

Research Article

Electricity Load Forecasting Using Support Vector Regression with Memetic Algorithms

Zhongyi Hu, Yukun Bao, and Tao Xiong

Department of Management Science and Information Systems, School of Management, Huazhong University of Science and Technology, Wuhan 430074, China

Correspondence should be addressed to Yukun Bao; yukunbao@mail.hust.edu.cn

Received 25 September 2013; Accepted 19 November 2013

Academic Editors: C. Koroneos and X. Zhou

Copyright © 2013 Zhongyi Hu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Electricity load forecasting is an important issue that is widely explored and examined in power systems operation literature and commercial transactions in electricity markets literature as well. Among the existing forecasting models, support vector regression (SVR) has gained much attention. Considering the performance of SVR highly depends on its parameters; this study proposed a firefly algorithm (FA) based memetic algorithm (FA-MA) to appropriately determine the parameters of SVR forecasting model. In the proposed FA-MA algorithm, the FA algorithm is applied to explore the solution space, and the pattern search is used to conduct individual learning and thus enhance the exploitation of FA. Experimental results confirm that the proposed FA-MA based SVR model can not only yield more accurate forecasting results than the other four evolutionary algorithms based SVR models and three well-known forecasting models but also outperform the hybrid algorithms in the related existing literature.

1. Introduction

Electricity load forecasting has always been the essential part of efficient power system planning and operation. Specially that it is not only critical for automatic generation control, reliable operation, and resource dispatch but is also a fundamental piece of information used for energy transactions in competitive electricity markets [1]. Inaccurate forecast of power load leads to a great deal of loss for power companies, and a 1% increase in forecasting error implied a 10 million increase in operating costs [2]. However, the electricity load is inevitably affected by various factors such as climate factors, social activities, and seasonal factors, thus making it difficult to be accurately predicted.

During the past several decades, numerous approaches have been proposed for electricity load forecasting. Traditional methods, such as autoregressive moving average model (ARMA) [3], exponential smoothing models [4, 5], and regression models [6, 7] are often difficult to model the electricity load with high accuracies due to the nonlinearity of the load inherently. On the other hand, with the development of intelligence techniques in recent years, many studies have tried to apply the artificial intelligence techniques to improve

the forecasting accuracy of load. Among them, neural networks (NN) have received much share of attention, and a great number of studies have reported successful results in the load forecasting [8]. Reference [9] proposed a practical method using NN combined similar days approach, which resulted in a reliable forecasts for one-to-six hour-ahead electricity load. Reference [10] proposed an adaptive artificial neural network with particle swarm optimization (PSO) used to adjust the network's weights; computational results indicated that the proposed model can obtain higher forecasting precision with traditional BP algorithm. Reference [11] applied Bayesian neural network in the short-term load forecasting and the results of the proposed model gains better performance than that of conventional neural networks. Neural network with novel learning algorithm based on a modified harmony search technique also gives better results than several benchmarks [12]. Reference [8] gave a comprehensive review and evaluation of the neural networks for short-term load forecasting. However, the NN has a large number of parameters to be tuned and suffers from the danger of over-fitting.

Different from NN which minimizes the empirical error based on the empirical risk minimization principle (ERM), support vector regression (SVR) implements the structural

risk minimization principle (SRM) by minimizing an upper bound to the generalization error [13]. This leads to excellent generalization performance with SVR, which has been shown to outperform other nonlinear forecasting techniques including NN based forecasting models [14]. As one of the fields in time series forecasting using SVR [14–20], electricity load forecasting using SVR as well as its varieties has been well studied. Reference [16] proposed a locally weighted support vector regression to solve the short-term load forecasting problem, and the experimental results proved the superior performance of the proposed model compared with some published models. Combined with fuzzy c-means (FCM) and particle swarm optimization (PSO), a SVR based model is studied to forecast the short-term load of a city [21]. Reference [17] proposed a TF- ϵ -SVR model with trend fixed and seasonal adjustment to improve the forecasting accuracy of the electricity demand.

However, the generalization ability of SVMs highly depends on the adequate setting of parameters [22–25], such as penalty coefficient, kernel parameters, and the width of loss function. Therefore, the selection of the optimal parameters is of critical importance to obtain a good performance in handling electricity load forecasting task with SVR. Recently, various studies try to improve the forecasting accuracy of electricity load when the SVR model is used. Reference [18] employs simulated annealing algorithms (SA) to choose the parameters of SVR, and computational results show that the SA based SVR model achieves better performance for load forecasting compared with autoregressive integrated moving average (ARIMA) model and the general regression neural networks (GRNN) model. Reference [26] proposed a differential evolution algorithm based SVR model to forecast the annual load. Reference [27] proposed a LSSVM based load forecasting model with the fruit fly algorithm used to automatically choose the parameters of LSSVM; experimental results show that the proposed model outperforms some other alternative models. Pai et al. conducted a series of relevant researches by using genetic algorithm (GA) [28, 29], chaotic particle swarm optimization (CPSO) [30], artificial bee colony algorithm (ABC) [31], immune algorithm (IA) [32, 33], and hybrid algorithm [34] for parameters determination of SVR to improve the forecasting accuracy of the electricity load. However, GA and some other evolutionary algorithms (EAs) are not guaranteed to find the global optimum parameters of a SVR model, though they are generally good at finding “acceptable good” or near-optimal solutions to problems. More specifically, although they are good at exploring the solution space and detecting the region of attraction of the global optimum efficiently, they lack the abilities to perform a refined tuning search locally [24, 35].

Memetic algorithms (MAs), a powerful algorithmic paradigm that combines the evolutionary algorithms (EAs) with problem-specific local searcher (LS), have been successfully applied in a wide variety of areas [36–38]. MA has the ability to exploit the complementary advantages of EAs (generality, robustness, and global search efficiency), and problem-specific local search (exploiting application-specific problem structure, rapid convergence toward local minima) [39]. In our previous study [24], MA is proposed to tune

the parameters of SVM in classification problems. However, there are, if any, few works related to MA that have been reported in the load forecasting literature on the issue of SVR parameters optimization. Notice that there are three important parameters in SVR, whereas SVM for classification has only two. With the increase of dimensions and the change of structure complexity for the optimization problem, the performance of MA is a big challenge. As such, it is of interest to involve the MAs for SVR parameters optimization in improving the prediction accuracy of STLF. In this study, by combining firefly algorithm (FA) and pattern search (PS), an efficient FA based memetic algorithm (FA-MA) is proposed to automatically determine the parameters of SVR for improving the forecasting accuracy of electricity load forecasting. In the proposed FA-MA, FA is responsible for the exploration of the search space and the detection of the potential regions with optimum solution, while PS is used to produce an effective exploitation on the potential regions obtained by FA. The performance of proposed FA-MA for parameters optimization in SVR is justified on two real-world cases against selected counterparts.

The rest of the study is organized as follows. Section 2 presents a brief review on SVMs. Section 3 elaborates on the FA-MA proposed in this study. The results with discussions are reported in Section 4. Finally, we conclude this study in Section 5.

2. Support Vector Regression

Given a set of training data, $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\} \subset R^n \times R$, where x_i is input vector, y_i is the target, and n denotes the number of the data items in the training set. Based on the structured risk minimization (SRM) principle [13], rather than finding minimum empirical errors, SVMs aim to generate a decision function (1) by minimizing a regularized risk function (2)

$$f(x) = \langle w, \phi(x) \rangle + b, \quad (1)$$

$$R = R_{\text{emp}} + \frac{1}{2} \|w\|^2$$

$$= \frac{C}{n} \sum_{i=1}^n L(y_i, f(x)) + \frac{1}{2} \|w\|^2, \quad (2)$$

$$L(y_i, f(x)) = \begin{cases} |y - f(x)| - \epsilon & \text{if } |y - f(x)| \geq \epsilon \\ 0 & \text{otherwise,} \end{cases} \quad (3)$$

where in (1), $\langle \cdot, \cdot \rangle$ denotes the inner product, w is the weight vector, that controls the smoothness of the model, and b is a parameter of bias. $\phi(x)$ is the high-dimensional feature space which is nonlinearly mapped from the input space x . In the regularized risk function given by (2), the first term R_{emp} or $(C/n) \sum_{i=1}^n L(y_i, f(x))$ is the empirical risk. In SVR, Vapnik's ϵ -insensitive loss function [40] given by (3) is often used to measure the empirical risk, and ϵ is called the tube size. The second term, $(1/2)\|w\|^2$, is the regularization term to be used as a measure of flatness or complexity of the function. Hence, C is referred to as the regularization constant and it specifies

the trade off between the empirical risk and the regularization term. Both C and ε are user-determined parameters.

By introducing two positive slack variables ξ and ξ^* , (2) is transformed into the following constrained form:

$$\begin{aligned} \text{Minimize} \quad & R = C \sum_{i=1}^n (\xi_i + \xi_i^*) + \frac{1}{2} \|w\|^2 \\ \text{Subjected to} \quad & y_i - \langle w, x_i \rangle - b \leq \varepsilon + \xi_i \\ & \langle w, x_i \rangle + b - y_i \leq \varepsilon + \xi_i^* \\ & \xi_i, \xi_i^* \geq 0, \quad i = 1, 2, \dots, n. \end{aligned} \quad (4)$$

According to Wolfe’s dual theorem and the saddle-point condition, the dual optimization problem of the above primal one is obtained as in the following form:

$$\begin{aligned} \max_{\alpha, \alpha^*} \quad & -\frac{1}{2} \sum_{i,j=1}^l (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) \langle \phi(x_i), \phi(x_j) \rangle \\ & - \varepsilon \sum_{i,j=1}^l (\alpha_i + \alpha_i^*) + \sum_{i,j=1}^l y_i (\alpha_i - \alpha_i^*) \\ \text{s.t.} \quad & \sum_{i,j=1}^l (\alpha_i - \alpha_i^*) = 0, \quad \alpha_i, \alpha_i^* \in [0, C] \end{aligned} \quad (5)$$

with

$$w = \sum_{i=1}^n (\alpha_i - \alpha_i^*) \phi(x_i), \quad (6)$$

where α_i, α_i^* are nonnegative Lagrange multipliers that can be obtained by solving the convex quadratic programming problem stated above.

Finally, based on the (6) and the trick of kernel function, the decision function given by (1) has the following explicit form:

$$f(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) K(x_i, x_j) + b. \quad (7)$$

Here, $K(x_i, x_j)$ is defined as kernel function. The value of the kernel function is equivalent to the inner product of two vectors x_i and x_j in the feature space $\phi(x_i)$ and $\phi(x_j)$; that is, $K(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle$. The elegance of using the kernel function is that one can deal with feature spaces of arbitrary dimensionality without having to compute the map $\phi(x)$ explicitly. Any function that satisfies Mercer’s condition [40] can be used as the kernel function. There are several typical examples of kernel function such as linear kernel, polynomial kernel, radial basis function (RBF), and sigmoid kernel. Each kernel has some parameters. Generally, among these kernel functions, RBF kernel (8) is strongly recommended and widely used for its performance and complexity [41] and thus SVR with RBF kernel function is the one studied in this study.

Consider

$$K(x_i, x_j) = \exp\left(\frac{-\|x_i - x_j\|^2}{2\delta^2}\right), \quad (8)$$

where δ is kernel parameter. The kernel parameter should be carefully chosen as it implicitly defines the structure of

the high-dimensional feature space $\phi(x)$ and thus controls the complexity of the model.

Overall, SVR is a powerful learning machine with strong theoretical foundations and excellent generalization performance. Note that before implementing the SVR with RBF kernel, there are three parameters (penalty parameter C , RBF kernel parameter δ , and width of ε loss function) to be set. Previous studies show that these three parameters play an important role in the success of SVR [42]. In this study, to determine these parameters and to improve the forecasting accuracy of SVR in electricity load forecasting, a firefly algorithm (FA) based memetic algorithm (FA-MA) is proposed in Section 3.

3. Memetic Algorithm for Parameters Selection of SVR

Memetic algorithms (MAs), one of the recent growing areas in computational intelligence, is first coined by Moscato and Norman [43]. Inspired by Darwinian principles of natural evolution and Dawkins’ notion of meme, MA has come to light as an union of population based stochastic global evolutionary algorithm and local improvement procedures. As a designed a hybridization, MAs are expected to make full use of the balance between exploration and exploitation of the search space to complement the advantages of population based methods and local based methods. Nowadays, MAs have revealed their successes with high performance and superior robustness across a wide range of problem domains; detail reviews are reported in [44, 45]. Since often there are no free lunches, the hybridization can be more complex and expensive to implement. Considering the effectiveness of firefly algorithm which is introduced recently and can be even superior to the GA and PSO [46–48], this study proposed a FA based memetic algorithm with pattern search as a local individual learner, to improve the forecasting accuracy of electricity load forecasting model using SVR.

In the following subsections, we will explain the implementation of the proposed FA-MA for parameters optimization in details.

3.1. Initialization. In the proposed FA-MA, each firefly (or individual) is a parameter set of the SVR model and can be denoted as $\mathbf{x}_i = \langle C, \delta, \varepsilon \rangle$. A set of fireflies is called a swarm or population. Traditionally, initial swarm is often generated randomly in firefly algorithm or other evolutionary algorithms. To guarantee an initial swarm with reliability and diversity, Latin hypercube sampling (LHS) method is applied to generate a random sample set. With the use of LHS, we first split the search space into subspaces and then try to take randomly the values within each subspace to achieve an initial sample set which is representative of the whole search space. Hence, it can guarantee the initial samples to be relatively uniformly distributed over each dimension, which is proved to be superior to random initialization [49].

3.2. Fitness Function. Since the ultimate goal of the SVR model is to forecast the future electricity load with high

accuracy (i.e., known as generalization ability), it is important to choose such fitness function which can estimate the generalization ability when determining the parameters in SVR with FA-MA. In this study, the data is split into three parts which are training set, validation set, and testing set. The training set is used to train the SVR model with a certain parameter set, and the validation set is deserved to assess the generalization ability of the established forecasting model. The parameter set with lowest mean squared percentage error (MAPE) (For convenience, the formulation of MAPE is given in Section 4.2.) in the validation set is selected as the optimal solution. That is to say, MAPE in the validation set is used as the fitness function. In the proposed firefly algorithm based memetic algorithm, the brightness or light intensity of a firefly is determined by the fitness function.

3.3. Exploration with Firefly Algorithms (FAs). Firefly algorithm, first introduced by Yang et al. [46, 47], is a swarm based intelligent metaheuristic. The FA mimics the social behavior of fireflies which move and communicate with each other based on their flashing characteristics, such as brightness, frequency, and the time period. Specially, the superiority of FA against genetic algorithms (GAs) and particle swarm optimization (PSO) in existing studies [46, 47] motivates us to use the FA to explore the search space.

In FA, each firefly is assumed to be attracted to other ones regardless of their sex, and the attractiveness is proportional to their brightness. Besides, as mentioned before, the brightness of a firefly is determined by the fitness function. To minimize the fitness defined in Section 3.2, the brightness can simply be minus of the MAPE.

The movement of a firefly i attracted by another more attractive firefly j can be formulated as (Details of the definition are shown in Yang [46].)

$$v_i = v_i + \beta_0 e^{-\gamma r_{ij}^2} (v_j - v_i) + \alpha \left(\text{rand} - \frac{1}{2} \right), \quad (9)$$

where second term is the attraction of firefly j to firefly i , and the third term is the randomization of the movement. γ is a absorption coefficient, r_{ij} is the Cartesian distance between two fireflies i and j . β_0 is the attractiveness at $r_{ij} = 0$, α is a randomization parameter, rand is a random number generator uniformly distributed in $[0, 1]$. As recommended by [46], $\gamma = 1$, $\beta_0 = 1$, and $\alpha \in [0, 1]$ are used in this study. Besides, α is often replaced by a αS_k where the scaling parameters S_k is determined by the actual scales of the problem.

3.4. Refinement with Pattern Search. In the proposed FA-MA, pattern search is employed to conduct exploitation of the parameters solution space. Pattern search (PS), a simple effective optimization technique, has already been successfully used in parameters optimization in previous studies [50]. By examining the neighborhood of the current solution, pattern search is very effective to exploit the local regions. In addition, its convergence to local minima for constrained problems as well as unconstrained problems has been proven in [51]. Thus, it is deserved to enhance the local exploitation of the FA in

proposed memetic algorithm. In some sense, the main objective of PS is to conduct individual learning by exploiting small local regions effectively in relatively short periods of time.

Pattern search investigates nearest neighborhood of the current solution and tries to find a better move. If all neighbors fail to produce an improvement, then the search step is reduced. This search stops until the search step gets sufficiently small, ensuring the convergence to a local minimum. The pattern search is based on a pattern P_k that defines the neighborhood of current solution. A well often used pattern is five-point unit-size rood pattern which can be represented by the generating matrix P_k in (10):

$$P_k = \begin{bmatrix} 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 \end{bmatrix}. \quad (10)$$

The procedure of pattern search is outlined in Algorithm 1. Δ_0 denotes the default search step of PS, Δ is a search step, p_k is a column of P_k , and Ω denotes the neighborhood of the current solution. The termination conditions are the maximum iteration is met or the search step gets a predefined small value. To balance the amount of computational budget allocated for exploration versus exploitation, $\Delta_0/8$ is experimentally selected as the minimum search step.

3.5. Description of the Proposed FA-MA. Based on the population initialization, firefly algorithm based exploration, and pattern search based individual learning, a FA-MA is illustrated in Algorithm 2.

It can be seen that FA-MA not only applies the FA to effectively perform exploration for promising solution in the whole search space but also employs pattern search to perform exploitation for individual learning in local spaces. To guarantee an initial swarm with diversity, Latin hypercube sampling method is applied to generate a random sample set. In addition, it is important to balance the exploration and exploitation under limited computational budget in MA. Hence, in this study, each firefly undergoes local refinement with a specified probability $pl(x_k)$, and the selection probability is defined by a roulette wheel section scheme with linear scaling [24]:

$$pl(x_k) = \frac{f_{\max}(\mathbf{P}) - f(x_k)}{\sum_{y \in \mathbf{P}} (f_{\max}(\mathbf{P}) - f(y))}, \quad (11)$$

where f is a fitness function (i.e., MAPE in this study) and $f_{\max}(\mathbf{P})$ is the maximum fitness value among the current population \mathbf{P} . With this selection probability, a firefly with better fitness value gains more chance to be selected for exploitation.

Since both exploration and exploitation are stressed and balanced, it is expected to have good ability for improving the load forecasting with SVR. In the next section, we will investigate the performance of the proposed FA-MA.

4. Experimental Results

4.1. Experimental Setup. To verify the electricity load forecasting performance of the proposed SVR-MA model, two

```

Begin
  Initialize: predefine the default search step  $\Delta_0$ , initialize current solution  $p_0$ ;
   $\Delta = \Delta_0$ ;
  While (Termination conditions are not satisfied)
     $\Omega = \{p_0 + \Delta * p_k \mid \text{for each column } p_k \text{ in } P_k\}$ ;
    Evaluate the nearest neighbors in  $\Omega$ ;
    If (there are improvements in the  $\Omega$ )
      Update the current solution to the best neighbor in  $\Omega$ ;  $\Delta = \Delta_0$ ;
    Else
      Decrease the search step  $\Delta = \Delta_0/2$ ;
    End If
  End While
End
    
```

ALGORITHM 1: Pseudocode of pattern search for individual learning.

```

Begin
  Initialize the parameters in MA, such as, swarm size  $n$ , scaling parameters  $S_k$ , the maximum iteration number, and default search step  $\Delta$  in PS;
  Generate the initial population of fireflies  $\{x_i, i = 1, 2, \dots, n\}$ , using Latin Hypercube Sampling (LHS) method;
  Evaluate the fitness of each firefly in the population: use each firefly to train the SVR based forecasting model and calculate the fitness (MAPE);
  While (Termination conditions are not satisfied)
    For  $i = 1: n$  in all  $n$  fireflies
      For  $j = 1: i$  in all  $n$  fireflies
        If ( $I_j > I_i$ ) firefly  $i$  move to a new position according to (11); End If
        Evaluate the fitness of each firefly and update its light intensity;
      End For
    End For
    Rank the fireflies and find the current best;
    %%% Conduct individual learning with pattern search with a probability pl
    For  $k = 1: n$  in all  $n$  fireflies
      If ( $\text{rand} > \text{pl}(x_k)$ )
        Conduct individual learning for firefly  $x_k$  with pattern search (as shown in Algorithm 1);
      End If
    End For
  End While
  Output the best firefly as the final optimal solution
End
    
```

ALGORITHM 2: Pseudocode of proposed firefly algorithm based memetic algorithm.

real-life cases are considered in this study. The first one is the hourly observations from Pennsylvania-New Jersey-Maryland (PJM) power system, which is a well-established electricity market in U.S. The data consists of 18 months of hourly observations, from January 1, 2010, to 31 June, 2011 (data are available from PJM Interconnection, <http://www.pjm.com>). The series consists of 13104 hourly observations. The second one is the monthly electric load of Northeast China which has been investigated in the existing literature [34]. This data consists of 64 monthly observations with the date from January 2004 to April 2009.

As a preprocessing stage, several missing load values are filled in by the average of the neighboring values. By adopting linear transformation (in (12)), the series are linearly scaled to the range [0, 1]. The main advantage of scaling is to avoid

attributes in greater numeric ranges dominating those in smaller numeric ranges. Another advantage is to prevent numerical difficulties during the calculation [41]

$$x' = \frac{x - \min_A}{\max_A - \min_A}, \tag{12}$$

where x is an original value of attribute A , x' is the scaled value, \min_A is the minimum of attribute A , and \max_A is the maximum of attribute A . It should be noted that the forecasting value will be rescaled back following the reverse of the linear transformation and the forecasting performance is calculated based on the original scale of the data.

4.2. Performance Measures. To assess the forecasting performance of a model, many accuracy measures can be used [52].

TABLE 1: Performance metrics and their formulas in regression problems.

Metrics	Formula
MAPE	$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \left \frac{y_{t+i} - \hat{y}_{t+i}}{y_{t+i}} \right \times 100$
MASE	$\text{MASE} = \frac{1}{N} \sum_{i=1}^N \left \frac{y_{t+i} - \hat{y}_{t+i}}{(1/(t-1)) \sum_{j=2}^t y_j - y_{j-1} } \right $
DS	$\text{DS} = \frac{\sum_{i=2}^N d_i}{N-1} \times 100$ $d_i = \begin{cases} 1, & \text{if } (y_{t+i} - y_{t+i-1})(y_{t+i} - \hat{y}_{t+i-1}) \geq 0 \\ 0, & \text{otherwise} \end{cases}$

N is the number of forecasting periods, y_{t+i} is the actual value at period $t+i$, \hat{y}_{t+i} is the forecasting value at period $t+i$, and \bar{y} is the mean of all values. In this study, the day-ahead (24 hours) short-term load is forecasted recursively, so the number of forecasting periods N equals 24.

However, it is hard to say whether one accuracy measure is better or worse than the other [53]. Besides, different metrics may evaluate the quality of the forecasting performance from different perspectives. In this study, three accuracies are selected to assess the prediction performance, they are mean absolute percentage error (MAPE), mean absolute scaled error (MASE), and directional symmetry (DS). The definitions of them can be found in Table 1. MAPE is one often used metric, which measure the percentage error between the actual and predicted values. The smaller the values of MAPE, the closer the predicted values to the actual values. MASE is a scaled error which is scaled by a naïve forecast model. MASE is less than one if the forecast is better than the naïve method, and the smaller the values of MASE, the better the naïve method. It is highly recommended in recent study as it is less sensitive to outliers and easy to be interpreted [52]. DS provides an indication of the accuracy of the predicted direction and the large value suggests a better predictor.

Furthermore, a nonparametric Wilcoxon's signed-rank test [54] is performed to determine if there is significant difference between the two approaches based on the prediction error of the testing data sets. This test performs a two sample rank test for the difference between two population medians. Since the population distributions of the performance measures are unknown, a nonparametric test is suggested for the performance comparison of the two models [55].

4.3. Results and Discussions. In the first case, the 24 step-ahead electricity load is predicted directly. The data are divided into three parts: training set, validation set, and testing set. The periods and number of observations of each set are shown in Table 2. As mentioned in Section 3.2, the training set and validation set are used to determine the optimal parameters, and then the forecasting model is established in the integrated training set (training set and validation set). At last, the testing set is used to assess the out-of-sample forecasting performance of the proposed model with optimal

TABLE 2: Training, validation, and testing set for the first sample case.

Data sets	Period	No. of observation
Training set	1/1/2010–12/31/2010	365 * 24
Validation set	1/1/2011–3/31/2011	90 * 24
Testing set	4/1/2011–6/30/2011	91 * 24

TABLE 3: MAPE (%) of SVR model with different parameter determination methods.

Period	FA-MA	FA	GA	PSO	SA
April	1.24	1.51	1.67	1.73	1.95
May	1.34	1.53	1.77	1.83	2.00
June	1.48	1.78	1.71	2.00	2.14
ALL	1.35	1.61	1.72	1.85	2.03

TABLE 4: MASE of SVR model with different parameter determination methods.

Period	FA-MA	FA	GA	PSO	SA
April	0.33	0.38	0.44	0.47	0.51
May	0.37	0.43	0.50	0.53	0.51
June	0.41	0.48	0.48	0.50	0.56
ALL	0.37	0.42	0.47	0.50	0.53

parameters obtained by memetic algorithms. Considering the short-run trend, daily and weekly periodicity characteristics of hourly load, the hourly load values of the last one day, and the similar hours in the previous 30 days are selected as the input variables set of the forecasting model. Then, the input variables are selected from the variables set by a filter method which maximizes the mutual information using forward-backward selection strategy [56].

To verify the improvement of forecasting accuracy with our proposed memetic algorithm in SVR based forecasting model, four well-known evolutionary algorithms (EAs) including genetic algorithm (GA), particle swarm optimization (PSO), simulated annealing (SA), and firefly algorithm (FA) are selected to determine the parameters (C , δ , and ϵ) in SVR based load forecasting model. The experiments are implemented in MATLAB 2012a using computer with Intel Core 2 Duo CPU T5750, 2.00 GHz, and 2 G RAM. The parameters' search space in SVR is defined as an exponentially growing space: $\log_2 C \in [-6, 6]$, $\log_2 \gamma \in [-6, 6]$, and $\log_2 \epsilon \in [-6, 6]$. The parameters in each EA are controlled based on initial experiments. More specifically, the population size of each method is set as 30, and the stopping criterions of each method are set as follows: the number of iterations reached 150 or there is no improvement in the fitness for 50 consecutive iterations. The scaling parameters S_k in firefly algorithm and FA-MA are both set as 1, which is 1/6 percentage of the maximum of search space.

The forecasting results of different EAs based SVR forecasting model in each separate month and the whole testing period are illustrated in Tables 3, 4, and 5. For the purpose of reducing statistical errors, the results in Tables 3–5 are average results of 30 independently trial runs. From Tables 3–5, several observations can be drawn. Firstly, compared with GA,

TABLE 5: DS (%) of SVR model with different parameter determination methods.

Period	FA-MA	FA	GA	PSO	SA
April	96.49	94.74	93.78	93.15	92.77
May	95.73	94.32	93.69	93.08	92.54
June	95.30	93.30	93.50	92.56	91.50
ALL	95.84	94.12	93.66	92.93	92.27

TABLE 6: Time consuming of SVR model with different parameter determination methods.

	FA-MA	FA	GA	PSO	SA
CPU time (min)	27.3	20.7	25.6	21.9	22.5

PSO, and SA, the FA based forecasting model can obtain the best performance in most of the periods for each metric, which imply the superior ability of determining the parameters in SVR and thus improving the forecasting performance. Secondly, by using the proposed MA to determine the parameters in SVR forecasting model, the forecasting results outperform the FA based forecasting model. The superior performance against FA can be contributed to the integration of pattern search for finely exploitation and the balance between exploration and exploitation in proposed MA. Thirdly, the proposed MA has the lowest MAPE and MASE, with the largest DS, which confirms the superiority of MA in improving the forecasting accuracy by enhancing the parameters determination process of SVR forecasting model.

Table 6 reports the time consumption of each evolutionary algorithm in selecting the optimal parameters in SVR forecasting model. From Table 6, we can see that the proposed FA-MA is a little more time consuming than that of the other four methods, which is mainly due to the finely exploitation with pattern search. However, in the real-world applications, this computational time is acceptable within a day-ahead decision making framework, and considering the forecasting performance of FA-MA in improving the forecasting, it can be used as an alternative method to improve the forecasting accuracy when the support vector regression (SVR) is used.

Furthermore, three well-known forecasting model, including radial basis function neural network (RBFNN), MLP neural network trained by LM (Levenberg-Marquardt), and autoregression integrated moving average (ARIMA), are selected to compare the day-ahead forecasting performance of the proposed FA-MA based SVR model. For the sake of fair comparison, the above two neural network based models have the same process in data preprocess, input selection, and parameters tuning to our proposed FA-MA based SVR model. While for ARIMA, the *forecast* package [57] in *R* is used to forecast the load. Tables 7, 8, and 9 show the comparison of average results of three separated months and the whole period. From Tables 7–9, it can be observed that the ARIMA is the worse one in each month and the whole period, which is mainly due to the linearity assumption. Besides, the proposed FA-MA based SVR model outperforms all the other forecasting models in terms of MAPE, MASE, and DS. Moreover, to verify the significance of accuracy improvement of proposed FA-MA based SVR model, a nonparametric

TABLE 7: MAPE of four forecasting models.

Period	FA-MA	RBF	MLP-LM	ARIMA
April	1.24	2.37	2.35	4.91
May	1.34	2.38	2.39	5.00
June	1.48	2.45	2.51	5.20
ALL	1.35	2.40	2.42	5.04

TABLE 8: MASE of four forecasting models.

Period	FA-MA	RBF	MLP-LM	ARIMA
April	0.33	0.60	0.59	0.72
May	0.37	0.60	0.63	0.75
June	0.41	0.64	0.66	0.78
ALL	0.37	0.61	0.63	0.75

TABLE 9: DS of four forecasting models.

Period	FA-MA	RBF	MLP-LM	ARIMA
April	96.49	90.12	90.34	85.51
May	95.73	89.15	89.01	85.31
June	95.30	89.78	89.19	84.40
ALL	95.84	89.68	89.51	85.07

Wilcoxon’s signed-rank test is used to test the significant different of FA-MA based SVR with three other models. The significant test shows that our proposed model is statistically superior to others at the 0.05 significance levels. In addition, to give a graphical view about the forecasting performance of the proposed model, the curves of real values, forecast values, and forecast errors are shown in Figure 1. It is obvious that the forecast curve accurately predicts the real values and only minor errors are obtained. The figure further illustrates the effectiveness of our proposed forecasting model.

In addition, to further verify the proposed performance of FA-MA against the existing hybrid methods, the second case, obtained from previous studies [34], is applied here. Similar to previous studies [34], the last 7 months are predicted. The seasonal mechanism effects stated in [34] are also taken into consideration. Table 10 shows the actual values and the forecasting load obtained by different forecasting models. The TF- ϵ -SVR-SA reports the results from [17]; the SVR model is optimized by SA. For CGASA and S-CGASA, the forecasts were generated by the SVR model with or without the seasonal mechanism, respectively. The parameters in these two SVR models were tuned by a hybrid algorithm, namely, chaotic genetic algorithm-simulated annealing algorithm (CGASA). In the last two columns, the SVR models were optimized by our proposed FA-MA algorithm. The only difference between them is that the S-FA-MA takes the seasonal mechanism effects into account. As illustrated in Table 10, the proposed FA-MA gains smaller MAPE, MASE than TF- ϵ -SVR-SA and CGASA. However, the results of SVRFA-MA is worse than those of S-CGASA, which is mainly due to an involvement of a seasonal mechanism in S-CGASA. Similar to S-CGASA, S-FA-MA also makes full use of the seasonal effects; it generates superior performance to S-CGASA in terms of MAPE, MASE and has competitive performance

TABLE 10: Comparison with existing hybrid algorithms.

Period	Actual	TF- ϵ -SVR-SA	CGASA	S-CGASA	FA-MA	S-FA-MA
Oct.08	181.07	184.5035	177.3	175.6385	175.9047	178.2513
Nov.08	180.56	190.3608	177.4428	185.21	184.5484	184.2637
Dec.08	189.03	202.9795	177.5848	189.907	195.4447	188.9679
Jan.09	182.07	195.7532	177.7263	181.9693	185.5828	181.7957
Feb.09	167.35	167.5795	177.8673	163.2805	161.4537	161.9352
Mar.09	189.3	185.9358	178.0078	182.1747	184.854	181.9227
Apr.09	175.84	180.1648	178.6806	177.6289	177.2037	176.1128
MAPE (%)		3.799	3.731	1.901	2.433	1.583
MASE		0.576	0.554	0.237	0.326	0.217
DA (%)		83.333	33.333	83.333	83.333	83.333

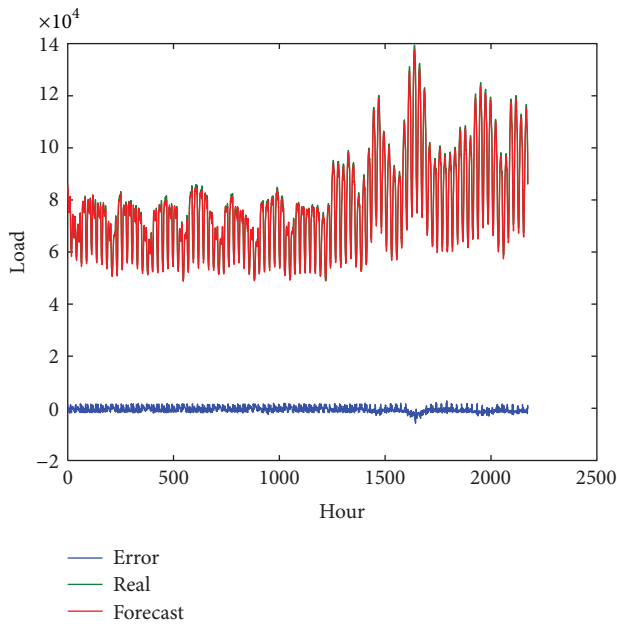


FIGURE 1: Curves of real values, forecast values, and errors of proposed FA-MA based SVR model.

with S-CGASA in terms of DA. Thus, in this case, we can conclude that the proposed FA-MA is superior to the existing hybrid SA, CGASA in improving the forecasting accuracy of SVR models. The outstanding forecasting performance of our proposed FA-MA against the existing hybrid algorithm (i.e., CGASA) is caused by the following reasons. Firstly, based on the framework of memetic algorithm, both global exploration and local exploitation are enhanced in the proposed FA-MA, which not only can avoid the premature convergence but also ensure searching capability. Secondly, a roulette wheel section scheme is applied in the proposed FA-MA to select the individuals to be refined, which generates a good balance to the exploration and exploitation.

5. Conclusions

Electricity load forecasting is an important issue to operate the power system reliably and economically. In this study, to improve forecasting accuracy of electricity load forecasting using support vector regression (SVR), a firefly algorithm

(FA) based memetic algorithm (FA-MA) was presented. In the proposed FA-MA, FA was employed to explore the search space and detect the potential regions, while pattern search (PS) was used to conduct the individual learning to improve the exploitation ability of FA. With the proposed FA-MA used to determine the parameters of SVR, a novel forecasting model, FA-MA based SVR, was presented to forecast the electricity load with two real cases. In the first case, four evolutionary algorithms (FA, GA, PSO, and SA) based SVR forecasting models and three well-known models (RBFNN, MLP-LM, and ARIMA) were selected to compare the forecasting performance. Computational results show that the proposed FA-MA could effectively improve the forecasting accuracy of SVR compared with some other evolutionary algorithms based SVR. Meanwhile, the FA-MA based SVR forecasting model could outperform the selected counterparts significantly. In the second case, comparison results show that the proposed FA-MA is superior to the existing hybrid algorithm in the literature.

However, in this study, only the historical load values are taken into consideration to forecast the electricity load, and some exogenous variables (i.e., temperature, humidity) are also very important to improve the forecasting accuracy. Other topics include more extensive comparison with other models, developing more efficient memetic algorithms and seasonal adjustment. Extensive experimental studies in other forecasting problems and benchmark functions can be investigated. Future work will be on the research of the above cases.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

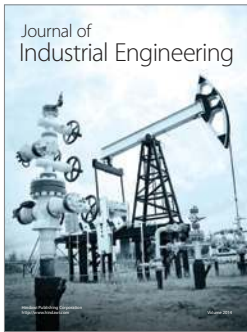
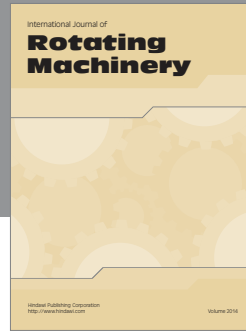
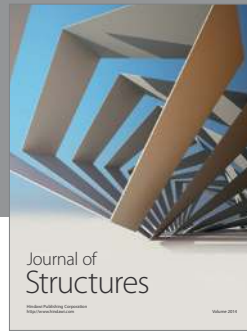
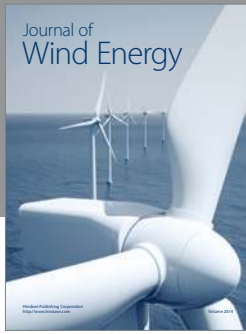
Acknowledgments

This work was supported by the Natural Science Foundation of China under Project no. 70771042, the Fundamental Research Funds for the Central Universities (2012QN208-HUST), the Ministry of Education in China (MOE) Project of Humanities and Social Science (Project no. 13YJA630002), and a Grant from the Modern Information Management Research Center at Huazhong University of Science and Technology (2013WZ005 2012WJD002).

References

- [1] S. Fan and L. Chen, "Short-term load forecasting based on an adaptive hybrid method," *IEEE Transactions on Power Systems*, vol. 21, no. 1, pp. 392–401, 2006.
- [2] D. W. Bunn and E. D. Farmer, *Comparative Models for Electrical Load Forecasting*, John Wiley & Sons, New York, NY, USA, 1985.
- [3] S. Saab, E. Badr, and G. Nasr, "Univariate modeling and forecasting of energy consumption: the case of electricity in Lebanon," *Energy*, vol. 26, no. 1, pp. 1–14, 2001.
- [4] A. P. Douglas, A. M. Breipohl, F. N. Lee, and R. Adapa, "The impacts of temperature forecast uncertainty on bayesian load forecasting," *IEEE Transactions on Power Systems*, vol. 13, no. 4, pp. 1507–1513, 1998.
- [5] J. H. Park, Y. M. Park, and K. Y. Lee, "Composite modeling for adaptive short-term load forecasting," *IEEE Transactions on Power Systems*, vol. 6, no. 2, pp. 450–457, 1991.
- [6] A. Goia, C. May, and G. Fusai, "Functional clustering and linear regression for peak load forecasting," *International Journal of Forecasting*, vol. 26, no. 4, pp. 700–711, 2010.
- [7] H. A. Amarawickrama and L. C. Hunt, "Electricity demand for Sri Lanka: a time series analysis," *Energy*, vol. 33, no. 5, pp. 724–739, 2008.
- [8] 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.
- [9] P. Mandal, T. Senjyu, N. Urasaki, and T. Funabashi, "A neural network based several-hour-ahead electric load forecasting using similar days approach," *International Journal of Electrical Power and Energy Systems*, vol. 28, no. 6, pp. 367–373, 2006.
- [10] Z. A. Bashir and M. E. El-Hawary, "Applying wavelets to short-term load forecasting using PSO-based neural networks," *IEEE Transactions on Power Systems*, vol. 24, no. 1, pp. 20–27, 2009.
- [11] P. Lauret, E. Fock, R. N. Randrianarivony, and J.-F. Manicom-Ramsamy, "Bayesian neural network approach to short time load forecasting," *Energy Conversion and Management*, vol. 49, no. 5, pp. 1156–1166, 2008.
- [12] N. Amjady and F. Keynia, "A new neural network approach to short term load forecasting of electrical power systems," *Energies*, vol. 4, no. 3, pp. 488–503, 2011.
- [13] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer, New York, NY, USA, 1995.
- [14] N. Sapankevych and R. Sankar, "Time series prediction using support vector machines: a survey," *IEEE Computational Intelligence Magazine*, vol. 4, no. 2, pp. 24–38, 2009.
- [15] Y. Bao, T. Xiong, and Z. Hu, "Forecasting air passenger traffic by support vector machines with ensemble empirical mode decomposition and slope-based method," *Discrete Dynamics in Nature and Society*, vol. 2012, Article ID 431512, 12 pages, 2012.
- [16] E. E. Elattar, J. Goulermas, and Q. H. Wu, "Electric load forecasting based on locally weighted support vector regression," *IEEE Transactions on Systems, Man and Cybernetics C*, vol. 40, no. 4, pp. 438–447, 2010.
- [17] J. Wang, W. Zhu, W. Zhang, and D. Sun, "A trend fixed on firstly and seasonal adjustment model combined with the ϵ -SVR for short-term forecasting of electricity demand," *Energy Policy*, vol. 37, no. 11, pp. 4901–4909, 2009.
- [18] P.-F. Pai and W.-C. Hong, "Support vector machines with simulated annealing algorithms in electricity load forecasting," *Energy Conversion and Management*, vol. 46, no. 17, pp. 2669–2688, 2005.
- [19] T. Xiong, Y. Bao, and Z. Hu, "Multiple-output support vector regression with a firefly algorithm for interval-valued stock price index forecasting," *Knowledge-Based Systems*, vol. 55, pp. 87–100, 2014.
- [20] Y. Bao, T. Xiong, and Z. Hu, "Multi-step-ahead time series prediction using multiple-output support vector regression," *Neurocomputing*, 2013.
- [21] P. Duan, K. G. Xie, T. T. Guo, and X. G. Huang, "Short-term load forecasting for electric power systems using the PSO-SVR and FCM clustering techniques," *Energies*, vol. 4, no. 1, pp. 173–184, 2011.
- [22] O. Chapelle, V. Vapnik, O. Bousquet, and S. Mukherjee, "Choosing multiple parameters for support vector machines," *Machine Learning*, vol. 46, no. 1–3, pp. 131–159, 2002.
- [23] Y. K. Bao and Z. T. Liu, "A fast grid search method in support vector regression forecasting time series," in *Intelligent Data Engineering and Automated Learning—IDEAL 2006*, E. Corchado, H. Yin, V. Botti, and C. Fyfe, Eds., vol. 4224, pp. 504–511, Springer, Berlin, Germany, 2006.
- [24] Y. Bao, Z. Hu, and T. Xiong, "A PSO and pattern search based memetic algorithm for SVMs parameters optimization," *Neurocomputing*, vol. 117, pp. 98–106, 2013.
- [25] T. A. F. Gomes, R. B. C. Prudêncio, C. Soares, A. L. D. Rossi, and A. Carvalho, "Combining meta-learning and search techniques to select parameters for support vector machines," *Neurocomputing*, vol. 75, no. 1, pp. 3–13, 2012.
- [26] J. Wang, L. Li, D. Niu, and Z. Tan, "An annual load forecasting model based on support vector regression with differential evolution algorithm," *Applied Energy*, vol. 94, pp. 65–70, 2012.
- [27] H. Li, S. Guo, H. Zhao, C. Su, and B. Wang, "Annual electric load forecasting by a least squares support vector machine with a fruit fly optimization algorithm," *Energies*, vol. 5, no. 11, pp. 4430–4445, 2012.
- [28] P.-F. Pai and W.-C. Hong, "Forecasting regional electricity load based on recurrent support vector machines with genetic algorithms," *Electric Power Systems Research*, vol. 74, no. 3, pp. 417–425, 2005.
- [29] W. C. Hong, Y. Dong, W. Y. Zhang, L. Y. Chen, and B. K. Panigrahi, "Cyclic electric load forecasting by seasonal SVR with chaotic genetic algorithm," *International Journal of Electrical Power & Energy Systems*, vol. 44, no. 1, pp. 604–614, 2013.
- [30] W.-C. Hong, "Chaotic particle swarm optimization algorithm in a support vector regression electric load forecasting model," *Energy Conversion and Management*, vol. 50, no. 1, pp. 105–117, 2009.
- [31] W.-C. Hong, "Electric load forecasting by seasonal recurrent SVR (support vector regression) with chaotic artificial bee colony algorithm," *Energy*, vol. 36, no. 9, pp. 5568–5578, 2011.
- [32] W.-C. Hong, "Electric load forecasting by support vector model," *Applied Mathematical Modelling*, vol. 33, no. 5, pp. 2444–2454, 2009.
- [33] W.-C. Hong, Y. Dong, C.-Y. Lai, L.-Y. Chen, and S.-Y. Wei, "SVR with hybrid chaotic immune algorithm for seasonal load demand forecasting," *Energies*, vol. 4, no. 6, pp. 960–977, 2011.
- [34] W. Y. Zhang, W. C. Hong, Y. C. Dong, G. Tsai, J. T. Sung, and G. F. Fan, "Application of SVR with chaotic GASA algorithm in cyclic electric load forecasting," *Energy*, vol. 45, no. 1, pp. 850–858, 2012.
- [35] S. Areibi and Z. Yang, "Effective memetic algorithms for VLSI design = genetic algorithms + local search + multi-level clustering," *Evolutionary Computation*, vol. 12, no. 3, pp. 327–353, 2004.

- [36] Q. H. Nguyen, Y.-S. Ong, and M. H. Lim, "A probabilistic memetic framework," *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 3, pp. 604–623, 2009.
- [37] A. Elhossini, S. Areibi, and R. Dony, "Strength pareto particle swarm optimization and hybrid EA-PSO for multi-objective optimization," *Evolutionary Computation*, vol. 18, no. 1, pp. 127–156, 2010.
- [38] Y.-S. Ong, M. H. Lim, and X. Chen, "Memetic computation—past, present & future [research frontier]," *IEEE Computational Intelligence Magazine*, vol. 5, no. 2, pp. 24–31, 2010.
- [39] J. Tang, M. H. Lim, and Y. S. Ong, "Diversity-adaptive parallel memetic algorithm for solving large scale combinatorial optimization problems," *Soft Computing*, vol. 11, no. 9, pp. 873–888, 2007.
- [40] V. N. Vapnik, *Statistical Learning Theory*, John Wiley & Sons, New York, NY, USA, 1998.
- [41] C. W. Hsu, C. C. Chang, and C. J. Lin, "A practical guide to support vector classification," Department of Computer Science, National Taiwan University, 2003.
- [42] F. E. H. Tay and L. Cao, "Application of support vector machines in financial time series forecasting," *Omega*, vol. 29, no. 4, pp. 309–317, 2001.
- [43] P. Moscato and M. G. Norman, "A "memetic" approach for the traveling salesman problem implementation of a computational ecology for combinatorial optimization on message-passing systems," in *Proceedings of the International Conference on Parallel Computing and Transputer Applications*, 1992.
- [44] F. Neri and C. Cotta, "Memetic algorithms and memetic computing optimization: a literature review," *Swarm and Evolutionary Computation*, vol. 2, pp. 1–14, 2012.
- [45] X. Chen, Y.-S. Ong, M.-H. Lim, and K. C. Tan, "A multi-facet survey on memetic computation," *IEEE Transactions on Evolutionary Computation*, vol. 15, no. 5, pp. 591–607, 2011.
- [46] X. S. Yang, "Firefly algorithms for multimodal optimization," in *Stochastic Algorithms: Foundations and Applications*, O. Watanabe and T. Zeugmann, Eds., vol. 5792, pp. 169–178, Springer, Berlin, Germany, 2009.
- [47] X.-S. Yang, S. S. S. Hosseini, and A. H. Gandomi, "Firefly algorithm for solving non-convex economic dispatch problems with valve loading effect," *Applied Soft Computing Journal*, vol. 12, no. 3, pp. 1180–1186, 2012.
- [48] K. Chandrasekaran and S. P. Simon, "Optimal deviation based firefly algorithm tuned fuzzy design for multi-objective UCP," *IEEE Transactions on Power Systems*, vol. 28, no. 1, pp. 460–471, 2013.
- [49] M. D. Mckay, R. J. Beckman, and W. J. Conover, "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code," *Technometrics*, vol. 42, no. 1, pp. 55–61, 2000.
- [50] M. Momma and K. P. Bennett, "A pattern search method for model selection of support vector regression," in *Proceedings of the 2nd SIAM International Conference on Data Mining*, pp. 261–274, 2002.
- [51] E. D. Dolan, R. M. Lewis, and V. Torczon, "On the local convergence of pattern search," *SIAM Journal on Optimization*, vol. 14, no. 2, pp. 567–583, 2003.
- [52] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *International Journal of Forecasting*, vol. 22, no. 4, pp. 679–688, 2006.
- [53] J. S. Armstrong and F. Collopy, "Error measures for generalizing about forecasting methods: empirical comparisons," *International Journal of Forecasting*, vol. 8, no. 1, pp. 69–80, 1992.
- [54] F. X. Diebold and R. Mariano, "Comparing predictive accuracy," *Journal of Business and Economic Statistics*, vol. 13, pp. 253–263, 1995.
- [55] W. Conover, *Practical Nonparametric Statistics*, John Wiley & Sons, New York, NY, USA, 2nd edition, 1980.
- [56] A. Sorjamaa, J. Hao, N. Reyhani, Y. Ji, and A. Lendasse, "Methodology for long-term prediction of time series," *Neurocomputing*, vol. 70, no. 16–18, pp. 2861–2869, 2007.
- [57] R. J. Hyndman and Y. Khandakar, "Automatic time series forecasting: the forecast package for R," *Journal of Statistical Software*, vol. 27, no. 3, pp. 1–22, 2008.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

