A Generalized Long-term Load Forecasting Model Using Modified Grey Wolf Optimization Algorithm

Rizk M. Rizk-Allah¹, I. M. El-Desoky¹, A. N. Ayad¹

¹Faculty of Engineering, Menoufia University, Shebin El-Kom, Egypt
E-mail: rizk_masoud@yahoo.com

Abstract. Annual accurate predicting is important for long-term planning of the electricity company because it always seeks to achieve the reliability and sustainability for the electrical services. The process of long term load forecasting (LTLF) is implemented in several manners depending on random variables affecting the consumption of the load. This paper introduces two different strategies that have been utilized for annual electrical peak load (PL) for Jordan. The first strategy used in this paper is based on different historical features gathered from the national electric power company (NEPCO) such as (population growth, the gross domestic product of the country and the historical annual PL). The second strategy is based on the historical electrical peak load only. The work includes a comparison process using three different optimization algorithms particle swarm (PSO), grey wolf (GWO) and modified grey wolf (MGWO-based on the concept of aligned) to train a feed forward neural network (FFN). The proposed MGWO-FFN gives the best results of both the two strategies than GWO-FFN and PSO-FFN. Results of the two strategies are presented and indicate the superiority of the second strategy on the first one.

Key words: Grey wolf; Forecasting; Optimization; Feed forward net.

Abbreviations

| Acronym | Description |
|---------|-------------|
| LTLF    | Long term load forecasting |
| NEPCO   | National electrical power company |
| GWO     | Grey wolf optimization algorithm |
| MGWO    | Modified grey wolf optimization algorithm |
| PSO     | Particle swarm optimization algorithm |
| FFN     | Feed forward neural network |
| MGWO-FFN| Hybrid model of modified grey wolf algorithm based training a feed forward neural network |
| PL      | Electrical peak load consumption |
| RBF     | Radial base neural network |
| POP     | Number of population |
| GDP     | Gross domestic product |
| NEPCO   | National electric power company |
| MSE     | Mean square error |
1. Introduction
For more decades researchers attempt to find out various strategies to utilize it in electrical LTLF. In order to confront the challenges those face the electrical company which are represented in the increase in demand, urbanization and the fossil fuel depletion, it became obligatory to plan the form of power system to keep face those development and maintains the coverage of electricity consumption. For that reasons most of researchers spare no effort for finding more accurate methods that service the process of LTLF. They search all over the fields and directions to make the process of constructing new power system easier and gives accurate data about the electricity consumption in the future. Because of the importance of electric power and it has become the main lifeline of all nations, officials and governments deduct a large time to present more accurate studies about planning the future power systems.
Researchers continue in presenting more accurate strategies with different algorithms for the purpose of accuracy for LTLF process. This work presents a comparison between two different strategies for LTLF for Jordan. One of them uses more than one type of features that affect the consumption of electricity and is presented in [1]. Such strategy uses the historical data of population number, the gross domestic product and PL as input to the used prediction model to predict the annual PL of Jordan. This strategy demands more types of data which have correlations with the consumption of electricity to do the process of forecasting, for that reason researchers attempt to find out strategies that have advantages of accuracy and ease of use. So Wang et al [2] proposes a hybridized support vector regression and differential evolution algorithm to forecast the annual load and using a strategy that utilizes the historical electrical load only to forecast the annual electrical load. Such strategy is characterize with the ease of use and gives more accurate results.
Another branch of LTLF process is the model used, and for the recent decades more papers present new and powerful algorithms which heading to giving accurate results and minimizing the global error of forecasting process. Regression models [3, 4] and time series strategy [5, 6], have not the capability to give accurate results in nonlinear unsteady relationships. Neural network is utilized in forecasting process because its capability for presenting very good performance to recognize nonlinear patterns effectively. The invention of neural networks is in 1943 and presented in [7]. The variant types of NNs are: FFN [8] which is a type of MLP, Kohonen self-organizing network [9], RBF network [10], recurrent neural network [11], and spiking neural networks [12]. This paper utilizes the FFN type in which data flow is organized to be in uni-direction over the networks. However, the various types of NNs, they are needed to have the experience during the whole historical data. This experience is obtained through the process called learning process, in which a part of data is applied to the network and giving it the ability to adjust their variables. The classes of learning types are supervised [13, 14] and unsupervised [15, 16] learning. More algorithms have been utilized to do the process of learning NNs. Optimization algorithms have an essential rule to make the learning process more accurate.
This paper presents a comparison between two different strategies to forecast the annual PL for Jordan using three models. Such models are FFN trained based on three different optimization algorithms like GWO algorithm which is proposed in [17], PSO algorithm that is presented in [18] and MGWO. Results indicate that MGWO-FFN outperforms the other two models GWO-FFN and PSO-FFN during the two strategies used by calculating MSE that have minimum value. Historical data including POP, GDP and the annual PL of Jordan are obtained from NEPCO to be utilized in this work.
The other parts in this paper are organized as: Section 2 presents the proposed strategies used to forecast the Jordan's PL. Section 3 presents three optimization algorithms MGWO, GWO and PSO using to train FFN. Section 4 introduces the experimental works and results. Section 5 concludes this work.

2. Two different strategies for LTLF

2.1 Forecasting PL using different types of features

All of us know that the predicting process needs some features which affect the output, so that these features definitely have correlation between it and the predicted value. This strategy is used to forecast PL using more than one type of feature that affects the process. So GDP, GDP² and POP are the various features that more effects on the value of consumed PL. Table 1 shows the correlation factors between these features and PL and how they affect the manner of electricity consumption which are calculated using the MATLAB corrcoef function. The used data fall through the year 2000 to 2018. Forecasting using this strategy utilizes the mentioned features as the input to the proposed models and these features are sampled to have set of training samples and testing ones. Fig.1 clarifies the prediction process using Year, GDP, GDP² and POP as inputs to the model.

Table 1: correlation factor between features and PL

| Input variables | Correlation factors |
|-----------------|---------------------|
| PL-POP          | 0.933               |
| PL-GDP          | 0.997               |
| PL-GDP²         | 0.993               |
| PL-Year         | 0.996               |

Fig.1 forecasting process using more than one type of features
The disadvantages of this strategy is effort that lost to find the affected data and then the gathering process to all these types of data accurately is some difficult in order to give us accurate predicted load, as well as, this process needs more time to be done. But its advantage is the accuracy and it gives minimum MSE.

2.2 Forecasting PL using historical electrical load only

In this strategy historical PL data are used and forms the input features to the forecasting model. In it the input are the value of annual PL of the previous three years ($L_{n-3}, L_{n-2}, L_{n-1}$) and the output of this process is the PL of the recent year $L_n$. This process is done through the given years from 2000 to 2018 for Jordan and is implemented as: to forecast the PL in year 2003 we must input the values of PL of the previous four years (2000, 2001 and 2002), then to forecast the PL in year 2004 we must input the PL of years (2001, 2002 and 2003) and the value of PL in 2003 that is used in the second process is the real one. This process is called roll based forecasting strategy which apply a window through all the training and testing samples of the model. Fig.2 presents how is the prediction process using the roll based strategy is done.

![Fig.2 forecasting process using one type of features (PL)](image)

This strategy makes it easy to deal with LTLF because of using the uni-type of the affecting features and it also minimizes the incoming error of multiple types of input features. Another advantage of that strategy is the accuracy of results of both training and testing compared to the first strategy. So a comparison is done between the uses of those two strategies using the three models that are presented in the coming sections. The proposed models are hybrid between different optimization algorithms and FFN.
3. Various hybrid models based on FFN with different optimization algorithms

In this part the author reviews three different types of hybrid models, which train the FFN based on an optimization algorithm, such algorithms are PSO, GWO and MGWO. Each algorithm is equipped to adjust the nature of data and prepare FFN to fit the given pattern of the data. FFN requires some adjustable variables to give us the correct results according to the given data. These variables are the weights and the activation functions which are utilized in FFN. Fig.3 show the architecture of FFN used in this work. The figure shows that the number of the input features used in this work is four and the number of the hidden layer neurons is three and the number of the output layer neurons is one as follows. In FFN there are some steps for performing the candidate task, these steps are summarized in the following equations which describe the FFN's performance.

\[
h_{inj} = \sum_{i=1}^{n} (w_{inij}X_i) , \quad j = 1, 2, \ldots, h (1)
\]

Where \( n \) is the input feature size, \( h \) is the number of hidden layer neurons, \( w_{inij} \) presents the linking weights between the \( ith \) node in the input layer and the \( jth \) node in the hidden layer, \( X \) is the input feature and \( h_{in} \) is the input value for each hidden neurons. Then the hidden output is defined as:

\[
h_{outj} = \tanh(h_{inj}) , \quad j = 1, 2, \ldots, h (2)
\]

Equation (2) shows the use of the first activation function which is utilized to calculate the output of each hidden neuron. After that the output layer requires two

![Fig.3 FFN architecture with one hidden layer](image-url)
processes in order to gives us the result, one of them is the input process and is estimated as the following function:

\[ out_k = \sum_{j=1}^{h} (w_{jk} \cdot h_{jk}) \quad k = 1, 2, \ldots, m \]  

(3)

The other process is calculating the value of the output using another activation function as presented in the following function:

\[ y_k = \tan^{-1}(out_k) \]  

(4)

\( m \) is the number of output layer’s neurons and \( w_{jk} \) is the linking weight between the hidden layer neurons and the output ones. Choosing the appropriate weights is done using the proposed optimization algorithms by minimizing MSE of the trained samples. A detailed presentation is shown to know how each of the proposed algorithms is run to find the minimum solution for FFN.

### 3.1 Particle swarm optimization

It is a based action of any swarm sets like bees, wasps, ants and others; each particle action is based on its own intelligence. Fig.4 shows how each particle is characterized in the swarm for searching the global minimum of error. Each one is defined with its position and velocity so that we initiate their positions and velocities and then work to update them according finding the prior Gbest using minimum error as presented in the following equations:

\[ v_j(i) = v_j(i-1) + c_1 r_1 [P_{best,j} - x_j(i-1)] + c_2 r_2 [G_{best} - x_j(i-1)] \quad j = 1, 2, \ldots, N \]  

(5)

\[ x_j(i) = x_j(i-1) + v_j(i) \]  

(7)

Where \( c_1 \) and \( c_2 \) are the cognitive and social learning rates, respectively, and \( r_1, r_2 \) are uniformly distributed random numbers in the range 0 and 1. The parameters \( c_1 \) and \( c_2 \) denote the relative importance of the position of the particle itself to the position of the swarm. The values of \( c_1 \) and \( c_2 \) are usually assumed to be 2 so that \( c_1 r_1 \) and \( c_2 r_2 \) ensure that the particles would overfly the target about half the time.

### 3.2 Grey wolf optimization

Grey wolf is one of the predatory factions. Grey wolves are beholden as head of hunters, i.e. they hold the head of the food chain because of their strictly rules in social dominant hierarchy. The formation of the grey wolf pack is:

a. The alpha is the leader of the pack, so that it is responsible for making decisions. The alphas decisions are dictated to the pack.

b. The betais inferior wolf and its role is to help the alpha in decision making or other activities. The beta can be either male or female, and he/she is candidate to be the coming alpha.

c. Omega wolves always have to submit to alpha and beta. They are the last wolves that are allowed to eat.

The mathematical modeling of is presented in the following steps:
a. Encircling the prey and updating the position of each wolf according to the
position of the prey which presented in equations 8 and 9
\[
D = |\vec{C} \vec{X}_p (i) - \vec{X} (i) | \tag{8}
\]
\[
\vec{X} (i + 1) = \vec{X}_p (i) - \vec{A} \vec{D} \tag{9}
\]
Where \( i \) shows the current iteration, \( \vec{A} \) and \( \vec{C} \) are coefficient vectors, \( \vec{X}_p \) is the position vector of the prey, and \( \vec{X} \) indicates the position vector of the wolf. Vectors \( \vec{A} \) and \( \vec{C} \) are calculated as:
\[
\vec{A} = 2a \vec{r}_1 - a \tag{10}
\]
\[
\vec{C} = 2 \vec{r}_2 \tag{11}
\]
Parameter \( a \) denotes to a relation which presents a linearly decreasing from \([2 \text{ to } 0]\) over all iterations as shown in (12).
\[
a = 2 - 2 \frac{i}{I_{\text{max}}} \tag{12}
\]
Where \( I_{\text{max}} \) = no. of iterations and \( \vec{r}_1, \vec{r}_2 \) have random values between \([0 \text{ and } 1]\). As can be seen, the wolf's position \( \vec{X} \) can be updated till it hits the goal or the prey by reconfiguration of \( \vec{A} \) and \( \vec{C} \) vectors.

b. Hunting. Hunting of prey is usually guided by \( \alpha, \beta \) and \( \delta \) will participate occasionally.
\[
\vec{D}_{\alpha} = |\vec{C}_{1} \vec{X}_{\alpha} - \vec{X} |, \vec{D}_{\beta} = |\vec{C}_{2} \vec{X}_{\beta} - \vec{X} |, \vec{D}_{\delta} = |\vec{C}_{3} \vec{X}_{\delta} - \vec{X} | \tag{13}
\]
\[
\vec{X}_1 = \vec{X}_{\alpha} - \vec{A}_{1} \vec{D}_{\alpha}, \vec{X}_2 = \vec{X}_{\beta} - \vec{A}_{2} \vec{D}_{\beta}, \vec{X}_3 = \vec{X}_{\delta} - \vec{A}_{3} \vec{D}_{\delta} \tag{14}
\]
Where \( \vec{X}_1, \vec{X}_2 \) and \( \vec{X}_3 \) are the update positions of \( \alpha, \beta \) and \( \delta \) wolves, respectively, according to the best information \( \vec{X}_{\alpha}, \vec{X}_{\beta} \) and \( \vec{X}_{\delta} \) about the prey position. Then the final decision is taken by the following equation:
\[
\vec{X} (i + 1) = \frac{1}{3}(\vec{X}_1 + \vec{X}_2 + \vec{X}_3) \tag{15}
\]

3.3 The proposed modified grey wolf optimization
In order to improve the performance of GWO algorithm, the most important features of the wolves should be considered. Exploration and exploitation are the two features that distinguish the predatory nature of GWs. During all iterations, firstly ordinary GWO starts with gathering data by moving all the pack to explore the search area about the victim's position, and storing these data in best three wolves (\( \alpha, \beta \) and \( \delta \)).
The exploration and exploitation phases are modeled by the parameter $\lambda$, in which, if $\|\lambda\| \geq 1$ the pack is in the exploration phase and for $\|\lambda\| < 1$ exploitation starts. So the re-adjustment of the parameter $\alpha$ balances the two phases in order to fit the problem and giving best solution.

This work proposes a modified variant of GWO algorithm (MGWO) that makes a balance between the exploration and exploitation phases for more accurate data about the prey position (the global minimum) by introducing a relation for the acceleration parameter $\alpha$ that is calculated as:

$$\alpha = \exp(-0.05*i) * [2 - 2*i / I_{\max}]$$ (16)

The difference between the parameter $\alpha$ in the original GWO and the proposed MGWO is presented in fig. 4.

Another modification in GWO algorithm related to how to make the decision. In basic GWO the decision about the final position of the prey is taken according to the best positions of $\alpha$, $\beta$ and $\delta$ wolves and the mean of these positions is the victim's position. In the proposed MGWO the property of aligned is tracked, in which the decision making is according to order of dependence in which $\alpha$ in the first then $\beta$ and $\delta$. Equation (17) presents the new relation of the final decision in MGWO:

$$\bar{X}(i+1) = \frac{1}{6}(3*\bar{X}_1 + 2*\bar{X}_2 + \bar{X}_3)$$ (17)

This work can be summarized in the flowchart which presented in fig. 5.
Fig. 5 the flowchart of the process using the proposed models

1. Start
2. Input data base (historical data)
3. Forming input vector from the normalized training samples, initializing the two groups of weights (input-hidden and hidden-output) and then the combined weight vector is created.
4. Initiate each algorithm's parameters and defining the number of hidden layer neurons and the activation functions.
5. Initiate random positions of each algorithm's soldier.
6. Calculate fitness (MSE); update the positions of soldiers of each algorithm.
7. Apply variables and equations to find the best solution according to each algorithm.

- Are iterations exceeding?
  - Yes
  - Are runs exceeding?
    - Yes
    - Test the yielding MLP and get the predicted values.
    - End
  - NO
  - Run = run + 1
  - i = i + 1
4. Experimental results

In order to implement this comparison process a historical data is obtained from NEPCO for Jordan throughout years from 2000 to 2018 which are presented in Table 2. Such data are year, GDP in millions of JOD, POP in 1000 inhabitants and PL in MW. The comparison is done by introducing each of the two strategies with the needed data. Data must be preprocessed before using i.e. it must be normalized before it has applied to the model. The normalization process is done according to the equation (18) as follows:

\[
z(i) = \frac{x(i)}{x_{\text{max}}}, \quad i = 1, 2, \ldots, S
\]  

(18)

Where the number of samples is \(S\), \(x(i)\) is the value of the feature in the sample \(i\) and \(x_{\text{max}}\) is the maximum value of in the row of each feature. Then data are classified to training samples and testing samples, such that the classification is implemented to train data from year 2000 to 2015 and test the model to estimate PL of 2016, 2017 and 2018. The implementation of the work is using MATLAB program 2013a with an Intel® core™ i5-7200u, 2.5GHZ CPU, 4GB RAM with windows 10 professional LAP. Results of each of the two strategies used are presented as in the coming parts. Table 3 presents the parameters of each algorithm used in both strategies with some difference between each other.

Table 2: data used from NEPCO for Jordan's annual PL forecasting

| Year | GDP  | POP  | PL   |
|------|------|------|------|
| 2000 | 4764 | 4857 | 1206 |
| 2001 | 5044 | 4978 | 1225 |
| 2002 | 5437 | 5098 | 1370 |
| 2003 | 5658 | 5230 | 1387 |
| 2004 | 6152 | 5350 | 1514 |
| 2005 | 6695 | 5473 | 1710 |
| 2006 | 7199 | 5600 | 1860 |
| 2007 | 7757 | 5723 | 2130 |
| 2008 | 8378 | 5850 | 2230 |
| 2009 | 8665 | 5980 | 2300 |
| 2010 | 9030 | 6113 | 2560 |
| 2011 | 9391 | 6241 | 2680 |
| 2012 | 9891 | 6366 | 2790 |
| 2013 | 10484| 6494 | 2975 |
| 2014 | 11184| 6617 | 3000 |
| 2015 | 11436| 9159 | 3300 |
| 2016 | 11683| 9456 | 3250 |
| 2017 | 11885| 9702 | 3320 |
| 2018 | 12185| 9980 | 3380 |

4.1 Forecasting Jordan’s annual PL using more than one type of features

In this strategy users utilize some features to implement the annual PL forecasting for Jordan in the years from 2000 to 2018, that features are (Year, GDP, GDP² and POP), because of the correlation between these candidate features and PL. However this strategy gives more accurate results about PL prediction of Jordan’s power system, it requires more than one type of data which needs more efforts and the probability of existing error in gathering these data accurately.

Table 3: general parameters of each proposed model
Model Parameter
---
MGWO-FFN \( a \) decreases from 2 to 0 non-linear
GWO-FFN \( a \) decreases from 2 to 0 linearly
PSO-FFN \( w_{\text{max}}=0.9 \) (for both strategies), \( w_{\text{max}}=0.4 \) (for 1\(^{st}\) strategy) and 0.8 (for 2\(^{nd}\) one), \( c_1=c_2=1.5 \) (for 1\(^{st}\) strategy) and 1 (for 2\(^{nd}\) one)

For all algorithms
- Maximum iterations=70 (for both strategies)
- Search agents number=700 (for both strategies)
- Upper boundary=2 (for 1\(^{st}\) strategy) and 1.2 (for 2\(^{nd}\) one)
- Lower boundary=-3 (for both strategies)
- Number of runs=50 (for both strategies)

However, this work is done using a little number of samples; it gives accurate results in which the convergence achievement is met at minimum MSE of training and testing because of the powerful of the used models. Equation (19) shows the calculation of MSE:

\[
MSE = \frac{1}{S} \sum_{h=1}^{S} \left( \text{Actual}(h) - \text{Predicted}(h) \right)^2, \quad h = 1,2,\ldots,S
\]

where \( S \) is the number output samples.

FFN is organized to be suitable for this case in which the number of input neurons is four and the output is one with three neurons in hidden layer.

Fig. 6 presents the fitting curve of PL for the three proposed models MGWO-FFN, GWO-FFN and PSO-FFN from 2000 to 2015 in which MGWO-FFN achieves minimum MSE and gives more accurate results comparing to the other two models for both two strategies. Table 4 presents the PL results of the testing process for each model for the three years from 2016 to 2018 and the relative error which is calculated according to the equation:

\[
\text{Re\%}_k = \frac{(\text{Prediction}_k - \text{Actual}_k)}{\text{Actual}_k} \times 100
\]

where \( k \) is the order of the tested years. Fig. 7 shows the estimated relative error for each model for each of the testing years.

![Fig.6 fitting curve using the three models and the proposed first strategy](image)

**Table 4**: Results of the testing process for the annual PL consumption of Jordan using first strategy

| Year | Actual PL | MGWO-FFN | GWO-FFN | PSO-FFN |
|------|-----------|----------|---------|---------|
|      | PL \( \times 10^3 \text{ kwh} \) | Re\% | PL \( \times 10^3 \text{ kwh} \) | Re\% | PL \( \times 10^3 \text{ kwh} \) | Re\% |
| 2016 | 3250      | 3301     | 1.57    | 3313    | 1.94    | 3322    | 2.215 |
| 2017 | 3320      | 3336     | 0.48    | 3348.7  | 0.86    | 3340    | 0.6   |
| 2018 | 3380      | 3378     | -0.059  | 3392    | 0.36    | 3357.9  | -0.654 |
It is clear that FFN is well trained using the proposed MGWO so that the yielding error of testing process is fewer than using GWO and PSO.

4.2 Forecasting Jordan's annual PL using one type of features

This strategy permits users to implement annual PL predicting using the historical data of PL of the previous years that's why annual PL predicting has become easy. The higher the number of samples is the accurate predicting results are obtained, but this work is done using data through years 2000 to 2018 that may be not many to give more accurate results. Like the first one, samples are organized as: training samples from 2000 to 2015 and testing ones from 2016 to 2018. This strategy gives improved results compared to the using of the first strategy, but the more advantage is the ease of use and implementing the PL forecasting. According to equation (19) MSE is calculated and minimized using the proposed models MGWO-FFN, GWO-FFN and PSO-FFN and the yielding results indicates superiority of MGWO to train FFN which gives minimum error comparing to the other two algorithms. FFN is adjusted to suit for this case in which the number of input neurons is three and the output is one with two neurons in hidden layer.

Fig.8 shows the fitting curve of the training data with all proposed algorithms, in which the proposed MGWO-FFN acts very well and achieves the most minimum MSE comparing to GWO-FFN and PSO-FFN. Testing process utilizing this strategy is presented in Table 5 with the relative error in which the proposed MGWO training FFN very well comparing to GWO and PSO. Fig.9 presents the relative error analysis of each utilized model in which MGWO-FFN acts with good performance and achieves minimum relative error comparing to GWO and PSO.
Fig. 8 fitting curve using the three models and the proposed second strategy

Table 5: results of the testing process for the annual PL consumption of Jordan using second strategy

| Year | Actual PL | MGWO-FFN PL | Re% | GWO-FFN PL | Re% | PSO-FFN PL | Re% |
|------|-----------|--------------|-----|------------|-----|------------|-----|
| 2016 | 3250      | 3221.4       | -0.88 | 3108.4     | -4.36 | 3089       | -4.95 |
| 2017 | 3320      | 3298.7       | -0.642 | 3223.8     | -2.898 | 3251.6     | -2.06 |
| 2018 | 3380      | 3392.4       | 0.367 | 3349.1     | -0.91 | 3346.4     | -0.994 |

5. Conclusions

This paper implements a comparison between two strategies for LTLF using three different optimization algorithms to forecast annual PL consumption of Jordan. The first strategy depends on random variables as year, GDP, GDP² and POP, where it has a disadvantage of the possible error that may come from each of these features which leads to yielding error to the forecasting process. The second strategy uses the historical annual PL values to implement the forecasting process, where it has the advantage of ease of use and gives accurate results, but by contrast it needs more data.
to give more accurate results comparing to the first strategy. The implementation in this work is done with the algorithms namely MGWO-FFN, GWO-FFN and PSO-FFN. The proposed MGWO-FFN outperforms the GWO-FFN and PSO-FFN with less time of implementation and minimum MSE.

References

[1] Feilat EA, Talal Al-Sha’abi D, Momani MA. 2017 Long-term load forecasting using neural network approach for Jordan’s power system. Eng Press. 1(2): 43-50. doi:10.28964/EngPress-1-108
[2] Jianjun Wang, Li Li, DongxiaoNiu, Zhongfu Tan, 2012 An annual load forecasting model based on support vector regression with differential evolution algorithm, Applied Energy 94, 65–70.
[3] Antti Sorjamaa, JinHao, NimaReyhani, Yongnan Ji, AmauryLendasse, 2007 Methodology for long-term prediction of time series, Neurocomputer 70 (16–18) 861–869.
[4] H.M. Ai-Hamadi, S.A. Soliman, 2005 Long-term/mid-term electric load forecasting based on short-term correlation and annual growth, Electric Power Systems Research 74 (3), 353–361.
[5] Ruijun Dong, Witold Pedrycz, 2008 A granular time series approach to long-term forecasting and trend forecasting, Physica A: Statistical Mechanics and its Applications 387 (13), 3253–3270.
[6] S. Sp. Pappas, L. Ekonomou, P. Karampelas, D.C. Karamousantas, S.K. Katsikas, G.E. Chatzarakis, P.D. Skafas, 2010 Electricity demand load forecasting of the Hellenic power system using an ARMA model, Electric Power Systems Research 80 (3), 256–264.
[7] McCulloch WS, Pitts W (1943) A logical calculus of the ideas immanent in nervous activity. Bull Math Biophysics 5:115–133
[8] Bebis G, Georgiopoulos M (1994) Feed-forward neural networks. Potentials, IEEE 13:27–31
[9] Kohonen T (1990) The self-organizing map. Proc IEEE 78:1464–1480
[10] Park J, Sandberg IW (1993) Approximation and radial-basisfunction networks. Neural Comput 5:305–316
[11] Dorffner G (1996) Neural networks for time series processing, in Neural Network World
[12] Ghosh-Dastidar S, Adeli H (2009) Spiking neural networks. Int J Neural Syst 19:295–308
[13] Reed RD, Marks RJ (1998) Neurosmithing: supervised learning in feedforward artificial neural networks. Mit Press
[14] Caruana R, Niculescu-Mizil A (2006) An empirical comparison of supervised learning algorithms. In: Proceedings of the 23rd international conference on Machine learning, pp 161–168
[15] Hinton GE, Sejnowski TJ (1999) Unsupervised learning: foundations of neural computation. MIT press
[16] Wang D (2001) Unsupervised learning: foundations of neural computation. AI Mag 22:101
[17] S. Mirjalili, S. M. Mirjalili, A. Lewis, Grey Wolf Optimizer, Advances in Engineering Software, vol. 69, pp. 46-61, 2014, DOI: http://dx.doi.org/1
[18] J. Kennedy and R. Eberhart, “Particle swarm optimization,” in Neural Networks, 1995. Proceedings, IEEE International Conference on, 1995, pp. 1942-1948.