Accommodation of curtailed wind power by electric water heaters based on a new hybrid prediction approach

Yang YU^{1,2}, Zengqiang MI^{1,2}, Xiaoming ZHENG¹, Da CHANG¹

Abstract Wind power curtailment is of great importance with the increase of large-scale wind power connected to the grid. A new concept of redundant wind power accommodated by dispatching electric water heaters (EWHs) is developed in the paper. Precise predictions of wind power and EWHs load power are the basis for this work. A hybrid multi-kernel prediction approach integrating an adaptive fruit fly optimization algorithm (AFOA) and multi-kernel relevance vector machine (MKRVM) is proposed to deal with the sample distribution of multi-source heterogeneous features uncovered by an energy entropy method, where AFOA is used to determine the kernel parameters in MKRVM adaptively and avoid the arbitrariness. For the large computation of the prediction approach, parallel computation based on the Hadoop cluster is used to accelerate the calculation. Then, an

economic dispatching model for accommodating wind power is built taking into account the penalty of curtailed wind power and the operating cost of EWHs. The proposed scheme is implemented in an intelligent residential district. The results show that the optimization performance of the hybrid prediction approach is superior to those of four usual optimization algorithms in this case. Regular or orderly scheduling of EWHs enables accommodation of superfluous wind power and reduces dispatch cost.

Keywords Wind power curtailment, Electric water heaters, Dispatching model, Relevance vector machine, Prediction model

1 Introduction

The huge demand for energy in modern society has triggered a rapid development of wind power. By the end of 2016, the cumulative installed capacity of wind power had soared to 486749 MW around the world, 30% of which is contributed by China. Nevertheless, due to lack of peak-shaving of hydroelectric power and the existence of numerous combined heat and power (CHP) plants, a portion of wind power is curtailed in winter in northern China. From the statistical data on wind power generation in China, the average of curtailed wind power is about 15%.

To help dealing with the issue of curtailed wind power, many measures have been proposed, originating from both the supply-side and demand-side. In addition to CHP plants, heat pumps [1–3] and electric boilers [4, 5] represent two typical approaches on the supply-side to integrate a high level of wind power with large-scale bulk power grids. Economic comparison between heat pumps and

CrossCheck date: 21 June 2018

Received: 12 October 2017 / Accepted: 21 June 2018 / Published online: 3 September 2018
© The Author(s) 2018

✉ Yang YU
yym0401@163.com

Zengqiang MI
mizengqiang@sina.com

Xiaoming ZHENG
498624179@qq.com

Da CHANG
939328521@qq.com

¹ State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, North China Electric Power University, Baoding 071003, China

² Hebei Key Laboratory of Distributed Energy Storage and Micro-grid, North China Electric Power University, Baoding 071003, China

electric boilers in Denmark was evaluated in [6]. This indicated that using the option of heat storage technologies can lower the cost of system operation. Wind power curtailment on the supply-side will require an increase in the initial investment and take up land resources. This leads to the consideration of demand-side accommodation, using such as electric vehicles [7–9], distributed energy storage [10–12], thermostatically controlled load [13–16]. These measures on demand-side accommodation have gained general approval due to inexpensive cost, large capacity, and immense potential. In contrast to other thermostatically controlled loads, such as air-conditioners and refrigerators, electric water heaters (EWHs) have some particular features in curtailing wind power: ① EWHs make up a large proportion of the power load in residential appliances and their electric consumption is closely related to the daily load [17]; ② the characteristics of high thermal inertia and large storage capacity make EWHs participate in power regulation continuously, variably and widely; ③ as a pure resistor load, EWH responds to the instruction rapidly and it is fit for demand response (DR) load. A model of EWHs aggregated for wind generation is presented in [18, 19]. This provides a dual advantage for both the grid and participating customers. The dynamic behavior based on a partial differential equation for EWHs is modeled for DR in [17], and it shows a better performance than the conventional method. This study primarily concentrated on aggregated and dynamic modeling for EWHs, while utilization of EWHs to reduce wind power curtailment still lacks systematic discussion. On the basis of the aggregated model, thermostat setpoint control consisting of price information is shown in [18] to balance wind power.

Given the randomness and fluctuation of wind power, previous studies have indicated [20–22] that the prediction model matching the time scale of system dispatch is a fundamental issue on large scale wind power penetration into the grid, especially for short-term forecasting (24–48 hours in advance). That is the issue with we are concerned. A novel idea of reducing wind power curtailment by EWHs is proposed based on a dispatching model of economic accommodation, where the accurate prediction of wind power and EWHs' load power are of great importance. For wind power or EWHs' load power, the original data reveals multi-source heterogeneous characteristics, such as historical load, meteorological data (temperature, precipitation, air pressure, relative humidity, wind direction, wind speed), and electricity price. It implies that inherent dispersion, diversity and complexity of input data have to be taken into account in a big data environment for an accurate load forecast, which leads to limitations in computation speed and scale for conventional prediction algorithms. Hence, parallel processing frameworks are introduced into load prediction

[23–25]. However, the aim of parallel computation as reported in the literature is principally to accelerate calculation speed and improve prediction accuracy. Such work ignores the multi-source heterogeneous characteristics of input data. Recently, multi-kernel (MK) learning has developed to solve the multi-factor problems [26]. An improved MK support vector machine (SVM) is constructed in [27, 28], which yields a more satisfactory accuracy in contrast with single kernel (SK) SVM. A small section of current MK research is based on a simply weighted combination of global kernel function and local kernel function. For SVM, the kernel function must satisfy the Mercer condition and the penalty coefficient is difficult to determine. In response, the relevance vector machine (RVM) appeared [29, 30], in which one of the benefits is that its kernel function does not need to meet the Mercer condition and the limitation of the penalty coefficient is relaxed through a probability model. The choice of the kernel and its parameters is also a key problem for RVMs like SVM. Currently, the SK used in RVM is mainstream [31, 32]. There is a small number of studies on weighted MK-RVM relying on previous experience [33] or an incremental learning approach [34]. After the above analysis, from a perhaps surprising perspective of big data with MapReduce based parallel computation, a new prediction approach integrating MK-RVM and adaptive fruit fly optimization algorithm (AFOA) is presented which takes into account the characteristics of sample distribution. AFOA is used to determine the kernel parameters in MK-RVM adaptively; in this way subjectivity and arbitrariness can be avoided. The correctness and feasibility of the proposed algorithm is established through the actual sampling data from an intelligent residential district.

The novelty in this paper is to supply a new concept of reducing wind power curtailment by storage-type EWHs from the aspect of DR and an economic dispatching model is built which considers the penalty of wind power curtailment and the operating cost of EWHs. Accurate predictions of wind power and EWHs' load power form the basis of implementing the dispatching model. Therefore, a hybrid MK prediction approach integrating AFOA and multi-kernel relevance vector machine (MKRVM) is presented in consideration of sample distribution of multi-source heterogeneous features determined by the energy entropy method. To tackle the issue of computation speed of the presented prediction approach, a MapReduce model based parallel computation method is introduced into the prediction approach and used to speed up the operation. The structure of the paper is as follows: Sect. 2 describes the hybrid prediction approach. The dispatching model for wind power curtailed with EWHs is established in Sect. 3, and correctness and validity of the algorithm are

established in Sect. 4 through an intelligent residential district. Section 5 gives the conclusions.

2 Representation of AFOA-MKRVM prediction approach

2.1 Sample distribution of multi-source heterogeneous features

Let the number of factors used for prediction be M , and these factors' set is written as $\{X_i, i = 1, 2, \dots, M\}$, $X_i = [x_1, x_2, \dots, x_N] \in \mathbf{R}_{1 \times N}$. The predicted value is $Y_i = [y_1, y_2, \dots, y_T] \in \mathbf{R}_{1 \times T}$. A weighted arithmetic mean based sample distribution method is designed to handle multi-source heterogeneous features of input data.

For each factor, the weighted arithmetic mean value $\bar{x}$ is calculated through the sampling points and written in (1). Then, the Euclidean distance between x_i and $\bar{x}$ is calculated and denoted as d_i , where the maximum Euclidean distance is written as $d_{\max}$. With the aid of the concept of energy entropy, let energy entropy of the i -th sampling point be $E_i = d_i^2$. The discriminant function of the distribution characteristic can be defined in (2). A threshold c^* is set to identify the characteristic of the sampling data. If $c \leq c^*$, the sampling data represents the globality. In this case a global kernel function, such as a polynomial function, should be selected. On the other hand, if the sampling data appears in the locality, the kernel describing the local feature should be chosen, and the Gaussian radial basis function (RBF) is chosen in this context. After comprehensive consideration, c^* is set to 0.5.

$$\bar{x} = \sum_{i=1}^N x_i f_i / \sum_{i=1}^N f_i \quad (1)$$

$$c = \sum_{i=1}^H E_i / \sum_{i=1}^N E_i \quad (2)$$

where f_i is the frequency of x_i appearing in X_i ; H is the number of sampling points indicating that the energy entropy is less than $d_{\max}^2/4$.

2.2 Brief description of RVM

To overcome the difficulties encountered in SVM, such as the uncertainty of the penalty coefficient and the rigorosity of the Mercer principle of kernel function, an uncertainty analysis and sparse probabilistic model-based RVM prediction approach is proposed in [29] under a Bayesian framework. For a given training sample set $X'_i = [x'_1, x'_2, \dots, x'_n]$, where n is the number of the sampling points, the output vector is written as below:

$$Y' = f(X', \omega) + \varepsilon \quad (3)$$

where X' is the input vector; Y' is the output or target vector; ω is the weighted vector, $\omega = [\omega_0, \omega_1, \dots, \omega_n]$; ε is the random noise obeying the normal distribution with zero mean and σ^2 variance; f is the map from X' to Y' . The relation among the training set satisfies the independent distribution. Then, the prediction function of RVM can be written as:

$$f(X', \omega) = \sum_{i=1}^N \omega_i K(X', x'_i) + \omega_0 \quad (4)$$

where $K(X', x'_i)$ is the kernel function. The Gaussian prior probability $P(\omega|\alpha)$ of hyper-parameter α in accordance with ω is defined as:

$$P(\omega|\alpha) = \prod_{i=1}^N \frac{\alpha_i}{\sqrt{2\pi}} e^{-\frac{\omega_i^2 \alpha_i}{2}} \quad (5)$$

In terms of a Bayesian criterion, the posterior normal distribution $P(\alpha, \sigma^2|Y')$ with respect to ω and Y' can be obtained and the optimal solutions α_{MP} and σ_{MP}^2 for $P(\alpha, \sigma^2|Y')$ will be obtained through iterating the following equations step by step:

$$\alpha_{MP,k} = \frac{1 - \alpha_i \psi_{ii}}{\mu_i^2} \quad (6)$$

$$\sigma_{MP,k}^2 = \frac{\|Y' - \Phi(X')\mu\|^2}{N - \sum_{i=0}^N (1 - \alpha_i \psi_{ii})} \quad (7)$$

where ψ_{ii} is the element on the diagonal line for the posterior covariance matrix ψ , $\psi = (\sigma^{-2} \Phi^T(X') \Phi(X') + A)^{-1}$, $A = \text{diag}(\alpha_0, \alpha_1, \dots, \alpha_n)$; μ is the mean value, $\mu = \sigma^{-2} \psi \Phi^T(X') Y'$; $\Phi(X')$ is the kernel matrix, as shown in (8).

$$\Phi(X') = \begin{bmatrix} 1 & K(x'_1, x'_1) & K(x'_1, x'_2) & \cdots & K(x'_1, x'_n) \\ 1 & K(x'_2, x'_1) & K(x'_2, x'_2) & \cdots & K(x'_2, x'_n) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & K(x'_n, x'_1) & K(x'_n, x'_2) & \cdots & K(x'_n, x'_n) \end{bmatrix} \quad (8)$$

Eventually, the prediction value y^* can be obtained by the following equation:

$$y^* = \mu^T \varphi(x^*) \quad (9)$$

where $\varphi(x^*) = [1, K(x^*, x'_1), K(x^*, x'_2), \dots, K(x^*, x'_n)]$.

2.3 Generation of MK function

MK aggregation through fusing local and global features is an effective solution and can integrate the advantages of each kind of SK function [35]. There are many common SK functions, such as the linear function, Gaussian RBF, and the polynomial function. The polynomial function is good at extracting the global feature of sampling data, but it is poor in learning ability. RBF differs from the polynomial function with its strong learning ability and it is fit for handling local information. Through the organic combination of these two SKs shown in (10), the MK describing multi-source heterogeneous features will be improved. For each factor, a SK has been chosen through the energy entropy method described in Sect. 2.1.

$$K_{MKL}(x, y) = \sum_{i=1}^{M_1} \lambda_i K_{Gauss}^i(x, x_i) + \sum_{j=1}^{M_2} \theta_j K_{Poly}^j(x, x_j) \\ = \sum_{i=1}^{M_1} \lambda_i \exp\left(-\frac{\|x - x_i\|^2}{2\sigma_i^2}\right) + \sum_{j=1}^{M_2} \theta_j (xx_j + 1)^{p_j} \quad (10)$$

where $K_{Gauss}^i(x, x_i)$ and $K_{Poly}^j(x, x_j)$ are the Gaussian RBF kernel and polynomial kernel respectively; $\lambda_i, \theta_j \in [0, 1]$ are the weighted coefficients, $\sum_{i=1}^{M_1} \lambda_i + \sum_{j=1}^{M_2} \theta_j = 1$; M_1 and M_2 indicate the numbers of the Gaussian RBF kernel and polynomial kernel respectively, $M_1 + M_2 = M$; σ_i and p_j are the parameters of the Gaussian RBF kernel and polynomial kernel still to be determined.

The generation of the MK function is shown in Fig. 1. On account of the nonlinearity MK has difficulty in ensuring that the local optimal solution is obtained, the MK is constructed with multiple SK functions through a weighted linear method in the context. For different application fields, the number of SKs included in the MK should be determined in accordance with actual conditions.

Fig. 1 Production flow of multi-kernel considering sample distribution

For this study, the number of input data is 11. The features of sample distributions for the locality exist in load power, air quality index, air temperature, air pressure, relative humidity, vapour pressure, precipitation and EWHs' power, and RBF kernels are selected to describe them. The others are wind speed, wind direction and similar day, which should be exhibited by polynomial functions because of their global features.

2.4 Proposed FOA-MKRVM prediction method

The fruit fly optimization algorithm (FOA) is a new global optimization approach based on the foraging behavior of the fruit fly. For the constructed multi-kernel, there are $2 \times M$ unknown parameters, where the numbers of weighted coefficients and the kernel parameters are both M . To remove the subjectivity in parameter determination, FOA is used to optimize the parameters of MKRVM. Then, a hybrid AFOA-MKRVM prediction method with self-adaptive generation parameters is proposed. Its specific implementation is as follows:

Step 1: initialize parameters for the fruit fly swarm, including the population size N_e , the dimension D_e , the maximum iteration time G_e , the threshold of training error e , and the initial locations X_0 and Y_0 .

Step 2: initialize parameters for MK-RVM, including the weighted coefficients and the kernel parameters.

Step 3: start the optimal calculation for the fruit fly and use FOA to regulate the parameters in MK-RVM.

Step 4: update the location and the iteration step in terms of random direction. The individual fruit fly is iteratively evolved by smell search of random direction and adaptive step. The iteration step is adjusted adaptively through the optimal smell concentration in the former generation and the number of iterations in the current generation, which is shown as follows:

$$\begin{cases} t_i = \frac{v}{S_{i-1,opt}} \exp\left(-\vartheta \left(\frac{i}{G_{max}}\right)^\eta\right) + t_{min} \\ X_i = X_{i-1} + \zeta t_i \\ Y_i = Y_{i-1} + \zeta t_i \end{cases} \quad (11)$$

where t_i is the adaptive step; v is the regulatory coefficient in $(0, 1]$; $S_{i-1,opt}$ is the optimal smell concentration in the former generation of the fruit fly swarm; ϑ is the constraint coefficient in $(0, 1)$; η is an integer in $[1, 10]$; t_{min} is the minimum iteration step; ζ is the regulatory coefficient distributed randomly in $[-1, 1]$; X_i and Y_i are the location coordinates of the fruit fly.

Step 5: for the unknown location of food, the distance r_i is calculated and then the smell concentration value S_i is calculated as shown in (12).

$$\begin{cases} r_i = \sqrt{X_i^2 + Y_i^2} \\ S_i = \frac{1}{r_i} \end{cases} \quad (12)$$

Step 6: in terms of the smell concentration value S_i , the prediction values and their root mean square errors are outputted. Meanwhile, the fitness function $F(\bullet)$ is constructed through the smell concentration value S_i and the maximal smell concentration among the swarm is found. This is shown in the following:

$$\begin{cases} s_i = F(S_i) \\ C_{opt} = \max(s_i) \end{cases} \quad (13)$$

Step 7: retain the optimal smell concentration value and its location coordinates; then the fruit fly swarm will fly towards the location by means of vision.

$$\begin{cases} S_{opt} = C_{opt} \\ X = X_{opt} \\ Y = Y_{opt} \end{cases} \quad (14)$$

Step 8: repeat the iterations from **Step 2** to **Step 6** and judge whether the smell concentration is superior to that of the former generation. If the iteration time extends the maximal iteration time G_e or the training error reaches the threshold e , the iteration will be stopped.

2.5 Parallel prediction based on MapReduce model

In practice, for the same size of sample, the computation speed in MK learning prediction is significantly slower than SK learning prediction. To speed up the computation of FOA-MKRVM, a parallel prediction algorithm based on MapReduce model is implemented and its flow diagram is shown in Fig. 2.

Step 1: construct the training sample set including N_1 sampling points and the testing sample set including S sampling points.

Step 2: constitute the MK function in terms of Fig. 1.

Step 3: split the samples from the training set into Q subsets. Then, each subset consists of about N_1/Q training sampling points. These subsets are uploaded to the distributed file system of Hadoop.

Step 4: train Q map functions from **Step 3** in parallel through the hybrid FOA-MKRVM prediction method and output RVs of every subset as value.

Step 5: the RVs of every subset are aggregated to form a reduce function consisting of the entire RVs. The hybrid FOA-MKRVM prediction method is reused to train the entire RVs and form the final RVs. The parallel prediction model is ultimately constructed and is used to forecast the testing sample set.

Fig. 2 Flow diagram of parallel prediction based on MapReduce model

2.6 Performance evaluation

The mean absolute percentage error (MAPE) is selected as the index of performance evaluation and is shown as the following:

$$E_{MAPE} = \left(\sum_{t=1}^T |Y_t - y_t| / y_t \right) / n \times 100\% \quad (15)$$

where Y_t is the prediction value; y_t is the actual value. For the load forecast, a smaller value of MAPE indicates a higher precision accuracy.

Additionally, the general index of speedup ratio is employed to evaluate the parallel efficiency and is described as:

$$R_{SR} = T_{sd} / T_{pl} \quad (16)$$

where R_{SR} is the speedup ratio; T_{sd} and T_{pl} are the operation times of standalone and parallel respectively.

3 Dispatching model for curtailment of wind power with EWHs

A dispatching model of economic accommodation for wind power curtailment through EWHs is established. Economic accommodation of curtailed wind power aims to

minimize total coal consumption. If the integration of wind power does not lead to threatening the secure operation of thermal power plants, wind power curtailment will be reasonably executed, where EWHs are connected to maximize the cost saving.

3.1 Objective function

The objective function in economic accommodation of curtailed wind power is constructed as follows:

$$\min \sum_{t=1}^T \left[\sum_{i=1}^{z_1} F_{TP}(P_{TP}^{i,t}) + \sum_{i=1}^{z_2} (F_{CHP}(P_{CHP}^{i,t}, Q_{CHP}^{i,t}) + \beta P_{CW}^t + \chi P_{IEWH}^t) \right] \quad (17)$$

$$F_{CHP}(P_{CHP}^{i,t}, Q_{CHP}^{i,t}) = b_{1i} + b_{2i}P_{CHP}^{i,t} + b_{3i}Q_{CHP}^{i,t} + b_{4i}(P_{CHP}^{i,t})^2 + b_{5i}P_{CHP}^{i,t}Q_{CHP}^{i,t} + b_{6i}(Q_{CHP}^{i,t})^2 \quad (18)$$

where z_1 and z_2 are the numbers of thermal power (TP) plants and CHP plants; $P_{TP}^{i,t}$ and $P_{CHP}^{i,t}$ are the power generated at the time t for the i -th TP and CHP units individually; $Q_{CHP}^{i,t}$ is the thermal load at the time t for the i -th CHP unit; F_{TP} is the coal consumption of the TP unit and can be written as $F_{TP}(P_{TP}^{i,t}) = a_{1i}(P_{TP}^{i,t})^2 + a_{2i}P_{TP}^{i,t} + a_{3i}$, a_{1i} , a_{2i} and a_{3i} are the coefficients depicting its coal consumption; F_{CHP} is the coal consumption of the CHP unit; b_{1i} , b_{2i} , b_{3i} , b_{4i} , b_{5i} and b_{6i} are the coefficients depicting its coal consumption; P_{CW}^t is the curtailed wind power at the time t ; β is the penalty coefficient; P_{IEWH}^t is the dispatched power of EWHs at the time t ; χ is the cost coefficient.

3.2 Constraint conditions

Power load balance constraint:

$$\sum_{t=1}^T \left(\sum_{i=1}^{z_1} P_{TP}^{i,t} + \sum_{i=1}^{z_2} P_{CHP}^{i,t} \right) + P_W^t = P_L^t + P_{IEWH}^t \quad (19)$$

where P_W^t is the consumed wind power; P_L^t is the power of load at the time t .

Thermal load balance constraint:

$$\sum_{t=1}^T \sum_{i=1}^{z_2} Q_{CHP}^{i,t} = Q^t \quad (20)$$

where Q^t is the thermal load at time t .

Technology constraints of TP units:

$$P_{TP,\min}^i \leq P_{TP}^{i,t} \leq P_{TP,\max}^i \quad (21)$$

$$r_{TP,\downarrow}^i \Delta t \leq P_{TP}^{i,t} - P_{TP}^{i,t-1} \leq r_{TP,\uparrow}^i \Delta t \quad (22)$$

where $P_{TP,\min}^i$ and $P_{TP,\max}^i$ are the minimum and maximum power for the i -th TP unit; $r_{TP,\downarrow}^i$ and $r_{TP,\uparrow}^i$ are the rates of ramp up and ramp down for the i -th TP unit; Δt is the dispatching time, $\Delta t = 1$ h.

Technology constraints of CHP units:

$$P_{CHP,\min}^i \leq P_{CHP}^{i,t} \leq P_{CHP,\max}^i \quad (23)$$

$$r_{CHP,\downarrow}^i \Delta t \leq P_{CHP}^{i,t} - P_{CHP}^{i,t-1} \leq r_{CHP,\uparrow}^i \Delta t \quad (24)$$

where $P_{CHP,\min}^i$ and $P_{CHP,\max}^i$ are the minimum and maximum power for the i -th CHP unit; $r_{CHP,\downarrow}^i$ and $r_{CHP,\uparrow}^i$ are the rates of ramp up and ramp down for the i -th CHP unit.

Power constraint of wind power units:

$$P_W^t + P_{CW}^t = P_{WP}^t \quad (25)$$

where P_{WP}^t is the predicted power of wind power at the time t .

Power constraint of EWHs:

$$0 \leq P_{EWH}^t + P_{IEWH}^t \leq P_{EWHp} \quad (26)$$

$$0 \leq E_{EWH}^t + E_{IEWH}^t \leq E_{EWHp} \quad (27)$$

$$E_{EWH}^t = \int \kappa P_{EWH}^t dt \quad (28)$$

$$E_{IEWH}^t = \int \kappa P_{IEWH}^t dt \quad (29)$$

where P_{EWH}^t is the predicted power of EWHs at the time t ; P_{EWHp} is the total power of EWHs in the grid; E_{EWH}^t is the predicted electrical energy of EWHs during the period of time t ; E_{IEWH}^t is the electrical energy consumed by dispatched EWHs during the period of time t ; E_{EWHp} is the total electrical energy that can be consumed by EWHs; κ is the efficiency coefficient.

4 Case study

The experimental data originates from an intelligent residential district powered by a small regional power grid in China. It includes three TP units (#1, #2, #3), three CHP units (#4, #5, #6), and a wind farm with 150 MW rated power, which includes 56×1.5 MW doubly-fed induction generators and 33×2.0 MW full converted permanent magnet synchronous generators. The generated power of the wind farm is transmitted into the residential district through a booster substation. The parameters of TP and CHP units are shown in Table 1 and Table 2. The data on EWHs with 540 MW rated power and 810 MWh maximum storage capacity in the residential district are gathered into a control centre. EWHs can complete the energy storage during 1.5 hours. The sampling frequency of data is

Table 1 Parameters of TP units

Unit	Capacity (MW)	a_{1i}	a_{2i}	a_{3i}
#1	300	21.2678	0.1715	0.000189
#2	100	23.8905	0.2716	0.000261
#3	100	23.8945	0.2809	0.000237

one hour; the sampling data are utilized for day-ahead dispatch.

4.1 Verification of prediction approach

In accordance with the prediction requirement and data scale, a Hadoop cluster with five nodes is established in a virtual machine, where one personal computer (PC) works as the master node and the remaining four PCs are configured as the slave nodes. The master node as a central server is responsible for resource allocation and task assignment. The slave nodes execute the tasks of store and operation.

First, the sample distributions of multi-source heterogeneous features are discussed. Nine features are selected for wind power prediction, which are historical load power, wind speed, wind direction, air temperature, air pressure, relative humidity, precipitation, similar day, and air quality index. Simultaneously, six features of historical load power, air temperature, air pressure, relative humidity, precipitation, and similar day are chosen for aggregated power prediction of EWHs. Wind power curtailment in northern China occurs frequently in winter. Hence, one-week of data from December 7th to December 13th, 2015 were extracted to analyze the heterogeneous features through some calculations in terms of (1) to (3). The results are shown in Fig. 3, which shows that the features of sample distributions for historical load power, air quality index, air temperature, air pressure, relative humidity, vapour pressure, precipitation and EWHs load power reflect local behavior. However, wind speed, wind direction, and similar day are distributed as globally.

Secondly, the training data collected from November 9th, 2015 to December 20th, 2015 is used to predict the load from December 21th, 2015 to December 27th, 2015. The parameters of the fruit fly swarm are shown as follows: the population size $N_e = 20$, maximum iteration time $G_e =$

1000, threshold of training error $e = 0.1$, the adaptive step $t_i = 0.1$, the regulatory coefficient $\nu = 0.6$, the constraint coefficient $\vartheta = 0.2$, $\eta = 5$, the minimum iteration step $t_{\min} = 0.1$, the regulatory coefficient $\varsigma = -0.2$. To evaluate the superiority of the proposed AFOA-MKRVM algorithm, the Gaussian RBF based single kernel RVM (RBF-SKRVM), the polynomial kernel based single kernel RVM (PK-SKRVM), least squares support vector machine (LS-SVM), and weighted least squares support vector machine (WLS-SVM) [37] are compared with the AFOA-MKRVM algorithm. The prediction curves on December 22th, 2015 through the five algorithms are plotted in Fig. 4 and the results of MAPE for seven days are listed in Table 3 and Table 4. In general, Table 3 and Table 4 indicate that the MAPE of prediction values of RBF-SKRVM is lower than that of PK-SKRVM. This is to some extent induced by the local behavior of the majority of sample distributions. However, the prediction accuracy of the proposed AFOA-MKRVM is the highest; especially since it can catch the highest and lowest points illustrated in Fig. 4. Relatively, the prediction accuracy in LS-SVM is the lowest because it neglects the distribution features in the data. The satisfactory prediction both in wind power and in EWHs load power establishes the perfect property of the hybrid MK algorithm in resolving multi-source heterogeneous features through a combination of various approaches.

To compare with the operation time of standalone and parallel, the sampling data are divided into six categories in terms of one-week data: C_1 to C_6 , such as C_1 which includes the first week data from November 9th, 2015 to November 15th, 2015, or C_2 which includes two weeks data from November 9th, 2015 to November 22th, 2015, and C_6 includes six weeks from November 9th, 2015 to December 20th, 2015. For every type of data, the presented algorithm operating in standalone and parallel is conducted respectively, and Fig. 5 gives the comparative result. It can be seen that the speedup ratio of parallel to standalone almost linearly increases with the number of clusters and data scale. For the same number of clusters, the greater data scale resolved in parallel brings a faster speedup ratio relatively; the tendency of the increase in speedup ratio with the number of cluster will gradually reduce a little.

Table 2 Parameters of CHP units

Unit	Capacity (MW)	b_{1i}	b_{2i}	b_{3i}	b_{4i}	b_{5i}	b_{6i}
#4	300	2.005	0.434	0.123	0.0009	0.00019	0.00021
#5	150	2.236	0.401	0.128	0.0008	0.00018	0.00026
#6	150	2.226	0.412	0.126	0.0007	0.00017	0.00025

Fig. 3 Sample distributions of multi-source heterogeneous features

4.2 Validation of dispatching model

The total load power, thermal load, wind power, and EWHs load power of a typical day in winter are selected

and plotted in Fig. 6, which demonstrates that the curtailment of wind power frequently occurs in the night from 11 p.m. to 7 a.m.. It is restricted by the minimum power outputs (MPOs) of TP and CHP units. For a TP unit, its

Fig. 4 Prediction results of wind power and EWHs load power with five algorithms

MPO is a constant, which is half of the generating power in that case. However, the MPO for a CHP varies with its thermal load. A big power will be forced to output with a large thermal load in the nighttime for heating supply. Combined with a large output of wind farm and a small load power, wind power will have to be curtailed to avoid the forced outage of TP and CHP plants, as shown in Fig. 6.

To reveal of the relationship between thermal load and accommodation rate of wind power, the dispatching model shown in (17) is solved by a genetic algorithm (GA) with invariable load power and the result is plotted in Fig. 7. It implies that the accommodation ratio of wind power declines with the augmentation of thermal load for the constraint relation of electricity to heat in CHP units.

Hence, when the supply side has been limited, the idea of supplementing the regulation of load power through EWHs from the demand side is a valid path to enlarge the system's peaking capability indirectly. Figure 8 compares the load power and wind power accommodation with and without dispatched EWHs respectively, and indicates that activation of EWHs during the period of wind power curtailment is actually able to increase the power consumed by load. The ratio of wind power connected into the grid is improved from 84.62% to 86.27%. Accordingly, the coal consumption is reduced from 9964.5 t to 9962.7 t as shown in Table 5. For this case, 540 MW/810 MWh EWHs can be used to a day-ahead dispatch. In terms of the prediction data, about 753.6 MWh or 93.04% EWHs has been heated that cannot be employed within the period of curtailed wind power. Therefore, a surplus of less than 10% EWHs can be dispatched to accommodate wind power, and this plays a weak role in wind power curtailment. In other words, a disorderly use of EWHs goes against the accommodation of wind power.

The relation among dispatched EWHs load power, curtailed wind power and coal consumption is uncovered by solving the constructed dispatch model with different orderly activation of EWHs and the result is given in Fig. 9, which shows that if about 351 MW EWHs are dispatched, the curtailed wind power will be totally accommodated in the case. Simultaneously, the cost of coal consumption obtains the lowest value. In addition, it is worth mentioning that heating EWHs orderly can shave the peak load in the daytime and shift this portion of power into the load valley in the nighttime. Inversely, the curtailed wind power and coal consumption increase nearly linearly with the reduction of dispatched EWHs' power.

Table 3 Prediction results of MAPE with five algorithms for wind power

Date	Prediction result of MAPE (%)				
	AFOA-MKRVM	RBF-SKRVM	PK-SKRVM	LS-SVM	WLS-SVM
1	2.40	4.49	7.06	8.15	7.89
2	2.60	4.79	6.46	7.56	7.02
3	1.99	4.00	7.30	9.15	8.76
4	2.05	4.32	6.85	8.01	7.61
5	2.02	4.11	7.71	9.88	9.24
6	1.84	3.76	8.03	10.12	9.99
7	2.00	3.85	7.37	9.65	9.02

Table 4 Prediction results of MAPE with five algorithms for EWHs load power

Date	Prediction result of MAPE (%)				
	AFOA-MKRVM	RBF-SKRVM	PK-SKRVM	LS-SVM	WLS-SVM
1	2.37	5.70	5.29	7.59	7.12
2	2.06	5.33	6.85	9.01	7.21
3	2.41	4.72	5.70	7.95	7.65
4	2.48	4.85	5.67	7.82	7.31
5	2.31	5.06	6.04	8.65	7.92
6	2.30	4.34	6.10	8.31	7.77
7	2.25	4.53	6.17	9.01	8.56

Fig. 5 Comparison of operation time for the proposed algorithm running in standalone or parallel **Fig. 7** Ratio of curtailed wind power with variation of thermal load power **Fig. 6** Mechanism analysis of wind power curtailment

To demonstrate the influence of EWHs on the overall cost, the cost ratio of EWHs to overall cost with different curtailed wind power is shown in Fig. 10. It indicates that the cost ratio of EWHs to overall dispatching cost reduces

with the increase of wind power curtailment. The maximum cost ratio of dispatching EWHs is about 0.2%. This illustrates that it is a relatively economical and feasible scheme in wind power accommodation.

5 Conclusion

This paper addresses the issue of wind power curtailment with an innovative way of dispatching storage-type EWHs. The conclusions are as follow:

- 1) In contrast to SK prediction such as RBF-SKRVM and PK-SKRVM, the proposed MK prediction algorithm combining AFOA and MKRVM has a higher precision of prediction, especially in resolving the sample distribution of multi-source heterogeneous features.
- 2) The parallel computation with a Hadoop cluster accelerates the computation speed of the prediction approach in a similar linear tendency with the number of clusters.

Fig. 8 Comparison of load power and wind power accommodation with and without dispatched EWHs

Table 5 Results comparison with and without dispatched EWHs

Operating mode	Coal consumption (t)	Curtailed wind power (MW)	Accommodation ratio of wind power (%)
Without EWHs	9964.5	350.69	15.38
With EWHs	9962.7	313.09	13.73

- 3) Wind power curtailment frequently occurs at night because of a low power load and a high thermal load; the accommodation ratio of wind power reduces with the increase of thermal load.
- 4) Dispatching of EWHs enables the reduction of the curtailment of wind power and decreases coal consumption; moreover, a regular dispatching of EWHs can smooth the curve of load power relatively and demonstrates better economic feasibility.

Our research ignores the limit of transmission capacity for the grid. However, some studies have shown [36] that the structure of the grid network is related to the power

Fig. 9 Coal consumption and dispatched EWHs power with different wind power accommodations

Fig. 10 Cost ratio of EWHs in overall cost with different wind power accommodations

transmission and the accommodation of wind power. Future work can build a dispatching model in more detail. In addition, the kernel functions used in the paper are the RBF kernel and the polynomial kernel, and this solution is a relatively restricted one. More work will be carried out on the MK prediction approach in the selection of kernel function.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Acknowledgements This work was supported by National Natural Science Foundation of China (No. 51407077), Fundamental Research Funds for the Central Universities of China (No. 2017MS095) and Technology Project of State Grid Corporation of China Headquarter (No. 5204BB16000F).

STATE GRID

STATE GRID ELECTRIC POWER RESEARCH INSTITUTE

References

- [1] Sichilalu S, Mathaba T, Xia X (2017) Optimal control of a wind-PV-hybrid powered heat pump water heater. *Appl Energy* 185:1173–1184
- [2] Cooper SJG, Hammond GP, Mcmanus MC et al (2016) Detailed simulation of electrical demands due to nationwide adoption of heat pumps, taking account of renewable generation and mitigation. *IET Renew Power Gener* 10(3):380–387
- [3] Papaefthymiou G, Hasche B, Nabe C (2012) Potential of heat pumps for demand side management and wind power integration in the German electricity market. *IEEE Trans Sustain Energy* 3(4):636–642
- [4] Zhang N, Lu X, McElroy MB et al (2016) Reducing curtailment of wind electricity in China by employing electric boilers for heat and pumped hydro for energy storage. *Appl Energy* 184:987–994
- [5] Deng JL, Hu LX, Li JJ (2016) Analysis on mechanism of curtailed wind power accommodation and its economic operation based on electric boiler for peak-load regulation at secondary heat supply network. *Autom Electr Power Syst* 40(18):41–46
- [6] Nielsen MG, Morales JM, Zugno M et al (2016) Economic valuation of heat pumps and electric boilers in the danish energy system. *Appl Energy* 167:189–200
- [7] Gonzalez MV, Andersson G (2016) Self scheduling of plug-in electric vehicle aggregator to provide balancing services for wind power. *IEEE Trans Sustain Energy* 7(2):886–899
- [8] Shojaabadi S, Abapour S, Abapour M et al (2016) Simultaneous planning of plug-in hybrid electric vehicle charging stations and wind power generation in distribution networks considering uncertainties. *Renewable Energy* 99:237–252
- [9] Chandrashekar S, Liu Y, Sioshansi R (2017) Wind-integration benefits of controlled plug-in electric vehicle charging. *J Mod Power Syst Clean Energy* 5(5):746–756
- [10] Hayes BP, Wilson A, Webster R et al (2016) Comparison of two energy storage options for optimum balancing of wind farm power outputs. *IET Gener Transm Distrib* 10(3):832–839
- [11] Hajipour E, Bozorg M, Fotuhi-Firuzabad M (2015) Stochastic capacity expansion planning of remote microgrids with wind farms and energy storage. *IEEE Trans Sustain Energy* 6(2):491–498
- [12] Valizadeh HH, Lotfifard S (2015) Spatiotemporal modeling of wind generation for optimal energy storage sizing. *IEEE Trans Sustain Energy* 6(1):113–121
- [13] Broeer T, Fuller J, Tuffner F et al (2014) Modeling framework and validation of a smart grid and demand response system for wind power integration. *Appl Energy* 113:199–207
- [14] He H, Sanandaji BM, Poolla K et al (2014) Aggregate flexibility of thermostatically controlled loads. *IEEE Trans Power Syst* 30(1):189–198
- [15] Wang JX, Zhong HW, Xia Q et al (2016) Model and method of demand response for thermostatically-controlled loads based on cost-benefit analysis. *Autom Electr Power Syst* 40(5):45–53
- [16] Tindemans SH, Trovato V, Strbac G (2015) Decentralized control of thermostatic loads for flexible demand response. *IEEE Trans Control Syst Technol* 23(5):1685–1700
- [17] Xu Z, Diao R, Lu S et al (2014) Modeling of electric water heaters for demand response: a baseline PDE model. *IEEE Trans Smart Grid* 5(5):2203–2210
- [18] Pourmousavi SA, Patrick SN, Nehrir MH (2014) Real-time demand response through aggregate electric water heaters for load shifting and balancing wind generation. *IEEE Trans Smart Grid* 5(2):769–778
- [19] Fitzgerald N, Foley AM, Mckeogh E (2012) Integrating wind power using intelligent electric water heating. *Energy* 48(1):135–143
- [20] Sun YZ, Wu J, Li GJ et al (2009) Dynamic economic dispatch considering wind power penetration based on wind speed forecasting and stochastic programming. *Proc CSEE* 29(4):41–47
- [21] Krishnasamy U, Nanjundappan D (2016) Hybrid weighted probabilistic neural network and biogeography based optimization for dynamic economic dispatch of integrated multiple-fuel and wind power plants. *Int J Electr Power Energy Syst* 77:385–394
- [22] Zhang ZS, Sun YZ, Lin J et al (2012) Innovation for wind power: adding smart processors between prediction systems and dispatch systems. In: *Proceedings of IEEE transmission and distribution conference and exposition, Orlando, USA, 7–10 May 2012*, 4 pp
- [23] Srinivasan D, Tan SS, Cheng CS et al (1999) Parallel neural network-fuzzy expert system strategy for short-term load forecasting: system implementation and performance evaluation. *IEEE Trans Power Syst* 14(3):1100–1106
- [24] Kalaitzakis K, Stavrakakis GS, Anagnostakis EM (2002) Short-term load forecasting based on artificial neural networks parallel implementation. *Electr Power Syst Res* 63(3):185–196
- [25] Kintakis AM, Chrysopoulos A, Mitkas PA (2015) Agent-based short-term load and price forecasting using a parallel implementation of an adaptive PSO-trained local linear wavelet neural network. In: *Proceedings of 12th international conference on the European energy market, Lisbon, Portugal, 19–22 May 2015*, 5 pp
- [26] Rakotomamonjy A, Bach F, Cau S et al (2008) SimpleMKL. *J Mach Learn Res* 9(3):2491–2521
- [27] Zhou Y, Cui X, Hu Q (2015) Improved multi-kernel SVM for multi-modal and imbalanced dialogue act classification. In: *Proceedings of international joint conference on neural networks, Killarney, Ireland, 12–17 July 2015*, 8 pp
- [28] Alam S, Kang M, Pyun JY et al (2016) Performance of classification based on PCA, linear SVM, and multi-kernel SVM. In: *Proceedings of 8th international conference on ubiquitous and future networks, Vienna, Austria, 5–8 July 2016*, pp 987–989
- [29] Tipping ME (2001) Sparse bayesian learning and the relevance vector machine. *J Mach Learn Res* 1(3):211–244
- [30] Bishop CM, Tipping ME (2013) Variational relevance vector machines. *Comput Sci* 28(3):46–53
- [31] Tolambiya A, Kalra PK (2010) Relevance vector machine with adaptive wavelet kernels for efficient image coding. *Neurocomputing* 73(7–9):1417–1424
- [32] Liu X, Zhang XH, Yuan J (2010) Relevance vector machine and fuzzy system based multi-objective dynamic design optimization: a case study. *Expert Syst Appl* 37(5):3598–3604
- [33] Yang F, Sun W, Lin G et al (2016) Prediction of military vehicle's drawbar pull based on an improved relevance vector machine and real vehicle tests. *Sensors* 16(3):1–20
- [34] Blekas K, Likas A (2014) Sparse regression mixture modeling with the multi-kernel relevance vector machine. *Knowl Inf Syst* 39(2):241–264
- [35] Sun Z, Chen S, Qiao L (2014) A general non-local denoising model using multi-kernel-induced measures. *Pattern Recogn* 47(4):1751–1763
- [36] Shu YB, Zhang ZG, Guo JB et al (2017) Study on key factors and solution of renewable energy accommodation. *Proc CSEE* 37(1):1–8
- [37] Huang J, Gui G, Wan Q (2011) Robust location algorithm based weighted least-squares support vector machine (WLS-SVM) for non-linear-of-sight environments. *Int J Phys Sci* 6(25):5897–5905

Yang YU received the M.S. degree in electrical engineering in 2008 from Xi'an Jiaotong University, China and the Ph.D. degree in electrical engineering in 2016 from North China Electric Power University, China. He is currently an associate professor at the Department of Electrical Engineering, North China Electric Power University in Baoding. His research interests include electrical energy storage technology, and optimal dispatch of flexible load.

Zengqiang MI received the M.S. degree in electrical engineering in 1986 from North China Electric Power University, China. He is currently a professor at the Department of Electrical Engineering, North China Electric Power University in Baoding. His research interests include power system operation and control, electrical energy storage technology, and optimal dispatch of flexible load.

Xiaoming ZHENG received the M.S. degree in electrical engineering in 2013 from Yanshan University, China. He is currently pursuing a Ph.D. degree at the Department of Electrical Engineering, North China Electric Power University in Baoding. His research interests include electrical energy storage technology.

Da CHANG received the B.S. degree in electrical engineering in 2015 from Taiyuan University of Technology, China. He is currently pursuing an M.S. degree at the Department of Electrical Engineering, North China Electric Power University in Baoding. His research interests include electrical energy storage technology.

STATE GRID

STATE GRID ELECTRIC POWER RESEARCH INSTITUTE