Research Article

Forecasting Electrical Energy Consumption of Equipment Maintenance Using Neural Network and Particle Swarm Optimization

Xunlin Jiang, Haifeng Ling, Jun Yan, Bo Li, and Zhao Li

College of Field Engineering, PLA University of Science and Technology, Nanjing 210007, China

Correspondence should be addressed to Haifeng Ling; hf.ling@ymail.com

Received 1 June 2013; Revised 17 August 2013; Accepted 9 September 2013

1. Introduction

Equipment maintenance plays a key role in resuming and keeping regular operational capabilities of military equipment [1]. Today’s military equipment becomes more and more complicated and technology intensive, which has brought a heavy burden to equipment maintenance. As a result, the demand of electrical energy is increasing rapidly, and the trend is likely to continue in the future. On the one hand, the proportion of energy cost in the total maintenance cost increases rapidly, which was about 26% during 2006–2008 and 30% in 2009 and 2010, according to our historical data. Thus, an accurate forecasting of electrical energy consumption is very important for decision making in maintenance planning and implementation. On the other hand, the total electrical energy demand in China also has increased dramatically in recent years, which has caused many areas and industries to suffer (typically seasonal) electricity shortage [2–4]. In such cases, if we fail to anticipate the electrical energy consumption of maintenance adequately, prearranged maintenance plans will be interrupted or even abandoned, and a lot of important equipment cannot perform expected operations, which may cause significant economic losses or even catastrophic results. For example, in 2008 there are about 28% engineering machineries waiting for repair had their maintenance programs canceled due to unanticipated electricity shortage. In consequence, in some important disaster rescue operations in China in 2009 (such as the Ms6.0 Chuxiong Earthquake and the Ms6.4 Haixi earthquake), the rescue forces had to request a number of dozers, diggers, loaders, and other machineries from other areas, which reduced rescue efficiency.

There are a number of methods and models for forecasting electric energy consumption [5]. Conventional forecasting approaches, such as principal component analysis, least error square (LES), and regression techniques [5–8], become difficult and impractical to provide an adequate forecasting of electrical energy consumption, mainly due to the ever increasing complexity and diversity of equipment maintenance demands. Moreover, those approaches have some inherent limitation and disadvantages such as requirement for large training datasets, sensitivity to noise, and low capability of handling missing data. In recent years, methods and models based on artificial intelligence (AI), including artificial neural networks (ANNs) [9], fuzzy logic and expert systems [10], have received much interest and gained popularity in forecasting electric energy consumption.
Inspired by biological nervous systems, an ANN consists of an interconnection of a number of neurons. It uses processing elements connected by links of variable weights to form a black box representation of systems [11]. One of the most widely used ANNs is the multilayer feed-forward neural network, which consists of an input layer, one or more hidden layers, and an output layer. Figure 1 presents an example of such a three-layer neural network.

![Figure 1: An example of multilayer feed-forward neural network.](image)

One of the most useful and popular applications of ANNs is forecasting, the process of which can be divided into two steps: a learning step for constructing the network model based on training data and a forecasting step for predicting targets values for real data. The learning step is performed by iteratively processing a dataset of training tuples, comparing the network's output value for each tuple with the actual known target value, according to which adjusting the weights of the network so as to minimize the mean-squared error between the network's output value and the actual target value.

The most well-known neural network learning algorithm is backpropagation (BP) [12], which calculates the gradient of the error with respect to the weights for a given input by propagating error backwards through the network. However, BP involves long training times and its performance degrades rapidly when the problem complexity increases [13]. In the recent two decades, many intelligent heuristic algorithms, such as genetic algorithms (GAs) and particle swarm optimization (PSO) algorithms, have been applied to the problem of training different kinds of ANNs [14, 15], including the multilayer feed-forward neural network, and have demonstrated their efficiency and effectiveness in training ANNs for improving their forecasting accuracies [16, 17].

In this paper, we propose an approach of forecasting electrical energy consumption of equipment maintenance using ANN and PSO. We use a feed-forward neural network for modeling possible relationships between the input variables and the expected electrical energy consumption and develop a new adaptive PSO for training the network. Experimental results show that our approach is superior to some state-of-the-art methods.

The rest of the review paper is synthesized as follows: Section 2 reviews the related work of the application of ANNs in electrical energy consumption forecasting and the related problems, Section 3 presents our approach in detail, Section 4 presents the computational experiments and the real-world applications, and Section 5 concludes.

### 2. Related Work

ANNs have gained much popularity in forecasting electric loads and energy demands since 1990s. Kawashima [18] developed an ANN model for building energy prediction, which is trained by the BP algorithm. To improve performance, the algorithm is combined with a three-phase annealing method which gradually reduces the learning rate in order to improve accuracy in a relatively short time. In [19] Islam et al. proposed an ANN-based weather-load and weather-energy models, where a set of weather and other variables are identified for both models together with their correlations and contribution to the forecasted variables. The models were applied to the historical energy, load, and weather data available for the Muscat power system from 1986 to 1990, and the forecast results for 1991-1992 showed that monthly electric energy and load can be predicted within a maximum error of 6% and 10%. In [20] Al-Shehri employed an ANN for forecasting the residential electrical energy in the Eastern Province of Saudi Arabia, and the forecasting result was shown to be closer to the real data than that predicted by the polynomial fit model.

Hatziargyriou and Dialynas [21] proposed an adaptive ANN for mid-term energy forecasting. Their model transforms the input variables to differences or relative differences, in order to predict energy values not included in the training set. The ANN parameters, such as the number of neurons, initial values, and time periods of momentum term and training rate, are simultaneously selected by an optimization process. Sözen et al. [22] also forecasted net energy consumption in Turkey based on an ANN approach, which employs a logistic sigmoid transfer function. They showed that the statistical coefficients of multiple determinations in the experiments are equal to 0.9999 and 1 for training and test data, respectively. In [23] Ermis et al. presented an ANN model which is trained based on world energy consumption data from 1965 to 2004 and applied for forecasting world green energy consumption to the year 2050. It is estimated that world green energy and natural gas consumption will continue increasing after 2050, while world oil and coal consumption are expected to remain relatively stable after 2025 and 2045, respectively.

The work of Azadeh et al. [24] presented an integrated algorithm for forecasting monthly electrical energy consumption based on ANN, computer simulation, and design of experiments using stochastic procedures. The integrated algorithm was applied for forecasting monthly electrical energy consumption in Iran from 1994 to 2005, the result of which demonstrated the effectiveness of the algorithm. Yokoyama et al. [25] proposed a global optimization method called “Modal Trimming Method” for determining ANN parameters, where the trend and periodic change are first removed from time series data on energy demand, and the converted data is used as the main input to a neural network.
nine variables, maintenance, and minor maintenance. Thus we identify themaintenance levels includemajormaintenance, medium heavy equipment, medium equipment, and light equipment. The equipment is classified as two categories. The first category is related to the equipment to be repaired. The equipment is classified as two categories. The first category is related to the equipment to be repaired. The equipment is classified as two categories. The first category is related to the equipment to be repaired. The equipment is classified as two categories. The first category is related to the equipment to be repaired.

Preliminary normalization by the maximum and minimum values of each variable \( x \) is executed based on the following equation:

\[
x = c + \frac{x - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}} (b - c),
\]

where \( x_{\text{max}} \) and \( x_{\text{min}} \) are the upper and lower values of the original variable, respectively, and \( b \) and \( c \) are the corresponding values of the normalized variable.

Given a unit \( j \) in a hidden or output layer, the net input to unit \( j \), denoted by \( I_j \), is defined as

\[
I_j = \sum_i w_{ij}O_i + \sigma_j,
\]

where \( w_{ij} \) is the weight of the connection from unit \( i \) in the previous layer to unit \( j \), \( O_i \) is output of unit \( i \), and \( \sigma_j \) is acting as a threshold bias of the unit \( j \).

And we employ the widely used sigmoid transfer function to compute the output of unit \( j \) from its net input:

\[
O_j = \frac{1}{1 + \exp(-I_j)}.
\]

### 3. The Proposed Method

#### 3.1. Network Structure

Our method adopts a three-layer forward neural network, since an ANN can solve a problem by using one hidden layer, provided it has the proper number of neurons [27]. From the original database of equipment maintenance, the input variables are preliminary selected by a team of experts in the domains of equipment maintenance and electrical power, and a dataset of twenty-three fields is filtered out. In general, the input variables can be divided into two categories. The first category is related to the equipment to be repaired. The equipment is classified as heavy equipment, medium equipment, and light equipment. The maintenance levels include major maintenance, medium maintenance, and minor maintenance. Thus we identify nine variables \( N_{1,1} \sim N_{3,3} \), where the first index denotes the equipment class (1-heavy, 2-medium, 3-light) and the second denotes the maintenance level (1-major, 2-medium, 3-minor). The first category also contains two variables related to rated power and three variables related to maintenance time of all equipment. The second category consists of nine variables related to maintenance resources and environments. Table 1 lists the input variables of our ANN for electrical energy consumption forecasting.

| Category | Variables |
|----------|-----------|
| First    | \( N_{W} \) (the number of workshops), \( N_{E} \) (the number of maintenance lines), \( H_{M1} \) (the available machine-hours of principle maintenance machines), \( H_{M2} \) (the available machine-hours of auxiliary maintenance machines), \( H_{W1} \) (the available working hours of senior maintenance workers), \( H_{W2} \) (the available working hours of common maintenance workers), \( T_{a} \) (the lower limit of working temperature), \( T_{b} \) (the upper limit of working temperature), \( T_{e} \) (the average environment temperature) |
| Second   | \( N_{1} \sim N_{3,3}, P_{A} \) (the total rated power of electric apparatus of all equipment), \( P_{M} \) (the total rated motor power of all equipment), \( T_{f} \) (the expected total fault detection time of all equipment), \( T_{c} \) (the expected total fault correction time of all equipment), \( T_{r} \) (the expected total spare parts replacement time of all equipment) |

Their approach was validated by the application to predict the cooling demand in a building used for a benchmark test. Geem and Roper [26] proposed an ANN model with feed-forward multilayer perceptron, error back-propagation algorithm, momentum process, and scaled data and applied the model to estimate the energy demand for the Republic of Korea. Nevertheless, to our best knowledge, research on the application of ANN models together with intelligent training algorithms to the field of energy consumption forecasting is still rare.

### 3.2. PSO and Control Parameters

PSO is a stochastic, population-based optimization technique that enables a number of individual solutions, called particles, to move through a hyperdimensional search space to search for the optimum or optima. Each particle has a position vector \( \mathbf{x} \) and a velocity vector \( \mathbf{v} \), which are adjusted at iterations by learning from a personal best solution \( \mathbf{pbest} \), found by the particle itself and a global best solution \( \mathbf{gbest} \), found by the whole swarm so far at each dimension \( j \):

\[
\mathbf{v}_{i,j}^{(t+1)} = \omega \mathbf{v}_{i,j}^{(t)} + c_1 r_1 (\mathbf{pbest}_{i,j}^{(t)} - \mathbf{x}_{i,j}^{(t)}) + c_2 r_2 (\mathbf{gbest}_{j}^{(t)} - \mathbf{x}_{i,j}^{(t)}),
\]

\[
\mathbf{x}_{i,j}^{(t+1)} = \mathbf{x}_{i,j}^{(t)} + \mathbf{v}_{i,j}^{(t+1)},
\]

where \( \omega \) is the inertia weight for controlling the velocity, \( c_1 \) and \( c_2 \) are the acceleration coefficients reflecting the weighting of “cognitive” and “social” learning, respectively, and \( r_1 \) and \( r_2 \) are two random numbers ranging from 0 to 1. Empirical studies have shown that PSO has a high efficiency in convergence to desirable optima and performs better than GAs in many problems [28–34].

Most PSO algorithms use a linear decreasing inertia weight and constant acceleration coefficients in (5). In recent years, some work has been conducted on the nonlinear and nonmonotonous change of control parameters, which has shown the ability of increasing the adaptivity and diversity of the swarm. In the work of Yang et al. [35], the inertia weight is adapted based on two characteristic parameters in the search course of PSO, namely, speed factor and aggregation degree factor. Arumugam and Rao [36] used the ratio of the global best fitness and the average of particles’ personal best fitness to determine the inertia weight in each iteration. Panigrahi...
et al. [37] proposed another PSO method where different particles are assigned with different inertia weights based on the ranks of the particles. In [38] Nickabadi et al. proposed a new adaptive inertia weight approach that uses the success rate of the swarm as its feedback parameter to ascertain the particles’ situation in the search space.

Zhan et al. [30] proposed an adaptive PSO that differentiates the evolutionary states as exploration, exploitation, convergence, and jumping out by evaluating the population distribution and particle fitness and uses the state information to automatically control inertia weight and acceleration coefficients at run time. Ratnaweera et al. [39] introduced a time-varying acceleration coefficient strategy where the cognitive coefficient is linearly decreased and the social coefficient is linearly increased. In [40] Zheng et al. introduced a new adaptive parameter strategy to the comprehensive learning particle swarm optimizer, which uses evolutionary information of individual particles to dynamically adapt the inertia weight and acceleration coefficient at each iteration.

3.3. An Adaptive PSO Algorithm for Evolving the ANN. Since the search process of PSO is highly complicated in the ANN training problems, here we use a self-adaptive mechanism for both the inertia weight and the acceleration coefficients. The mechanism is based on the concept of the success of a particle at an iteration, that is, whether or not a particle improves its personal best at iteration t:

\[
s_i(t) = \begin{cases} 
1, & \text{if } f(p_{\text{best}}^{(t+1)}) < f(p_{\text{best}}^{(t)}) \\
0, & \text{else.}
\end{cases}
\]  

(6)

Based on that we compute the rate of success of the whole swarm S:

\[
\rho^{(t)} = \frac{\sum s_i^{(t)}}{|S|}.
\]  

(7)

That is, \( \rho \in [0,1] \) is the percentage of the particles which have had an improvement in the last iteration. A high value of \( \rho \) indicates a high probability that the particles have converged to a nonoptimum point or slowly moving toward the optimum, while a low value of \( \rho \) indicates that the particles are oscillating around the optimum without much improvement [38]. And we use the following equation for dynamically changing inertia weight of PSO:

\[
w^{(t)} = e^{\rho^{(t)} - 1}.
\]  

(8)

PSO is a population-based optimization method, which is preferred to encourage the exploration during the early stages and to enhance exploitation during the latter stages. Here we also employ a mechanism that linearly decreases the cognitive coefficient and increases the social coefficient as follows:

\[
c_1^{(t)} = c_{1\text{max}} - \frac{t}{t_{\text{max}}} (c_{1\text{max}} - c_{1\text{min}}),
\]  

(9)

\[
c_2^{(t)} = c_{2\text{min}} + \frac{t}{t_{\text{max}}} (c_{2\text{max}} - c_{2\text{min}}).
\]  

The dataset of our forecasting problem is partitioned into three sets: a training set, a validation set, and a testing set. Our adaptive PSO contains a main procedure for evolving a swarm of ANNs with different structure (different number of the hidden nodes and the connection density), which is outlined in Algorithm 1, and a subprocedure for evolving a swarm of ANNs with the same structure but different connection weights, which is outlined in Algorithm 2. For each particle that represents an ANN, its fitness is defined as the root mean-squared error (RMSE) of the training dataset, which is calculated as follows:

\[
\text{RMSE} = \sqrt{\frac{\sum_{k=1}^{n} (y_k - y_k^*)^2}{N}},
\]  

(10)

where \( y_k \) is the network output of the \( k \)th training tuple and \( y_k^* \) is its desired output. The fitness of particles in the main-swarm is evaluated as the best ANN obtained by its subswarms.

The flowchart of the algorithm is shown in Figure 2.

4. Computational Experiment

The proposed method was applied to the actual historical data collected from seven equipment maintenance organizations.
(1) Randomly initialize a sub-swarm of $n'$ particles, each of which representing the weights of the $i$th ANN;
(2) Let $t' = 0$; set the pbest of each particle to its original position;
(3) For $t' = 1$ to $t'_{\text{max}}$ do
(4) For each particle in the $i$th sub-swarm do
(5) Move the particle in the sub search space according to (4) and (5);
(6) Update the pbest of the particle;
(7) Update the gbest of the $i$th sub-swarm;
(8) Update the control parameters of the inner loop according to (6)–(9);
(9) return the gbest of the $i$th sub-swarm as the $i$th ANN;

**Algorithm 2**: The pseudocode of the sub-PSO procedure.

**Figure 2**: The flowchart of the proposed approach.

from 1999 to 2011. We used the data from 1999 to 2006 and the data of 2008 as the training dataset, the data of 2007 and 2010 as the validation dataset, and the data of 2009 and 2011 as the testing dataset. Table 2 presents a summarization of the datasets.

We, respectively, implement three other approaches: the ANN trained by BP (denoted by ANN), the ANN trained by basic PSO (denoted by ANN-PSO), and the nonlinear regression model optimized by PSO (denoted by R-PSO) [41], in order to comparatively evaluate the performance of our method (denoted by ANN-APSO).

The experiments are conducted on a computer with Intel i5 3450 (4 × 3.1 G) multicore processor and DDR3 8 GB memory. The basic parameters of our approach are set as $c_{1\min} = c_{2\min} = 0.5$, $c_{1\max} = c_{2\max} = 2.0$, $t_{\max} = 1000$, $t_{\max} = 80$, $n = 40$, and $n' = 100$. The maximum running time of the other approaches is set as they consume nearly the same CPU time as our method.

Tables 3 and 4, respectively, present the experimental results of the four approaches on the testing datasets of 2009 and 2011. As we can see from the results, on every test data tuple, our ANN trained by adaptive PSO achieves the
Table 2: The summarization of the historical datasets on electrical energy consumption of equipment maintenance, where $N_E$ is the number of equipment repaired and $E_C$ is the total electrical energy consumption (in thousand kWh).

| Year | Tuples | Org1 $N_E$ | $E_C$ | Org2 $N_E$ | $E_C$ | Org3 $N_E$ | $E_C$ | Org4 $N_E$ | $E_C$ | Org5 $N_E$ | $E_C$ | Org6 $N_E$ | $E_C$ | Org7 $N_E$ | $E_C$ |
|------|--------|------------|------|------------|------|------------|------|------------|------|------------|------|------------|------|------------|------|
| 1999 | 12     | 48         | 39.2 | 65         | 50.5 | 52         | 39.9 | 37         | 33.6 | 28         | 25.0 | 50         | 49.3 | 55         | 48.7 |
| 2000 | 15     | 50         | 43.6 | 60         | 49.3 | 55         | 48.7 | 58         | 60.4 | 55         | 52.4 | 50         | 49.3 | 55         | 48.7 |
| 2001 | 15     | 54         | 58.2 | 64         | 60.3 | 56         | 57.1 | 56         | 61.0 | 58         | 60.0 | 50         | 49.3 | 55         | 48.7 |
| 2002 | 17     | 57         | 66.7 | 63         | 68.2 | 55         | 62.2 | 58         | 70.2 | 57         | 67.5 | 38         | 33.6 | 28         | 25.0 |
| 2003 | 18     | 61         | 72.8 | 66         | 81.6 | 59         | 68.9 | 60         | 77.8 | 58         | 72.3 | 36         | 33.6 | 28         | 25.0 |
| 2004 | 21     | 66         | 91.0 | 68         | 88.3 | 58         | 70.3 | 61         | 85.0 | 68         | 92.6 | 60         | 82.2 | 55         | 72.1 |
| 2005 | 20     | 72         | 113  | 68         | 95.8 | 62         | 81.7 | 65         | 102  | 70         | 107  | 68         | 98.0 | 59         | 83.5 |
| 2006 | 21     | 74         | 118  | 72         | 106  | 69         | 105  | 68         | 113  | 69         | 115  | 72         | 121  | 61         | 102  |
| 2007 | 21     | 75         | 123  | 78         | 137  | 65         | 116  | 73         | 130  | 75         | 136  | 78         | 148  | 63         | 103  |
| 2008 | 18     | 89         | 161  | 78         | 145  | 75         | 150  | 82         | 153  | 92         | 169  | 80         | 136  | 56         | 50.3 |
| 2009 | 21     | 72         | 132  | 79         | 151  | 72         | 138  | 78         | 156  | 76         | 148  | 82         | 160  | 65         | 116  |
| 2010 | 21     | 80         | 158  | 82         | 159  | 70         | 139  | 76         | 153  | 76         | 155  | 81         | 155  | 66         | 118  |
| 2011 | 21     | 83         | 169  | 80         | 160  | 72         | 145  | 75         | 149  | 80         | 171  | 80         | 151  | 67         | 133  |

Table 3: The MAPE (%) associated with the four approaches on the testing dataset of 2009.

| Organization | ANN  | ANN-PSO | R-PSO | ANN-APSO |
|--------------|------|---------|------|----------|
| #1           | 1.816\textsuperscript{+} | 1.220\textsuperscript{+} | 1.093\textsuperscript{+} | 0.835\textsuperscript{+} |
| #2           | 1.493\textsuperscript{+} | 0.000\textsuperscript{+} | 0.355\textsuperscript{+} | 0.000\textsuperscript{+} |
| #3           | 5.267\textsuperscript{+} | 1.801\textsuperscript{+} | 1.103\textsuperscript{+} | 1.004\textsuperscript{+} |
| #4           | 2.205\textsuperscript{+} | 0.934\textsuperscript{+} | 0.783\textsuperscript{+} | 0.297\textsuperscript{+} |
| #5           | 1.315\textsuperscript{+} | 0.303\textsuperscript{+} | 0.526\textsuperscript{+} | 0.067\textsuperscript{+} |
| #6           | 3.655\textsuperscript{+} | 1.500\textsuperscript{+} | 1.218\textsuperscript{+} | 0.722\textsuperscript{+} |
| #7           | 0.452\textsuperscript{+} | 0.000\textsuperscript{+} | 0.000\textsuperscript{+} | 0.000\textsuperscript{+} |

In column 2–4, \textsuperscript{+} means that the result of ANN-APSO has statistically significant improvement over that of the comparative algorithm (using paired one-tailed $t$-test at a 0.05 level of significance).

Table 4: The MAPE (%) associated with the four approaches on the testing dataset of 2011.

| Organization | ANN  | ANN-PSO | R-PSO | ANN-APSO |
|--------------|------|---------|------|----------|
| #1           | 3.357\textsuperscript{+} | 1.472\textsuperscript{+} | 1.164\textsuperscript{+} | 0.767\textsuperscript{+} |
| #2           | 6.060\textsuperscript{+} | 1.833\textsuperscript{+} | 2.261\textsuperscript{+} | 0.935\textsuperscript{+} |
| #3           | 8.248\textsuperscript{+} | 2.707\textsuperscript{+} | 2.095\textsuperscript{+} | 1.363\textsuperscript{+} |
| #4           | 5.131\textsuperscript{+} | 0.896\textsuperscript{+} | 0.803\textsuperscript{+} | 0.520\textsuperscript{+} |
| #5           | 3.690\textsuperscript{+} | 0.422\textsuperscript{+} | 0.693\textsuperscript{+} | 0.193\textsuperscript{+} |
| #6           | 2.851\textsuperscript{+} | 1.658\textsuperscript{+} | 1.827\textsuperscript{+} | 0.079\textsuperscript{+} |
| #7           | 0.941\textsuperscript{+} | 0.081\textsuperscript{+} | 0.235\textsuperscript{+} | 0.000\textsuperscript{+} |

In column 2–4, \textsuperscript{+} means that the result of ANN-APSO has statistically significant improvement over that of the comparative algorithm (using paired one-tailed $t$-test at a 0.05 level of significance).

(i) On the dataset of 2009, our ANN-APSO has significant performance improvement over ANN in all cases and over ANN-PSO and R-PSO in most cases.

(ii) On the dataset of 2011, the ANN-APSO always has significant performance improvement over the other three methods.

In summary, ANN-APSO provides the best forecasting accuracy on the testing datasets, and it is much more reliable and robust than the other three approaches.
5. Conclusion

The paper presents an approach for forecasting electrical energy consumption of equipment maintenance based on ANN trained by PSO. The ANN model is used for modeling possible relationships between the input variables and the expected electrical energy consumption, and a new PSO algorithm with adaptive control parameters is developed for evolving the ANN. Experimental results demonstrate that our approach performs better than some state-of-the-art methods on the historical data of equipment maintenance in China. And we believe that the proposed method can be applied to a wider range of forecasting problems in engineering.

In practice, we find that it is often difficult to get exact values of a certain number of input variables, and currently we just use their expected or estimated values. To further improve the adaptivity and flexibility of our approach, we are testing the use of fuzzy numbers and fuzzy-neural networks in forecasting and adapting our ANN-APSO algorithm for training and evolving the fuzzy-neural networks. Our future work also includes improving the algorithm by information sharing between different subswarms.

Acknowledgments

This work was supported by the Natural Science Foundation of China under Grant nos. 61020106009 and 61105073. The authors are grateful to the reviewers’ valuable comments that help to improve the paper greatly.

References

[1] Z. Q. Zhang, J. P. Wang, X. Y. Duan et al., Introduction To Equipment Technical Support, Military Science Press, Beijing, China, 2001.
[2] P. Crompton and Y. Wu, “Energy consumption in China: past trends and future directions,” Energy Economics, vol. 27, no. 1, pp. 195–208, 2005.
[3] D. Niu and M. Meng, “Research on seasonal increasing electric energy demand forecasting: a case in China,” Chinese Journal of Management Science, vol. 18, no. 2, pp. 108–112, 2010.
[4] N. Wada, O. Saito, Y. Yamamoto, T. Morioka, and A. Tokai, “Evaluating effects on the flow of electrical and electronic equipment and energy consumption due to alternative consumption patterns in China,” Journal of Risk Research, vol. 15, no. 1, pp. 107–130, 2012.
[5] H. K. AlFares and M. Nazeeruddin, “Lectic load forecasting: literature survey and classification of methods,” International Journal of Systems Science, vol. 33, no. 1, pp. 23–24, 2002.
[6] G. A. N. Mhamalu and M. E. El-Hawary, “Load forecasting via suboptimal seasonal autoregressive models and iteratively reweighted least squares estimation,” IEEE Transactions on Power Systems, vol. 8, no. 1, pp. 343–348, 1993.
[7] T. Haida and S. Muto, “Regression based peak load forecasting using a transformation technique,” IEEE Transactions on Power Systems, vol. 9, no. 4, pp. 1788–1794, 1994.
[8] A. Al-Shehri, “A simple forecasting model for industrial electric energy consumption,” International Journal of Energy Research, vol. 24, no. 8, pp. 719–726, 2000.
[9] H. S. Hippert, C. E. Pedreira, and R. C. Souza, “Neural networks for short-term load forecasting: a review and evaluation,” IEEE Transactions on Power Systems, vol. 16, no. 1, pp. 44–55, 2001.
[10] K.-H. Kim, J.-K. Park, K.-J. Hwang, and S.-H. Kim, “Implementation of hybrid short-term load forecasting system using artificial neural networks and fuzzy expert systems,” IEEE Transactions on Power Systems, vol. 10, no. 3, pp. 1534–1539, 1995.
[11] K. Chau, “A review on the integration of artificial intelligence into coastal modeling,” Journal of Environmental Management, vol. 80, no. 1, pp. 47–57, 2006.
[12] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, “Learning representations by back-propagating errors,” Nature, vol. 323, no. 6088, pp. 533–536, 1986.
[13] S. Chen and Z. Wang, “Acceleration strategies in generalized belief propagation,” IEEE Transactions on Industrial Informatics, vol. 8, no. 1, pp. 41–48, 2012.
[14] G. F. Miller, P. M. Todd, and S. U. Hegde, “Designing neural networks using genetic algorithms,” in Proceedings of the 3rd International Conference on Genetic Algorithms, J. David Schaffer, Ed., pp. 379–384, Morgan Kaufmann Publishers, San Francisco, Calif, USA, 1989.
[15] C. Zhang, H. Shao, and Y. Li, “Particle swarm optimization for evolving artificial neural network,” in Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, vol. 4, pp. 2487–2490, October 2000.
[16] J. D. Schaffer, D. Whitley, and L. J. Eshelman, “Combinations of genetic algorithms and neural networks: a survey of the state of the art,” in Proceedings of the International Workshop on Combinations of Genetic Algorithms and Neural Networks (COGANN’92), pp. 1–37, 1992.
[17] M. Settles, B. Rodebaugh, and T. Soule, “Comparison of genetic algorithm and particle swarm optimizer when evolving a recurrent neural network,” in Genetic and Evolutionary Computation, E. Cantú-Paz, J. A. Foster, K. Deb et al., Eds., vol. 2723 of Lecture Notes in Computer Science, pp. 148–149, Springer, New York, NY, USA, 2003.
[18] M. Kawashima, “Artificial neural network backpropagation model with three-phase annealing developed for the building energy predictor shootout,” ASHRAE Transactions, vol. 100, no. 2, pp. 1096–1103, 1994.
[19] S. M. Islam, S. M. Al-Alawi, and K. A. Ellithy, “Forecasting monthly electric load and energy for a fast growing utility using an artificial neural network,” Electric Power Systems Research, vol. 34, no. 1, pp. 1–9, 1995.
[20] A. Al-Shehri, “Artificial neural network for forecasting residential electrical energy,” International Journal of Energy Research, vol. 23, no. 8, pp. 649–661, 1999.
[21] N. D. Hatzigiorgiou and E. N. Dialynas, “An optimized adaptive neural network for annual midterm energy forecasting,” IEEE Transactions on Power Systems, vol. 21, no. 1, pp. 385–391, 2006.
[22] A. Sözen, M. A. Akçayol, and E. Arcaklioğlu, “Forecasting net electricity demand via suboptimal seasonal autoregressive models and iteratively reweighted least squares estimation,” IEEE Transactions on Power Systems, vol. 8, no. 1, pp. 41–48, 2012.
[23] E. Cantú-Paz, J. A. Foster, K. Deb et al., Eds., vol. 2723 of Lecture Notes in Computer Science, pp. 148–149, Springer, New York, NY, USA, 2003.
[24] A. Al-Shehri, “Artificial neural network for forecasting residential electrical energy,” International Journal of Energy Research, vol. 23, no. 8, pp. 649–661, 1999.
[25] N. D. Hatzigiorgiou and E. N. Dialynas, “An optimized adaptive neural network for annual midterm energy forecasting,” IEEE Transactions on Power Systems, vol. 21, no. 1, pp. 385–391, 2006.
[26] A. Sözen, M. A. Akçayol, and E. Arcaklioğlu, “Forecasting net electricity demand via suboptimal seasonal autoregressive models and iteratively reweighted least squares estimation,” IEEE Transactions on Power Systems, vol. 8, no. 1, pp. 41–48, 2012.
[27] K. Ermis, A. Midilli, I. Dincer, and M. A. Rosen, “Artificial neural network analysis of world green energy use,” Energy Policy, vol. 35, no. 3, pp. 1731–1743, 2007.
[28] A. Azadeh, S. F. Ghaderi, and S. Sohrabkhani, “A simulated-based neural network algorithm for forecasting electrical energy consumption in Iran,” Energy Policy, vol. 36, no. 7, pp. 2637–2644, 2008.
[25] R. Yokoyama, T. Wakui, and R. Satake, “Prediction of energy demands using neural network with model identification by global optimization,” Energy Conversion and Management, vol. 50, no. 2, pp. 319–327, 2009.

[26] Z. W. Geem and W. E. Roper, “Energy demand estimation of South Korea using artificial neural network,” Energy Policy, vol. 37, no. 10, pp. 4049–4054, 2009.

[27] S. Haykin, Neural Networks: A Comprehensive Foundation, Prentice-Hall, Englewood Cliffs, NJ, USA, 1994.

[28] J. Kennedy, “Bare bones particle swarms,” in Proceedings of the IEEE Swarm Intelligence Symposium, pp. 80–87, 2003.

[29] N. M. Kwok, D. K. Liu, K. C. Tan, and Q. P. Ha, “An empirical study on the settings of control coefficients in particle swarm optimization,” in Proceedings of the IEEE Congress on Evolutionary Computation (CEC ’06), pp. 3165–3172, Vancouver, Canada, July 2006.

[30] Z. H. Zhan, J. Zhang, Y. Li, and H. S. H. Chung, “Adaptive particle swarm optimization,” IEEE Transactions on Systems, Man, and Cybernetics B, vol. 39, no. 6, pp. 1362–1381, 2009.

[31] Y. J. Zheng, Q. Song, and S. Y. Chen, “Multiobjective fireworks optimization for variable-rate fertilization in oil crop production,” Applied Soft Computing, vol. 13, no. 11, pp. 4253–4263, 2013.

[32] P. Lu, S. Chen, and Y. Zheng, “Artificial intelligence in civil engineering,” Mathematical Problems in Engineering, vol. 2012, Article ID 145974, 22 pages, 2012.

[33] Y. J. Zheng, S. Y. Chen, Y. Lin, and W. L. Wang, “Bio-inspired optimization of sustainable energy systems: a review,” Mathematical Problems in Engineering, vol. 2013, Article ID 354523, 12 pages, 2013.

[34] Y. J. Zheng and S. Y. Chen, “Cooperative particle swarm optimization for multiobjective transportation planning,” Applied Intelligence, vol. 39, no. 1, pp. 202–216, 2013.

[35] X. Yang, J. Yuan, and H. Mao, “A modified particle swarm optimizer with dynamic adaptation,” Applied Mathematics and Computation, vol. 189, no. 2, pp. 1205–1213, 2007.

[36] M. S. Arumugam and M. V. C. Rao, “On the improved performances of the particle swarm optimization algorithms with adaptive parameters, cross-over operators and root mean square (RMS) variants for computing optimal control of a class of hybrid systems,” Applied Soft Computing Journal, vol. 8, no. 1, pp. 324–336, 2008.

[37] B. K. Panigrahi, V. R. Pandi, and S. Das, “Adaptive particle swarm optimization approach for static and dynamic economic load dispatch,” Energy Conversion and Management, vol. 49, no. 6, pp. 1407–1415, 2008.

[38] A. Nickabadi, M. M. Ebadzadeh, and R. Safabakhsh, “A novel particle swarm optimization algorithm with adaptive inertia weight,” Applied Soft Computing Journal, vol. 11, no. 4, pp. 3658–3670, 2011.

[39] A. Ratnaweera, S. K. Halgamuge, and H. C. Watson, “Self-organizing hierarchical particle swarm optimizer with time-varying acceleration coefficients,” IEEE Transactions on Evolutionary Computation, vol. 8, no. 3, pp. 240–255, 2004.

[40] Y. J. Zheng, H. F. Ling, and Q. Guan, “Adaptive parameters for a modified comprehensive learning particle swarm optimizer,” Mathematical Problems in Engineering, vol. 2012, Article ID 207318, 11 pages, 2012.

[41] M. AlRashidi and K. M. EL-Naggar, “Long term electric load forecasting based on particle swarm optimization,” Applied Energy, vol. 87, no. 1, pp. 320–326, 2010.
