A New Spinning Reserve Requirement Prediction with Hybrid Model

Mohammad Ghiasi\textsuperscript{a,b}, Esmaeil Ahmadinia\textsuperscript{a,c}, Milad Janghorban Lariche\textsuperscript{d}, Houman Zarrabi\textsuperscript{e} and Rolando Simoes\textsuperscript{f}

\textsuperscript{a}Power Control Center (PCC), Tehran Metro, Tehran Urban and Suburban Railway Operation Company (TUSRC), Tehran, Iran; \textsuperscript{b}Department of Electrical Engineering, Sciences and Research Branch, Islamic Azad University, Tehran, Iran; \textsuperscript{c}Department of Engineering, Payame Noor University, Tehran, Iran; \textsuperscript{d}Abadan School of Medical Sciences, Abadan, Iran; \textsuperscript{e}ICT Research Institute, Tehran, Iran; \textsuperscript{f}Department of Electrical Engineering, Northeastern University, Boston, MA, USA

ABSTRACT
Ancillary services is used to refer to a variety of operations beyond generation and transmission which are requested to maintain grid stability, security and reliability of power system. These services generally consist, frequency control, Spinning Reserves (SR) and operating reserves. Accordingly, an accurate day ahead forecast of SR requirement helps the Independent System Operator to manage a reliable and economic operation of the power system. This prediction model needs strong and accurate method to tackle the complexity, non-stationary and volatility of this signal. Hence, a new hybrid forecasting model is proposed in this paper, to solve the SR requirement. The proposed structure consists of three stage Neural Network (NN) based forecast engine with different learning algorithms. Also, the input signal of this forecast engine is filtered by a new feature selection model to find the high relevancy and low redundancy of features. The proposed strategy is implemented and tested on real data of Pennsylvania–New Jersey–Maryland (PJM) through the comparison with other techniques. Obtained numerical results demonstrate the validity of proposed method.

ARTICLE HISTORY
Received 23 November 2017
Accepted 1 April 2018

KEYWORDS
Spinning reserve forecast; feature selection; forecast engine; hybrid neural network

1. Introduction
An Independent System Operator (ISO) is an establishment organized at the direction or aim of the Federal Energy Regulatory Commission (FERC). In the areas where an ISO is organized, it coordinates, controls and monitors the operation of the electrical power system. Furthermore, FERC defines the ancillary services as necessary services to support the transmission of electric power from seller to consumers given the commitments of control areas and transmitting utilities within those control areas to maintain reliable operations of the interconnected transmission system [1].

Ancillary services are the proficiency services and functions supplied by the electric grid that facilitate and support the continuous flow of electricity so that supply will uninterruptedly meet demand. Accordingly, SR is an important ancillary service. Scheduling sufficient SR helps power systems to tackle unanticipated generator outages and major load forecasting errors without load shedding [2].

The certain frameworks for market clearing, usually used in many electricity markets, suppose that the needed SR capacity is predetermined, for example, based on the largest unit criterion or as a percentage of the load demand [3]. Contrary to their simplicity, the deterministic
frameworks may lead to uneconomical and/or insecure operation of the power system, since they ignore the uncertainty sources of the power system such as occurrence of generator and branch contingencies and load prediction error. A more detailed discussion about this issue is presented in [4]. To remedy the variety of uncertainties in relevant decision-making, some recent research works have proposed stochastic frameworks to consider the uncertainty sources of power system in defining the reserve capacity requirement of the system [1,5–9].

Actually, the accuracy of prediction model is one of the most important factors for solving this problem. In this paper, a new prediction method proposed to calculate SR capacity requirement based on uncertainty sources in power system. In [3], the SR capacity requirement is predicted based on an Adaptive Wavelet Neural Network. In [10], a probabilistic model of security-constrained unit commitment to minimize the cost of energy, spinning reserve and possible loss of load is proposed. Probabilistic methodology for estimating spinning reserve requirement in microgrids is presented in [11] which is based on multi-step method of wavelet and Mixed Integer Linear Programming (MILP).

Combination of Multi-Layer Perceptron (MLP) neural network based on Levenberg–Marquadt (LM) learning algorithm and Real Coded Genetic Algorithm (RCGA) training mechanism is proposed in [12], for solving the mentioned SR problem. However, the mentioned models couldn’t provide good accuracy of prediction model.

In this paper, a hybrid forecast engine based on three-stage NN is proposed. All three stages are trained with different training mechanism. Also, the weights of this hybrid forecast engine has been optimized by PSOTVAC/BFA algorithm which is introduced recently as strong optimization method [13]. Also, before predicting the input signal, a new fuzzy clustering method is implemented for filtering the features. In this model, the proposed improved fuzzy system filtered the features based on high relevancy and low redundancy.

The main contributions of this paper can be summarized as follows:

1. A novel and efficient forecast method for SR requirement is proposed. The forecast method is composed of a hybrid optimization algorithm and modified hybrid neural network (MHNN). It has high learning capability and can avoid from overfitting problem and trapping in local minima and dead bands.

2. A new feature selection model has been presented based on hybrid model for filtering the input features.

The remaining parts of the paper are organized as follows. In Section 2, the proposed feature selection is presented. Section 3, presents the forecast model. Obtained numerical results from the proposed strategy are presented in Section 4. Section 5 concludes the paper.

2. The Proposed Feature Selection

2.1. Input Set Structure

First of all, it is very important to recognize the construction and behavior of input signal (SR) in prediction process. Accordingly, a bulk candidate set of input features is structured considering the statistical behavior of the SR signal. Considering such a more set of signal enables us to employ the maximum information value of the available data such that no likely informative and benefit feature is missing. But, the large number of these features cannot be useful directly for using in forecast engine. Hence, these features should be filtered by appropriate filter named feature selection.

As presented in [12] the SR time series has daily periodicity specifications such that its pattern almost repeats every 24 hours. Furthermore, it is clear that high relevancy between the load demand and SR requirement is exist. Consequently, the following set of candidate inputs can be considered for SR requirement forecast:

\[
\{ SR_{t-1}, SR_{t-2}, \ldots, SR_{(t-200)}, L_{(t-1)}, L_{(t-2)}, \ldots, L_{(t-200)} \} \tag{1}
\]

where the lagged value of SR requirement and the load demand is presented by SR and LD, respectively. So, the daily and weekly periodicity specifications of both signals included through the candidate set of inputs [14].

Finally, for using the presented signal directly for forecast engine based on filtering the features, a minimum set of the most informative features will be selected based on the proposed feature selection method as described in the following:

2.2. Hybrid Feature Selection

According to the special conditions in SR forecasting problem, occasionally we encounter with many attributes whereas some of them no longer have helpful data and just complex the situation. Hence, feature selection is one of the very critical aspect that has a highly regarded recommendation [15]. This section introduce, a new feature selection which have greater effect on risk and return and appropriate analysis of algorithms results a novel feature selecting method in two stages is established. In this problem we deal with so many features that are either useless or have low information value. So, dealing with these features
is time wasting without any good results. Feature selection methods are generally categorized into three main groups as: (1) filter methods, (2) wrapper methods and (3) hybrid methods [16]. In this paper, we applied a hybrid model based on combination of filter and function-based clustering model to extract a set of effective features as presented in Figure 1.

### 2.2.1. Filter Model

As mentioned in [15], seven algorithms have been defined as Filter methods of: Info Gain, Gain Ratio, Chi square, Relief-f, One R, consistency and CFS. Moreover, Symmetrical Uncertainty and SVM algorithm are also implemented for weighting the features [15]. In this manuscript, to compare the importance of each feature based on mentioned models, a comprehensive analysis was proposed on the features and eventually the weightings of features are presented.

### 2.2.2. Function-Based Clustering Method

After reaching to the features weights by various filter based algorithms we have ‘n’ feature with ‘m’ feature’s weight and then a model to determine the important features’ clustering is needed which is between these weighted attributes. Consequently, we develop the presented function-based clustering model in [16]. This method is based on hierarchical divisive clustering model which starts with one cluster consisting of all objects ($X_{n \times m}$).

---

**Figure 1.** Proposed feature selection.

**Figure 2.** The flowchart of proposed hybrid PSOTVAC/BFA technique.

**Figure 3.** The proposed forecast engine structure.
For the object $x_1, \ldots, x_n$, we define the vector of group membership of objects as $z = (z_1, \ldots, z_n)^T$, where $z \in Z$, and $Z$ is the space of sign vectors denoted as:

$$Z = \{ z = (z_1, \ldots, z_n)^T | z_i = \pm 1 \}$$

(2)

All objects that are depended with an entry of 1 in z are categorized into one group, whereas the others with an entry of -1 are categorized into the other group.

Then by using the model of multivariate analysis of variance denoted as follows:

$$x_i = \mu + z_i \gamma + \epsilon_i, \quad i = 1, 2, \ldots, n$$

(3)

where $\epsilon_i$ is the error vectors and assumed to be normally distributed with a zero mean and a common covariance matrix $V$, i.e. $N(0, V)$. Additionally, $\epsilon_i$ and $\epsilon_j$ ($i \neq j$) are assumed to be independent. Then by maximum similarity, the clustering problem is presented as a least squares optimization problem as;

$$\min_{a, \beta, z \in Z} \{ (z - \alpha 1 - X \beta)^T (z - \alpha 1 - X \beta) \}$$

(4)

In a same time, the unknown vector of cluster membership and the coefficients of the linear clustering function are approximated. The calculation of the clustering function-based model will be converted to that of sign analysis [16], and by solving the problem, two clusters is obtained. Then, one of these groups based on higher within-group dispersion matrix is further categorized into two dissimilar sub-groups. The procedure continues till some stopping criterion has been satisfied which is based on within-group dispersion and/or between-group dispersion matrices.

By this method we use the benefits of several filter methods and use these weighting attributes by function-based clustering model to make more accurate decision of effective feature.

### 3. The Proposed Forecasting Engine Model

This section introduced the proposed hybrid forecast engine to predict the SR requirements. Accordingly, we present the proposed forecast engine structure and then, the proposed optimization method will be introduced.

#### 3.1. The Proposed HNN Structure

It is clear that combination of different neural networks (NNs) can potentially improve their learning ability in a complicated process [17–20]. But, in these methods the input data shared among their building blocks which mean the mature sharing between these blocks. To tackle the mentioned problem, in this paper a new hybrid structure of forecast engine is presented which includes to three main stages. All of these stages consist of a MLP NN as a forecast engine. In this model since all NNs have the same structure, the weights can be directly used by the next stage and then it can growths the obtained knowledge of the previous one. Additionally, selecting a variety of MLP training algorithms of the NNs, the HNN benefits from wider learning capability. The three NNs of HNN have LM (Levenberg–Marquardt), SCG (Scaled Conjugate Gradient Backpropagation) and CGP (Conjugate Gradient Backpropagation with Polak-Ribiére updates) learning algorithms. Further discussions
3.2. Hybrid Intelligent Algorithm

3.2.1. Particle Swarm Optimization (PSO)

This paper composed the ability of a version particle swarm optimization (PSO) and bacteria foraging algorithm (BFA), to find the solution of the clustering model in proposed fuzzy controller. The PSO is one of the optimization techniques and a kind of evolutionary computation technique which is launched by the Aberhart Rasel [24]. The characteristics of this algorithm are in literature as [17];

- The method is improved from excavation on swarm such as fish schooling as well as bird flocking.

In this algorithm, the position of each agent is defined by XY axis position and also the velocity is expressed by VX (defines the velocity of X axis) as well as VY (introduce the velocity of Y axis). Each agent knows its best value so far \( p_{best} \) and its XY location. This data is comparison of personal experiences of each agent. Moreover, each agent knows the best amount so far in the group \( g_{best} \) value \( p_{best} \).

This information is comparison of knowledge of how the other agents around them have performed. Accordingly, each agent tries to update its position by [24];

- The current locations \((x, y)\),
- The current speed and velocities \((VX, VY)\),
- The interval among the current location and \( p_{best} \),
- The interval among the current location and \( g_{best} \).

This improvement can be introduced by the definition of velocity and the place of particle. Velocity of each agent can be improved as:

\[
x_i(t + 1) = x_i(t) + v_i(t + 1)
\]

\[
V_i(t + 1) = ow_i(t) + c_1 r_1(t)[p_{best}(t) - x_i(t)] + c_2 r_2(t)[leader(t) - x_i(t)]
\]

where

\(x_i\) position of agent \(i\) at iteration \(k\)  
\(v_i\) velocity of agent \(i\) at iteration \(k\)  
\(o\) inertia weighting  
\(c_{1,2}\) coefficient of slope  
\(r_{1,2}\) rand random number between 0 and 1  
\(leader\) archive of unbeatable particles  
\(p_{best}\) \(p_{best}\) of agent \(i\)  
\(g_{best}\) \(g_{best}\) of the group

More description of this algorithm with the value of parameters is presented in [24].

- PSO with Time-Varying Acceleration Coefficients (PSO-TVAC)

This algorithm is extended version of PSO which is described in [17]. The equation of PSO-TVAC for velocity updating can be expressed as:

\[
V_i(t + 1) = ow_i(t) + c_1 r_1(t)[p_{best}(t) - x_i(t)] + c_2 r_2(t)[leader(t) - x_i(t)]
\]

3.2.2. Bacteria Foraging Algorithm (BFA)

BFA is the second algorithm which is hybrid with PSOTVAC and is based on the hypothesis that animals research for nutrients which maximizes their energy intake \(E\) per unit time \(T\) spent for foraging [18]. The Escherichia coli bacterium is probably the best understood micro organism. Generally the bacteria move for a longer interval in a friendly space. This algorithm is based on four main steps in literature as:

- Chemo-tactic behavior of E. Coli

The bacterium sometimes tumbles after a tumble or tumbles after a run [19]. This alternation between the two
modes will go the bacterium, and this enables it to ‘search’ for nutrients. If \( \theta(j, k, l) \) represent the position of the every individual in the population of \( S \) bacterial at the \( j \)th chemotactic step, \( k \)th reproduction stage and \( l \)th removal, the motion of bacterium may be presented by:

\[
\theta^*(j + 1, k, l) \approx \theta(j, k, l) + C(i)\phi(j)
\]

(8)

where \( C(i) (i = 1, 2, \ldots, S) \) is the size of the stage taken in the accidental direction specified by the tumble. \( \phi(j) \) is the random direction of movement after a tumble and \( J(i, j, k, l) \) is the fitness, which also denote the cost at the location of the \( i \)th bacterium \( \theta(j, k, l) \in R^n \). Also if at \( \theta(j + 1, k, l) \) the cost \( J(i, j + 1, k, l) \) is better (lower) than at \( \theta(j, k, l) \), then another stage of size \( C(i) \) in this similar orientation will be taken. Otherwise, bacteria will tumble via taking another stage of size \( C(i) \) in random direction \( \phi(j) \) in order to seek better nutrient environment.

- **Swarming**

An interesting group trend has been obtained for several motile species of bacteria including \( E. coli \) and \( Salmonella typhimurium \) [17] to achieve the function to model the cell-to-cell signaling with an attractant and a repellant. The \( E. coli \) swarming mathematical equation can be represented by:

\[
J_w(\theta, P(j, k, l)) = \sum_{i=1}^{S} J_w^i(\theta, \theta^*(j, k, l))
\]

(9)

\[
= \sum_{i=1}^{S} \left[ -d_{\text{attract}} \exp(-\omega_{\text{attract}} \sum_{m=1}^{p} (\theta^m - \theta^*_{m})^2) \right]
\]

\[
+ \sum_{i=1}^{S} \left[ -h_{\text{repellent}} \exp(-\omega_{\text{repellent}} \sum_{m=1}^{p} (\theta^m - \theta^*_{m})^2) \right]
\]

The \( J_w(\theta, P(j, k, l)) \) is the additional cost function added to the actual objective function (for minimization) to present a time varying objective function. The parameters of \( d_{\text{attract}}, \omega_{\text{attract}}, h_{\text{repellent}} \) and \( \omega_{\text{repellent}} \) are set as follows:

\[
\omega_{\text{attract}} = 0.2, \omega_{\text{repellent}} = 10, d_{\text{attract}} = h_{\text{repellent}}
\]

(10)

\( S \) total number of bacteria

\( p \) number of parameters to be optimized which are indicated in each bacterium

\( \theta = [\theta_1, \theta_2, \ldots, \theta_p]^T \) is a point in the \( p \)-dimensional search domain

\( d_{\text{attract}} \) depth of the attractant released by the cell

\( w_{\text{attract}} \) measure of the width of the attractant signal

\( h_{\text{repellent}} = d_{\text{attract}} \) height of the repellent effect

\( w_{\text{repellent}} \) measure of the width of the repellent

- **Reproduction**

According to the rules of evolution, single will reproduce themselves in appropriate conditions in a certain way. For bacterial, a reproduction stage takes location after all chemotactic steps.

\[
J_{\text{health}}^i = \sum_{j=1}^{N+1} J(i, j, k, l)
\]

(11)

where \( J_{\text{health}}^i \) = health of bacterium \( i \)

- **Elimination-dispersion**

In evolutionary procedure, removing and scattering events can occur such that bacteria in a region are killed or a group is dispersed into a new section of the environment due to some effects. They have the effect of possibly destroying chemotactic progress, however they also have the influence of helping in chemotaxis, since dispersal may place bacteria near good food sources. In this paper, the popular penalty function method employs functions to decrease the fitness of the particle in ratio to the magnitude of the limitations oscillations. The penalty parameters are chosen carefully to recognize between feasible and impractical solution. The evaluation function is defined as follows [19]:

\[
f(P_i) = \sum_{j=1}^{m} F_j(P_i) + \alpha \left[ \sum_{j=1}^{m} P_j - (P_D + P_L) \right]^2 + \beta \left[ \sum_{j=1}^{m} P_j(\text{violation}) \right]^2
\]

where \( \alpha \) and \( \beta \) are the penalty factors.

The main flowchart of proposed model is presented in Figure 2.

### 3.2.3. Combination of the Proposed Forecast Engine with Proposed Algorithm

Combination of proposed PSOTVAC/BFA algorithm with hybrid forecast engine is presented in Figure 3.

As presented in this figure, at first NN1 is trained by the LM learning algorithm. The training phase is interrupted according to the early stopping criteria for preventing the overfitting problem [22]. After that, the evaluated weights are transferred to the proposed PSOTVAC/BFA algorithm. By this transferring, PSOTVAC/BFA keeps the training procedure of NN1 by converting this procedure as an optimization problem. The objective function of the optimization problem is the forecast error of NN1 that should be minimized. Error of validation samples or validation error [25] is adopted as the error function in the training process for all of the learning algorithms of the HNN (LM, SCG and CGP) and PSOTVAC/BFA. It is clear that the NNs and the PSOTVAC/BFA component have the same training samples and validation samples such that
the weights can be transferred between them. Accordingly, the PSOTVAC/BFA tries to better minimize the validation error of NN1 after its training algorithm is terminated. Although LM, SCG and CGP are computationally efficient learning algorithms, they search the solution area in a particular direction [26]. Hence, it is possible that the mentioned algorithms trapped in a local minimum to learn the nonlinear prediction procedure, without being able to escape out. Against, the PSOTVAC/BFA algorithm, with its strong exploration ability, can search the solution space in different paths which makes the training procedure, more strong and accurate.

In this optimization process at first, the proposed PSOTVAC/BFA algorithm is initialized. Decision variables of the PSOTVAC/BFA algorithm, are selected as the weights of the NNs based forecasting engine. Then, the objective function of the proposed algorithm is evaluated. The formulation of objective function is presented as:

\[
\text{WMAPE} = \frac{1}{N_h} \sum_{t=1}^{N_h} \frac{|SR_{\text{act}(t)} - SR_{\text{for}(t)}|}{SR_{\text{act}(t)}}
\]

where \(SR_{\text{act}(t)}\) and \(SR_{\text{for}(t)}\) represent actual and forecasted value of SR requirement for hour \(t\), respectively; \(N_h\) indicates forecast horizon in terms of hour. For WMAPE, \(N_h = 168\).

After that, the PSOTVAC/BFA compares the value of the objective function with its value in the previous iteration of PSOTVAC/BFA. By this comparison it can select the best value of objective function [13]. If the stopping criterion is satisfied, it will go to the next step; otherwise, it returns to calculation of objective function again and keeps this procedure. Finally, the individual of PSOTVAC/BFA algorithm leading to the lowest objective function value is selected as the PSOTVAC/BFA solution. The weights of this optimal solution are loaded to NN\(_j\), and it goes to the next step of forecast engine. After optimizing the final NN (NN\(_j\)), all of the forecast engine is trained and ready for the prediction of the future values of SR requirement.

### 4. Numerical Results

In this section the proposed prediction method is tested by the real data of two well-known Pennsylvania–New Jersey–Maryland (PJM) electricity market which has been obtained from its website [27]. Also, this point should be mentioned that, the sequential reserve procurement trend was taken in the first years of operation of the California Independent System Operator (ISO) and was demonstrated to be a poor reserve market design prone to manipulations by the market players [28]. By considering the [29], even without market power, such sequential markets are sensitive to price reversals (where the lower quality reserves mean a higher price) and create wrong motivations and wrong presentation of bids. Furthermore, the functional make-up of the energy and reserves markets means another asthenia of this model in cases when the original energy schedule does not cause the obligation of enough capacity so as to confront the reserve requirements.

For example, the operator of system has to refer to out-of-market handworks that, obviously, cause losses of social welfare [28]. According to the presented shortcoming, California experienced bulk price spikes within the summer of 2000 that cause the shut of the market until suitable and new rules were created. Market reform objected at instauration the market of ancillary service has been done in California. The preliminary objectives of the reform have been to improve revenue and liquidity in the ancillary service market and decrease purchase cost. Key elements of the California reform have been to move from a sequential to a simultaneous procurement auction for the different services and to allow the ISO to use the replacing between the different reserve types. More details about the California reform is presented in [29].

In this section for the starting point, the obtained features from proposed filtering section s presented which selected features among the 400 candidate inputs for the PJM electricity market in 4 June 2005 as:

\[
\{SR_{(t-1)}, SR_{(t-3)}', SR_{(t-8)}', SR_{(t-14)}', SR_{(t-24)}', SR_{(t-48)}', L_{(t-1)}', L_{(t-22)}', L_{(t-48)}', L_{(t-168)}, L_{(t-168)}', L_{(t-53)}', L_{(t-52)}', L_{(t-168)}, L_{(t-168)}', L_{(t-13)}', L_{(t-13)}', L_{(t-54)}', L_{(t-77)}', L_{(t-99)}', L_{(t-102)}', L_{(t-108)}'\}
\]

According to the presented features, 25 candidate inputs are selected for the proposed test day, which results in the large filtering ratio of 400/25 = 16. Where, the candidate set of inputs is efficiently refined by the proposed fuzzy clustering mechanism model. So, the proposed forecast engine is fed by the obtained features as input in this test day. The specifications of short-run trend (e.g. selection of \(SR_{(t-1)}', SR_{(t-3)}', L_{(t-1)}', \text{and } L_{(t-3)}\)), daily periodicity (such as selection of \(SR_{(t-24)}', SR_{(t-48)}', L_{(t-24)}', \text{and } L_{(t-48)}\)) and weekly periodicity (such as selection of \(SR_{(t-168)}', \text{and } L_{(t-168)}'\)) of the SR requirement signal can be seen from presented features.

For providing an appropriate insight about the proposed forecast engine’s performance in the training phase, a typical curve based on its error and training is represented in Figure 4. Accordingly, Mean Square Error (MSE) is considered as the error function for 24 h of the day before the forecast day. So, the training set is considered between the day of 19 November 2005 over the PJM electricity market. Moreover, termination mechanism is the most important aspect in forecast engine training process. Where, sometimes premature convergence can
be occurred in this procedure. On the other hand, the test error will decrease through the initial iterations in training phase for NN-based forecast engine. However, when the overfitting occurs in this procedure, the error will increase on the test set [30]. For this purpose, for each step of proposed NN based forecast engine by increasing the error function the training procedure will stop by replacing the last values of weight (based on proposed optimization method). As seen from this figure, the LM half cycle of the first, second and third iterations has 33, 92–57 = 35 and 157–112 = 45 training epochs (pass through the entire training set), respectively, determined by the early stopping technique. The PSOTVAC/BFA half cycle of the first, second and third iterations has 57–33 = 24, 112–92 = 20 and 178–157 = 21 generations, respectively, determined by the stopping condition of the PSOTVAC/BFA. It can be seen that the validation error decreases by 0.44/2.1e−6 = 201523 times.

Bearing in mind, the validation error is prediction error for the unseen data, this shows great generalization capability of the proposed hybrid model of forecast engine for the problem of SR requirement prediction. To give a better insight about the performance of the proposed hybrid training mechanism, the training phase of the first LM half cycle is continued after the early stopping point and represented by dashed-dotted line in Figure 4. It can be considered that the first LM half cycle can be considered as a MLP neural network with a single LM learning algorithm that is a conventional NN. According to the presented dashed-dotted line in this figure which is continued from the first step of LM learning, by more continuing the training phase of the single LM, not only the validation error does not decrease, but also this trend increases which proofs the overfitting problem in this procedure. So, after trapping in a local minimum in this phase, the LM cannot escape from this point and this problem leads to an overfitting problem. But, by considering the remaining hybrid process based on PSOTVAC/BFA algorithm, it can be considered that, this prediction trend is decreased monotonically in each step based this hybrid method. This trend demonstrates that the proposed PSOTVAC/BFA algorithm could escape the LM learning algorithm from the local points. Even considering the best point of the single LM, determined by the early stopping technique, its validation error decreases by. Where, this point can claim that, the validation error of the proposed hybrid forecast engine is efficient and better than the validation error of the conventional NN.

Tables 1 and 2 present the obtained results for SR requirement prediction of the PJM electricity market. Four test weeks corresponding to four seasons of year 2004 are considered in this experiment that are the fourth week of February, May, August and November for winter, spring, summer and fall seasons, respectively. For fair comparison of proposed method, all of the considered conditions in [12] have been implemented in this paper. Weekly Mean Absolute Percentage Error (WMAPE), as a well-known error measure, is considered in this paper and its obtained results is presented in Table 1 for MLP neural network with single LM learning algorithm and the proposed NN-based forecast engine (MLP with the hybrid training mechanism), support vector machine (SVM), radial basis function neural network, ARIMA and with comparison of the Enhanced Particle Swarm Optimization (EPSO)-based NN [12]. In this table, the last row presents the average WMAPE values for the four test weeks. Also, in this testing the same candidate set are considered, which by this method we can evaluate the effectiveness of the proposed forecast engine.

Furthermore, the provided results for weekly error variance $\sigma^2_{e, week}$ as a forecast uncertainty, are indicated in Table 2. In this status, the small value of variance means more stable prediction. $\sigma^2_{e, week}$ is defined as follows:

$$\sigma^2_{e, week} = \frac{1}{N_h} \sum_{t=1}^{N_h} \left( \frac{SR_{act}(t) - SR_{for}(t)}{SR_{act}(t)} \right)^2 - \text{WMAPE}$$

For $\sigma^2_{e, week}$ the forecast horizon $N_h = 168$. By considering to Table 2, it is clear that, the proposed forecast engine has lower $\sigma^2_{e, week}$ values than the MLP with single LM learning algorithm in all test weeks, which indicates the superiority of the proposed forecast engine in terms of forecast stability.

For more presentation of capabilities of proposed forecast engine, it is compared with four prediction models presented in [12,31]. For the sake of a fair comparison, the
same conditions have been considered between the proposed method and them. The mentioned models includes transfer function model with demand, transfer function model without demand, a naïve model and the standard exponentially weighted moving average, denoted by M1, M2, M3 and M4, respectively. Obtained results from the mentioned strategies are depicted in Tables 3 and 4. Accordingly, the proposed method has considerably lower WMAPE and $\sigma_e^2$ week than all four models of [31] and [12] in all test weeks, which demonstrates the superiority of proposed model and optimization algorithm compared with EPSO and the presented models in [31].

Moreover, the graphical view of the proposed forecast method, its results for the summer test week of the PJM electricity market based on the highest value presented in Table 1, presented in Figure 5. By this figure, it can be considered that the proposed method could forecast with low error variation.

5. Conclusion

According to the important role of SR in ancillary service in electricity market, an accurate and strong prediction method is very important for solving the forecasting problem in this field, where it can be very useful in ISO and market participants. Accordingly, we proposed a new prediction model in this paper based on new feature selection and hybrid NN-based forecast engine. Moreover, the proposed forecast engine is based on three-stage NNs which are learned by different learning algorithms. Also, all weights of the proposed NNs are tuned by a new meta-heuristic method named PSOTVAC/BFA. By this method, high convergence rate and good global search ability have been provided for proposed model. To demonstrate the superiority of proposed model, a real world data of the PJM is considered for analyzing. Consequently, obtained results are compared with other presented techniques for this problem which proofs the validity and superiority of proposed method.

References

[1] Noruzi A, Banki T, Abedinia O, et al. A new method for probabilistic assessments in power systems, combining monte carlo and stochastic-algebraic methods. Complexity 2015;21(2):100–110.
[2] Ahmadian I, Abedinia O, Ghadimi N. Fuzzy stochastic long-term model with consideration of uncertainties for deployment of distributed energy resources using interactive honey bee mating optimization. Front Energy. 2014;8(4):412–425.
[3] Abedinia O, Amjadi, N, Ghadimi N. Solar energy forecasting based on hybrid neural network and improved metaheuristic algorithm. Comput Intell. 2017;34(1):241–260.
[4] Mohammadi M, Talebpour F, Safae, E, et al. Small-scale building load forecast based on hybrid forecast engine. Neural Process Lett. 2017;1–23.
[5] Eskandari Nasab M, Maleksaeedi I, Mohammadi M, et al. A new multiobjective allocator of capacitor banks and distributed generations using a new investigated differential evolution. Complexity. 2014;19(5):40–54.
[6] Abedinia O, Bekravi M, Ghadimi N. Intelligent controller based wide-area control in power system. Int J Uncertain Fuzz Knowledge Based Syst. 2017;25(01):1–30.
[7] Abedinia O, Ghadimi N. Modified harmony search algorithm based unit commitment with plug-in hybrid electric vehicles. J Artif Intell Electric Eng. 2013;2(6):49–62.
[8] Ghadimi N, Akbarimajd A, Shayeghi H, et al. Application of a new hybrid forecast engine with feature selection algorithm in a power system. Int J Ambient Energy. 2017;1–10.
[9] Ghadimi N, Akbarimajd A, Shayeghi H, et al. A new prediction model based on multi-block forecast engine in smart grid. J Ambient Intell Humaniz Comput. 2017;16.
[10] Liu Y, Wang W, Ghadimi N. Electricity load forecasting by an improved forecast engine for building level consumers. Energy. 2017;139:18–30.
[11] Gollou AR, Ghadimi N. A new feature selection and hybrid forecast engine for day-ahead price forecasting of electricity markets. J Intell Fuzzy Syst. 2017;32(6):4031–4045.
[12] Ghadimi N, Firouz MH. Short-term management of hydro-power systems based on uncertainty model in electricity markets. J Power Technol. 2015;95(4):265.
[13] Mirzapouri, F, Lakzaei M, Varamini G, et al. A new prediction model of battery and wind-solar output in hybrid power system. J Ambient Intell Humaniz Comput. 2017;1–11.
[14] Ebrahimian H, Barmayoon S, Mohammadi M, et al. The price prediction for the energy market based on a new method. Econ Res. 2018;31(1):313–337.
[15] Witten IH, Frank E. Data mining: practical machine learning tools and techniques. 3rd ed. San Francisco (CA): Morgan Kaufmann; 2011.

[16] B. Li, A new approach to cluster analysis: the clustering-function-based method. J R Stat Soc Series B Stat Methodolo 2006;68:457–476.

[17] Fan S, Liao JR, Yokoyama R, et al. Forecasting the wind generation using a two-stage network based on meteorological information. IEEE Trans. Energy Conversion. 2009;24(2):474–482.

[18] Leng H, Li X, Zhu J, et al. A new wind power prediction method based on ridgelet transforms, hybrid feature selection and closed-loop forecasting. Adv Eng Inform 2018;36:20–30.

[19] Ebrahimian H, Barmayoon S, Mohammadi M, et al. The price prediction for the energy market based on a new method. Econ Res 2018;31(1):313–337.

[20] Gonzalez AM, San Roque AM, Gonzalez JG. Modeling and forecasting electricity prices with input/output hidden Markov models. IEEE Trans Power Syst. 2005 May;20(2):13–24.

[21] Abedinia O, Amjady N. Short-term load forecast of electrical power system by radial basis function neural network and new stochastic search algorithm. Int Trans Electric Energy Syst. 2016;26(7):1511–1525.

[22] Abedinia O, Amjady N, Zareipour H. A new feature selection technique for load and price forecast of electrical power systems. IEEE Trans Power Syst. 2017;32(1):62–74.

[23] U.S. Federal Energy Regulatory Commission (FERC). Assessing the state of wind energy in wholesale electricity markets: Comments of Basin Electric Power Co-operative. FERC Docket No. AD04-13-000; 2005.

[24] Prasad M, Lin YY, Lin CT, et al. A new data-driven neural fuzzy system with collaborative fuzzy clustering mechanism. Neurocomputing. 2015;167:558–568.

[25] Hush DR, Horne BG. Progress in supervised Neural Networks. IEEE Signal Process Mag, 1993 Jan:8–39.

[26] Abedinia O, Amjady N. Short-term wind power prediction based on Hybrid Neural Network and chaotic shark smell optimization. Int J Precision Eng Manufact Green Technol. 2015;2(3):245–254.

[27] PJM Web Site. Available from: http://www.pjm.com.

[28] Galiana FD, Bouffard F, Arroyo JM, et al. Scheduling and pricing of coupled energy and primary, secondary, and tertiary reserves. Proc IEEE. 2005;93(11):1970–1983.

[29] Oren SS. Design of ancillary service markets. In: Proceedings of the 34th Annual Hawaii International Conference on System Sciences; 2001. p. 1–9.

[30] Open access same time information system (OASIS) of California electricity market. Available from: http://oasis.caiso.com/.

[31] Nogales FJ, Conejo AJ, Electricity price forecasting through transfer function models. J Oper Res Soc 2006;57(4):350–356.