecasting: A Tutorial. In Chen K, Wang L (eds): Trends in Neural Computation, Berlin, Heidelberg, Springer Berlin Heidelberg: 2007: 391-418, https://doi.org/10.1007/978-3-540-36122-0_16.
30. Li H, Zhang S, Zhang C et al. A novel unsupervised Levy flight particle swarm optimization (ULPSO) method for multispectral remote-sensing image classification. International Journal of Remote Sensing 2017; 38(23): 6970-6992, https://doi.org/10.1080/01431161.2017.1368102.
31. Liu C, Ye C. Bat algorithm with Levy flight characteristics. CAAI Trans Intell Syst 2013; 3: 240-246.
32. Liu T, Chen R, Xiao Y, Yang J. Stream-Based Short-Term Demand Forecasting Model Using ARIMA. Applied Mechanics and Materials 2012; 220-223: 315-318, https://doi.org/10.4028/www.scientific.net/AMM.220-223.315.
33. Mao L, Jiang Y. Medium-and long-term load forecasting based on partial least squares regression analysis. Power System Technology 2008; 32(19): 71-77.
34. Mazurek PA, Michałowska J, Koziel J et al. The intensity of electromagnetic fields in the range of GSM 900, GSM 1800 DECT, UMTS, WLAN in built-up areas. Przeglad Elektrotechniczny 2018; 94(12): 202-205, https://doi.org/10.15199/48.2018.12.45.
35. Michałowska J, Józwik J. Prediction of the parameters of magnetic field of CNC machine tools. Przeglad Elektrotechniczny, 95 (1): 134-136, https://doi.org/10.15199/48.2019.01.34.
36. Muneeswaran V, Rajasekaran MP. Performance evaluation of radial basis function networks based on tree seed algorithm. 2016 International Conference on Circuit, Power and Computing Technologies (ICCPCT), 2016: 1-4, https://doi.org/10.1109/ICCPCT.2016.7530267.
37. Niu D, Ji L, Tian J. Wavelet neural network embedded expert system used in short-term load forecasting. 2011 2nd IEEE International Conference on Emergency Management and Management Sciences, IEEE: 2011: 190-193.
38. Pai P, Hong W. Support vector machines with simulated annealing algorithms in electricity load forecasting. Energy Conversion and Management 2005; 46(17): 2669-2688, https://doi.org/10.1016/j.enconman.2005.02.004.
39. Pedregal DJ, Young PC. Development of improved adaptive approaches to electricity demand forecasting. Journal of the Operational Research Society 2008; 59(8): 1066-1076, https://doi.org/10.1057/palgrave.jors.2602447.
40. Przystupa K, Koziel J. Analysis of the quality of uninterruptible power supply using a UPS. 2018 Applications of Electromagnetics in Modern Techniques and Medicine (PTZE), IEEE: 2018: 191-194, https://doi.org/10.1109/PTZE.2018.8503204.
41. Przystupa K. An attempt to use FMEA method for an approximate reliability assessment of machinery. ITM Web of conferences, EDP Sciences: 2017; 15: 05001, https://doi.org/10.1051/itmconf/20171505001.
42. Przystupa K. Selected methods for improving power reliability. Przegląd Elektrotechniczny 2018; 94(12): 270-273, https://doi.org/10.15199/48.2018.12.62.
43. Razfar MR, Farshbaf ZR, Haghshenas M. Optimum surface roughness prediction in face milling by using neural network and harmony search algorithm. The International Journal of Advanced Manufacturing Technology 2011; 52(5): 487-495, https://doi.org/10.1007/s00170-010-2757-5.
44. Sachenko A, Kochan V, Turchenko V, Golovko V, Savitsky Y, Laopoulos T. Method of construction of training set for neural network that predicts drift of data acquisition module. Patent of Ukraine 50380. G06F15/18.
45. Schölkopf B, Smola A, Müller KR. Kernel principal component analysis. In Gerstner W, Germond A, Hasler M, Nicoud J-D (eds): Artificial Neural Networks - ICANN'97, Berlin, Heidelberg, Springer: 1997: 583-588, https://doi.org/10.1007/BFb0020217.
46. Senthilnath J, Das V, Omkar SN, Mani V. Clustering Using Levy Flight Cuckoo Search. Proceedings of Seventh International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA 2012), India, Springer: 2013: 65-75, https://doi.org/10.1007/978-81-322-1041-2_6.
47. Senthilnath J, Kulkarni S, Raghuram Dr. et al. A novel harmony search-based approach for clustering problems. International Journal of Swarm Intelligence 2016; 2(1): 66-86, https://doi.org/10.1504/IJSI.2016.077434.
48. Song W, Beshley M, Przystupa K et al. A software deep packet inspection system for network traffic analysis and anomaly detection. Sensors 2020; 20(6): 1637, https://doi.org/10.3390/s20061637.
49. Su J, Kochan O. Common mode noise rejection in measuring channels. Instruments and Experimental Techniques 2015; 58(1): 86-89, https://doi.org/10.1134/S0020441215010091.
50. Sultoni S, Abdullah AG. Real Time Facial Recognition Using Principal Component Analysis (PCA) And EmguCV. IOP Conference Series: Materials Science and Engineering 2018; 384: 012079, https://doi.org/10.1088/1757-899X/384/1/012079.
51. Sun S, Przystupa K, Wei M et al. Fast bearing fault diagnosis of rolling element using Lévy Moth-Flame optimization algorithm and Naive Bayes. Eksploatacja i Niezawodnosc - Maintenance and Reliability 2020; 22: 730-740, https://doi.org/10.17531/ein.2020.4.17.
52. Sun W, Wang Y, Zhang C. Forecasting CO2 emissions in Hebei, China, through moth-flame optimization based on the random forest and extreme learning machine. Environmental Science and Pollution Research 2018; 25(29): 28985-28997, https://doi.org/10.1007/s11356-018-2738-z.
53. Wang B, Zhao S, Zhang S. A distributed load forecasting algorithm based on cloud computing and extreme learning machine. Power System Technology 2014; 38(2): 526-531.
54. Wang G, Lu M, Dong Y, Zhao X. Self-adaptive extreme learning machine. Neural Computing and Applications 2016; 27(2): 291-303, https://doi.org/10.1007/s00521-015-1874-3.
55. Wang J, Jiang H, Wu Y, Dong Y. Forecasting solar radiation using an optimized hybrid model by Cuckoo Search algorithm. Energy 2015; 81: 627-644, https://doi.org/10.1016/j.energy.2015.01.006.
56. Yamany W, Fawzy M, Tharwat A, Hassanien A E. Moth-flame optimization for training Multi-Layer Perceptrons. 2015 11th International Computer Engineering Conference (ICENCO), 2015: 267-272, https://doi.org/10.1109/ICENCO.2015.7416360.

57. Yang J, Stenzel J. Short-term load forecasting with increment regression tree. Electric Power Systems Research 2006; 76(9-10): 880-888, https://doi.org/10.1016/j.epsr.2005.11.007.
58. Yang W, Zhou Q, Tsui KL. Differential evolution-based feature selection and parameter optimisation for extreme learning machine in tool wear estimation. International Journal of Production Research 2016; 54(15): 4703-4721, https://doi.org/10.1080/00207543.2015.1111534.
59. Yang X, Deb S. Engineering optimisation by cuckoo search. International Journal of Mathematical Modelling and Numerical Optimisation 2010; 1(4): 330-343, https://doi.org/10.1504/IJMMNO.2010.035430.
60. Yeromenko V, Kochan O. The conditional least squares method for thermocouples error modeling. 2013 IEEE 7th International Conference on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS), IEEE: 2013; 1: 157-162, https://doi.org/10.1109/IDAACS.2013.6662661.
61. Yu Z, Xiao L, Li H et al. Model Parameter Identification for Lithium Batteries Using the Coevolutionary Particle Swarm Optimization Method. IEEE Transactions on Industrial Electronics 2017; 64(7): 5690-5700, https://doi.org/10.1109/TIE.2017.2677319.
62. Zhang S, Ren S, Chen R et al. Short-term Power Load Forecast Based on Big Data Reduction and PCA-improved RBF Network. Acta Metrologica Sinica 2018; 39(3): 392-396.
63. Zhang Y, Yang S, Guo Z et al. Wind speed forecasting based on wavelet decomposition and wavelet neural networks optimized by the Cuckoo search algorithm. Atmospheric and Oceanic Science Letters 2019; 12(2): 107-115, https://doi.org/10.1080/16742834.2019.1569455.
64. Zhao Y, Liu J, Lyu Z, Ding Z. Structural damage identification based on residual vectors and Tree-seed algorithm. Acta Scientiarum Naturalium Universitatis Sunyatseni 2017; 56(4): 46-50.
65. Zhou P, Ang B, Poh KL. A trigonometric grey prediction approach to forecasting electricity demand. Energy 2006; 31(14): 2839-2847, https://doi.org/10.1016/j.energy.2005.12.002.