Article

A Novel Hybrid Short Term Load Forecasting Model Considering the Error of Numerical Weather Prediction

Guowei Cai [1,*], **Wenjin Wang** [1] **and Junhai Lu** [2]

[1] School of Electrical Engineering, Northeast Electric Power University, Jilin 132012, China; wangwenjin0829@126.com
[2] State Grid LiaoNing Electric Power Supply Co. Ltd., Shenyang 110000, China; lujunhai@ln.sgcc.com.cn
[*] Correspondence: caigw@mail.nedu.edu.cn; Tel.: +86-138-4322-5886

Abstract: In order to reduce the effect of numerical weather prediction (NWP) error on short term load forecasting (STLF) and improve the forecasting accuracy, a new hybrid model based on support vector regression (SVR) optimized by an artificial bee colony (ABC) algorithm (ABC-SVR) and seasonal autoregressive integrated moving average (SARIMA) model is proposed. According to the different day types and effect of the NWP error on forecasting prediction, working days and weekends load forecasting models are selected and constructed, respectively. The ABC-SVR method is used to forecast weekends load with large fluctuation, in which the best parameters of SVR are determined by the ABC algorithm. The working days load forecasting model is constructed based on SARIMA modified by ABC-SVR (AS-SARIMA). In the AS-SARIMA model, the ability of SARIMA to respond to exogenous variables is improved and the effect of NWP error on prediction accuracy is reduced more than with ABC-SVR. Contrast experiments are constructed based on International Organization for Standardization (ISO) New England load data. The experimental results show that prediction accuracy of the proposed method is less affected by NWP error and has higher forecasting accuracy than contrasting approaches.

Keywords: short term load forecasting (STLF); support vector regression (SVR); artificial bee colony (ABC); seasonal autoregressive integrated moving average (SARIMA)

---

1. Introduction

With the continuous development of smart grids, short term load forecasting (STLF) results have become the important basis for dynamic pricing in the power market. Electricity prices formulated based on results of STLF lead electricity consumption during the off-peak period, reduce differences between peak and valley loads and ensure the economic operation of the power system. Compared to the traditional power grid, the influence of STLF on the economic performance of the smart grid is more direct [1,2]. References [3,4] showed that a 1% increase of the prediction error would lead to an extra ten million pounds of cost in the UK.

Feature sets are the foundation of constructing STLF models. The historical load time series contain certain load trend and cycle information, it is must be included in feature sets. In addition to historical loads, other exogenous variables such as temperature and day types which affect the accuracy of STLF also should be considered. Particularly, there is a high correlation between temperature and load data, so adding the temperature variables can effectively improve the accuracy of STLF. However, the temperature data used to construct the feature sets is obtained from numerical weather prediction (NWP), where prediction errors exist, so the influence of NWP errors on load forecasting precision should be considered when establishing any forecast model [5,6].

Generally, the existing STLF methods can be divided into time series methods [7–11] and artificial intelligence methods [12–20]. The time series methods are used to establish the load forecasting model with historical load data, which mainly include exponential smoothing [7] and autoregressive integrated moving average (ARIMA) models [8–11]. The ARIMA models have the advantages of modeling simplicity, high computation speed and they are not affected by NWP errors [11]. Therefore, they are suitable for forecasting stable working day loads, but the relationship between load demand and exogenous variables is nonlinear, thus, it is difficult for ARIMA models to precisely forecast load in the scenarios in which the temperature and other exogenous variables change suddenly. Meanwhile, obtaining continuous weekends load series with uniform characteristics is difficult due to the long time interval between different weekends, so the time series method is less effective for the load forecasting during weekends.

Artificial intelligence methods mainly include the artificial neural network (ANN) [12–15] and support vector regression (SVR) [16–20]. ANN has excellent self-adaptive and nonlinear modeling ability. Therefore, ANN is widely used with high prediction accuracy in STLF. However, ANN needs a large number of training samples and easily falls into local optimal solutions [21–24]. Different from the traditional neural network using the empirical risk minimization principle, SVR is based on the principle of structural risk minimization. It has many advantages, such as obtaining global optimal solutions, avoiding the "curse of dimensionality", good generalization ability, and handling small samples [25–28]. In summary, SVR has better performance than ANN for STLF [17,18,25]. However, from the aspect of feature sets, the high accuracy of STLF prediction models based on SVR is largely dependent on the exogenous variables. When there are large errors in the temperature data of the feature sets, the forecasting accuracy of SVR is obviously decreased [4,21].

The selection of parameters and construction of feature sets play a pivotal role in the prediction precision of SVR [18,29,30]. Therefore, many parameter optimization algorithms are used to determine reasonable parameters of the SVR model, such as the genetic algorithm (GA) and particle swarm optimization (PSO). Nevertheless, GA suffers from the weakness of being time consuming and lacking memory function knowledge. PSO falls into local optima easily and its performance is affected by the particle parameters [31–34]. The artificial bee colony (ABC) algorithm is a novel swarm intelligence optimization algorithm [31–37]. The algorithm simulates the foraging behavior of bee colonies. Different types of bees play distinct roles in the foraging process. By collecting and sharing the food source, the bees find the best food source. The ABC algorithm resolves the conflict between expanding new solution space and searching exactly in the old solution space through cooperation between different types of bees. Therefore, compared with GA and PSO, the ABC algorithm overcomes the problem of local optimization, and has better performance [31–33].

To enhance the responsiveness of the time series models to exogenous variables and reduce the influence of the NWP error on the load forecasting results, a hybrid STLF model based on improved seasonal ARIMA (SARIMA) and SVR is proposed in this paper. Firstly, forecasting results of the methods in scenarios of different day types and NWP errors are analyzed. Secondly, various load forecasting models are constructed for different day types based on the characters of loads and predictors. The ABC-SVR model is constructed for forecasting load on weekends. SARIMA modified by ABC-SVR (AS-SARIMA) is constructed for forecasting load on working days. Finally, the International Organization for Standardization (ISO) New England data are used for comparative experiments to demonstrate the superiority of the proposed method in STLF.

2. Methodology

This section is used to introduce the principles of the methods used in the paper. The principle of the SARIMA model and the experiments using SARIMA for STLF are described in Section 2.1. The theory of the SVR model and the analysis of the impact of SVR parameters selection on prediction accuracy are presented in Section 2.2. The use of the ABC algorithm to determine the SVR parameters is introduced in Section 2.3.

2.1. Seasonal Autoregressive Integrated Moving Average

As a time series model, the SARIMA originates from the autoregressive moving average (ARMA) [8]. SARIMA is used to forecast periodic load series. Assuming that there is a nonstationary time series $\{x_t | t = 0, 1, \ldots, k\}$, a general SARIMA (p, d, q) $(P, D, Q)_s$ can be expressed as follows:

$$\phi(B)\Phi(B^s)\nabla^d(1 - B^s)^D x_t = \theta(B)\Theta(B^s)e_t \tag{1}$$

where x_t and e_t are the actual value and rand error at time period t respectively; p and q are corresponding orders of non-seasonal autoregressive polynomial $\phi(B)$ and moving average polynomial $\theta(B)$; d is the order of regular differences; P and Q are corresponding orders of seasonal autoregressive polynomial $\Phi(B^s)$ and moving average polynomial $\Theta(B^s)$; s is the period; D is the order of seasonal differences; B is a backshift operator, and satisfies $Bx_t = x_{t-1}$; ∇ is a differencing operator, and satisfies $\nabla = 1 - B$. It is assumed that e_t are independent and identically distributed random errors with mean of zero and variance of σ^2.

$\phi(B)$ and $\theta(B)$ can be described as follows:

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p \tag{2}$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q \tag{3}$$

$\Phi(B^s)$ and $\Theta(B^s)$ can be described as follows:

$$\Phi(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \cdots - \Phi_P B^{Ps} \tag{4}$$

$$\Theta(B^s) = 1 - \Theta_1 B^s - \Theta_2 B^{2s} - \cdots - \Theta_Q B^{Qs} \tag{5}$$

The steps of constructing the SARIMA $(p, d, q)(P, D, Q)_s$ are described as follows [38]:

(1) Get the period by analyzing the autocorrelation function (ACF), and obtain a new stationary time series by difference which eliminated the tendency and periodicity of the original series.
(2) Model identification: Achieve all reasonable combinations of p, q, P and Q through analyzing the ACF and partial autocorrelation function (PACF). Then, the primary model is determined by Akaike information criterion (AIC).
(3) Parameter estimation: Estimate the parameters of the model by means of maximum likelihood.
(4) Diagnostic checking: Decide whether the model is reasonable by residuals analysis. If the model is reasonable, it is determined as the final prediction model. Otherwise, repeat Steps 2–4.

In order to verify the forecasting accuracy of SARIMA, the SARIMA model is used to forecast the load from 30 January to 5 February 2012. The load data are obtained from ISO New England [5]. The forecast results are shown in Figure 1. Figure 1 shows that there are large errors in the experimental results of forecasting load on Monday and Saturday. Without consideration of the influence of the exogenous variables (such as temperature and day types) on load forecasting accuracy, the input of the SARIMA model only includes the historical load data. Therefore, there are obvious errors in the Monday and Saturday load forecasting results. This indicates that, when the exogenous variables change greatly (such as the data changes from workdays to holidays), the SARIMA model has poor load forecasting performance.
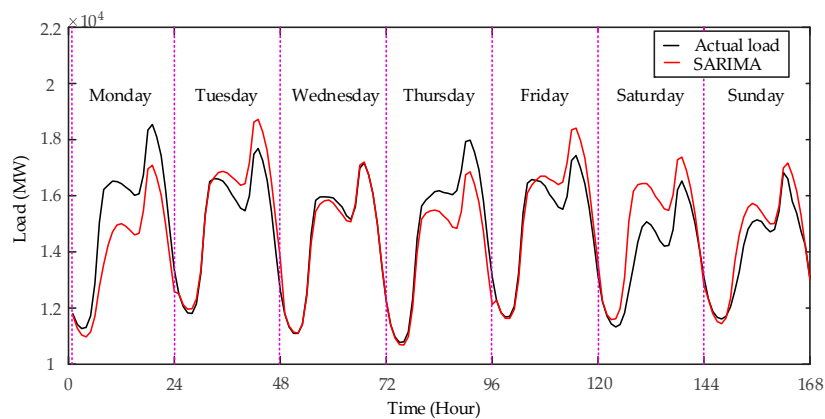
Figure 1. Curves of original load and load predicted by seasonal autoregressive integrated moving average (SARIMA).

2.2. Support Vector Regression

SVR is a novel regression model based on statistical learning theory. Compared to ANN, SVR improves the generalization capability and avoid falling to local optima. It has been proved that SVR has higher accuracy than ANN [25]. The theory of SVR is described as follows [18,23,30].

Given a training data set $\{(x_i, y_i), i = 1, 2, \ldots, n\}$, where n is the number of samples, x_i is the input vector, $x_i = [x_i^1, x_i^2, \ldots, x_i^d]^T$, d is the dimension of input vector, y_i is the corresponding output value. The nonlinear mapping function $H(x)$ is introduced to map the input space to the high dimensional feature space. Linear regression function is as Equation (6):

$$f(x) = wH(x) + b \tag{6}$$

where $f(x)$ represents the predicted value, w and b are weight vector and bias respectively.

ε-insensitive loss function is defined as Equation (7):

$$f_L(f(x), y, \varepsilon) = \begin{cases} 0, & |y - f(x)| \leq \varepsilon \\ |y - f(x)| - \varepsilon, & |y - f(x)| > \varepsilon \end{cases} \tag{7}$$

where $f_L(f(x), y, \varepsilon)$ is used to find an optimal hyperplane which can be used to divide the training samples into two subsets while the distance is maximized.

The objective function with the constraints is:

$$\begin{cases} \min \frac{1}{2} \; ||w||^2 + C \sum_{i=1}^{n} (\xi_i, \xi_i^*) \\ \text{s.t.} \begin{cases} y_i - wH(x_i) - b \leq \varepsilon + \xi_i \\ -y_i + wH(x_i) + b \leq \varepsilon + \xi_i^* \\ \xi_i \geq 0, \xi_i^* \geq 0 \end{cases}, i = 1, 2, \ldots n \end{cases} \tag{8}$$

where C is a parameter which trade off the empirical risk and regression function flatness, ξ_i and ξ_i^* are slack variables.

By introducing Lagrangian multiplies, Equation (8) can be described as follows:

$$\begin{cases} \max_{\beta, \beta^*} [-\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} (\beta_i - \beta_i^*)(\beta_j - \beta_j^*) K(x_i, x_j) - \sum_{i=1}^{n} (\beta_i + \beta_i^*)\varepsilon + \sum_{i=1}^{n} (\beta_i - \beta_i^*)y_i] \\ \text{s.t.} \begin{cases} \sum_{i=1}^{n} (\beta_i - \beta_i^*) = 0 \\ 0 \leq \beta_i, \beta_i^* \leq C \end{cases} \end{cases} \tag{9}$$

where $K(x_i, x_j)$ is the kernel function, satisfies $K(x_i, x_j) = H(x_i)H(x_j)$, β_i and β_i^* are Lagrangian multipliers. The regression function can be written as Equation (10):

$$f(x) = \sum_{i=1}^{n} (\beta_i - \beta_i^*)K(x_i, x) + b \tag{10}$$

The radial basis function (RBF) is easy to implement and has good ability to deal with the complex nonlinear relationships between the input and output vector of the samples [17,18,29,30]. Therefore, RBF is selected as kernel function of the SVR in this paper. The RBF is shown as Equation (11):

$$K(x, x_i) = \exp\left(-\frac{||x - x_i||^2}{2\sigma^2}\right) \tag{11}$$

where σ is the width of RBF.

The selection of parameters has a great influence on the prediction accuracy of the SVR. The parameter C is used to trade off the training error and model complexity. If C is too large, weak generalization ability and overfitting phenomena may appear. The parameter ε determines the number of support vectors of model. If ε is too large, there will be too few support vectors and the model will be too simple. If σ is too large, the RBF kernel will approximate the use of a linear kernel. Thus, the complexity and generalization ability of the SVR model is mainly determined by the selection of the parameters. It is needed to choose an appropriate parameter optimization algorithm to select the SVR parameters, so as to improve the prediction accuracy of SVR.

2.3. Artificial Bee Colony Algorithm

The ABC algorithm is an innovative kind of optimization method, which is applied to solve the real world problems by simulating the foraging behavior of bees [5,34]. The algorithm has many advantages such as simple operation, few parameters, robustness and avoiding local optimization. The ABC algorithm is used to select the parameters of SVR.

The bees can be classified into three groups in ABC algorithm: worker bees, onlooker bees and scout bees; the worker bees search for food sources; and they pass the information about nectar amounts to onlooker bees. The onlooker bees select the food source based on the information obtained from the worker bees, and further explore the nectar source. A food source position represents a solution of the optimization problem. The amount of nectar denotes the fitness value of a solution. If a worker bee abandons its food source, it will become a scout bee to search for a new food source. The initial positions are generated by Equation (12):

$$z_{ij} = z_{\min,j} + \text{rand}(0, 1)(z_{\max,j} - z_{\min,j}) \tag{12}$$

where $z_{\min,j}$ and $z_{\max,j}$ are corresponding boundary values for dimension index $j(j = 1, 2, \ldots, D)$, D is the dimension of food source position $z_i(i = 1, 2, \ldots, FN)$, FN is the number of food sources.

Then, the worker bee finds a new food source v_i in the neighborhood of z_i by Equation (13):

$$v_{ij} = z_{ij} + \varphi_{ij}(z_{ij} - z_{kj}) \tag{13}$$

where φ_{ij} is a random number in the range [–1, 1]. k is a random index in the range [1, *FN*], and satisfies $k \neq i$. If the fitness value of v_i is superior than that of z_i, the employ bee will replace z_i by v_i.

After obtaining the food source information shared by worker bees, an onlooker bee will select a food source to search. p_i is defined as the probability that the *i*th food source is selected by an onlooker bee:

$$p_i = \frac{fitness_i}{\sum\limits_{l=1}^{FN} fitness_l} \tag{14}$$

where $fitness_i$ is the nectar amount of the *i*th food source. As shown in Equation (14), the more nectar of the *i*th food source, the the higher probability that the food source is selected.

If a food source position has not been updated after *limit* cycles, the worker bee will abandon the food source and start to find a new one. The flowchart of the ABC algorithm is shown in Figure 2.
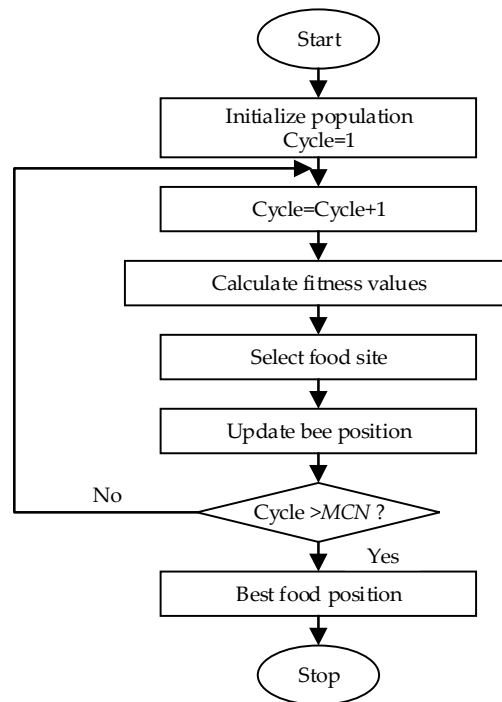

Figure 2. The flowchart of the artificial bee colony (ABC) algorithm. *MCN*: maximum cycle number.

3. The Proposed Short Term Load Forecasting Method

To construct the appropriate features as the input of the prediction model, the characteristics of the load data and the influence of the exogenous variables on load are analyzed in Section 3.1. The prediction accuracy of SARIMA and SVR under the circumstance of actual temperature and noisy temperature are compared in Section 3.2. To increase the prediction accuracy of SVR and SARIMA, the improved methods based on the two models are proposed in Sections 3.3 and 3.4, respectively. By comparing the performance of the improved models in the scenarios of different day types and NWP errors in Section 3.5, a new hybrid STLF model based on combing the advantages of the models is proposed in Section 3.6.

3.1. Feature Set Construction

The prediction accuracy of the SVR is highly dependent on the selection of its input variables. Therefore, it is necessary to construct a reasonable feature set as input of SVR. Considering the load data characteristics and the impact of exogenous variables such as temperature, day of week and time index on the load, the feature set is determined by the following analysis steps. The experiments of the paper are carried out on the basis of the hourly load data of ISO New England [5]. The load curve from 1 January to 1 March 2011 is shown in Figure 3. The loads during working days and weekends are separated by red dashed lines [39].
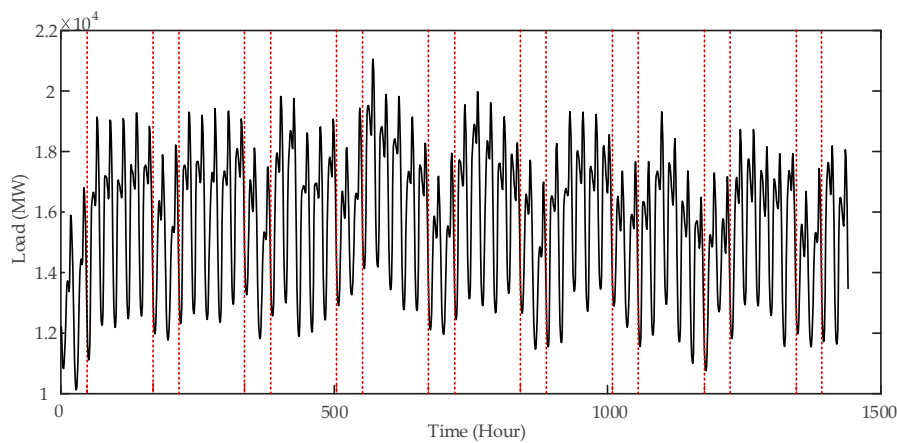
Figure 3. Load from 1 January to 1 March 2011.

(1)   Historical load. Figure 3 shows that the historical load data has the following characteristics: firstly, the load time series takes 24 h as a cycle. Secondly, the neighboring curve is similar and the load values on the different curve are close to each other at the same time. Finally, there is obvious difference between the load values during the weekends and working days. The load demands during weekends are less than working days. $L(t, d)$ is defined as the load value at the time $t(t = 1, 2, \ldots, 24)$ of forecasted day. When doing day-ahead load forecasting, the load value at every hour of forecasted day is unknown. The results of the selection of historical load features are described as follows: load at time t and $t - 1$ of the previous day, $L(t, d - 1)$ and $L(t - 1, d - 1)$, load at t of the seven days before forecasted day $L(t, d - 7)$, maximum and average load of the previous day, $L_{\max}(d - 1)$ and $L_{\mean}(d - 1)$, load at 24 of the previous day $L(24, d - 1)$ [40].

(2)   Day of week. Numbers from 1 to 7 represent Monday to Sunday, respectively.

(3)   Day type. Numbers 1 and 0 represent working days and weekends, respectively (the midweek holidays are identified as number 1).

(4)   Time index. The period of load time series is 24 h. Therefore, $T_{\sin}(t)$ and $T_{\cos}(t)$ [5] are defined as time variables to capture cycles. $T_{\sin}(t)$ and $T_{\cos}(t)$ can be calculated by Equations (15) and (16):

$$T_{\sin}(t) = \sin(2\pi t/24) \tag{15}$$

$$T_{\cos}(t) = \cos(2\pi t/24) \tag{16}$$

(5)   Temperature. The relationship between load and temperature in 2011 is shown in Figure 4. Figure 4 shows that load values are greatly influenced by temperature. The load values increase when the temperature is lower than 40 F or higher than 60 F. The temperature variables are selected as follows: the temperature at time t of the forecasted day $T(t)$ and the previous day $T(t, d - 1)$, the maximum and minimum temperature of the previous day, $T_{\max}(d - 1)$ and $T_{\min}(d - 1)$. Besides, the response time of load demand to temperature changes is larger than that of the sampling period. Therefore the average temperature of the past 3 h ($T_{av}(3)$), 6 h ($T_{av}(6)$), and 24 h ($T_{av}(24)$) are selected as temperature variables [40].
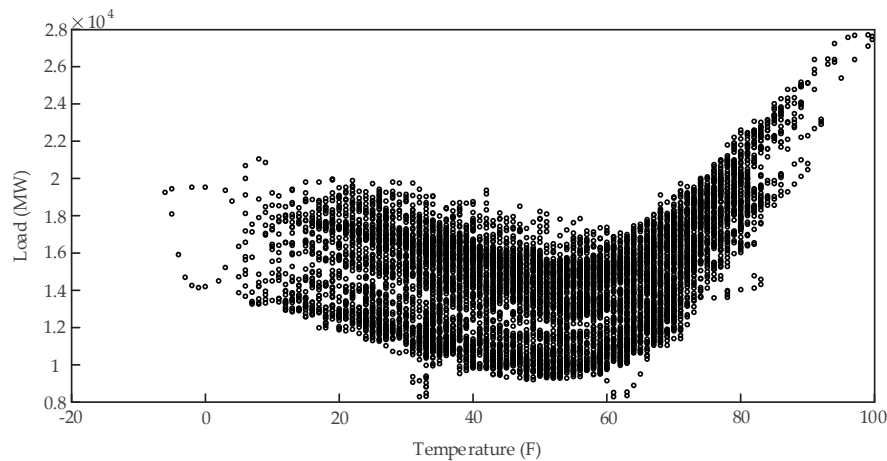
Figure 4. Relationship between load and temperature.

Table 1 lists the composition of the feature set.

Table 1. The feature set.

Type Number	Variables	Features
1	Historical load	$L(t, d-1)$, $L(t-1, d-1)$, $L(t, d-7)$, $L_{max}(d-1)$, $L_{mean}(d-1)$, $L(24, d-1)$
2	Day of week	Numbers from 1 to 7
3	Day type	1 (working days), 0 (weekends)
4	Time index	$T_{sin}(t)$, $T_{cos}(t)$
5	Temperature	$T(t)$, $T(t, d-1)$, $T_{max}(d-1)$, $T_{min}(d-1)$, $T_{av}(3)$, $T_{av}(6)$, $T_{av}(24)$

3.2. Comparison of Load Forecasting Accuracy between SARIMA and SVR

In order to analyze the influence of NWP error on the load forecasting accuracy, noisy temperature data is simulated by adding Gaussian noise to the actual temperature data. The mean value of the Gaussian noise is zero and the standard deviation is 0.6 °C [5,6]. SVR [41] and SARIMA are used to forecast the load obtained from ISO New England [5] from 6 to 12 February 2012. The mean absolute percentage error (*MAPE*) is adopted as criterion of error evaluation. The *MAPE* can be expressed as:

$$MAPE = \frac{1}{N} \sum_{t=1}^{N} \left| \frac{L(t) - \hat{L}(t)}{L(t)} \right| \times 100\% \tag{17}$$

where $L(t)$ is the actual load value, and $\hat{L}(t)$ is the forecasting value, N is the number of samples. The results of prediction are listed in Table 2 (AT denotes the actual temperature and NT denotes a noisy temperature).

Table 2. Mean absolute percentage error (*MAPE*) (%) of support vector regression (SVR) and SARIMA.

Day	SARIMA		SVR	
	AT	**NT**	**AT**	**NT**
Monday	5.74	5.74	3.13	3.17
Tuesday	0.98	0.98	1.35	1.44
Wednesday	2.84	2.84	2.13	2.13
Thursday	2.12	2.12	2.51	2.54
Friday	2.11	2.11	2.37	2.41
Saturday	4.50	4.50	2.82	2.75
Sunday	2.81	2.81	5.23	5.34
Average *MAPE*	3.01	3.01	2.79	2.83

Table 2 indicates that the whole prediction accuracy of SVR is higher than that of SARIMA in working days and weekends. However, the *MAPE* of the SVR increases significantly when the Gaussian noise is added to the actual temperature. Without considering the impact of variables such as day types and temperature on the load, the SARIMA model is established based on the load data, so the performance of SARIMA in STLF is not good. The forecasting accuracy of SARIMA is not affected by NWP error at the same time.

From the above analysis, we can further improve the prediction efficiency of SVR by optimizing the parameters, and modify the results of SARIMA by forecasting residuals. When predicting the residuals from SARIMA, the exogenous variables such as day types and temperature will be considered.

3.3. ABC Algorithm for Parameters Selection of SVR

From the analysis in the Section 2.2, it can be known that the forecast accuracy of SVR is largely dependent on the selection of parameters including C, σ and ε. To improve the performance of SVR in STLF, the ABC algorithm is used to determine the SVR parameters in this paper. The specific steps of constructing the ABC-SVR model are as follows:

(1) Initialize the parameters of ABC algorithm such as population of bees, maximum cycle number (*MCN*), abandonment cycle number (*limit*). The *i*th solution $z_i(i = 1, 2, \ldots, FN)$ of the algorithm is a vector with three elements including C, σ and ε:

$$z_i = (C_i, \sigma_i, \varepsilon_i) \tag{18}$$

where *FN* is the number of solution, C, σ and ε are within the range $[2^{-8}, 2^8]$.

(2) A worker bee finds a new solution in the neighborhood of the present solution, and calculates the fitness values of the two solutions. The fitness value is calculated by Equation (19):

$$fitness_i = \frac{1}{1 + e_i} \tag{19}$$

where e_i is the mean square error (MSE) of the SVR model, in which the elements of z_i are chosen as SVR parameters. MSE is defined as:

$$e_i = \frac{1}{N} \sum_{t=1}^{N} \left(L(t) - \hat{L}(t) \right)^2 \tag{20}$$

where $L(t)$ is the actual load value, and $\hat{L}(t)$ is the forecasting value, N is the number of samples.

(3) The onlooker bee selects a solution by calculating the probability of the solution by Equation (14), and updates the information of the present solution.

(4) When the number of cycles satisfies the abandonment criteria (*limit*), a new solution will be generated by Equation (12).

(5) Repeat Steps 2–4, until the number of cycles is equal to *MCN*.

(6) The elements of best solution are determined as parameters of SVR.

According to the above steps, the ABC-SVR model is constructed. Then, SVR and ABC-SVR are used to forecast the load from 20 to 26 February in the scenario of actual temperature and noisy temperature. The load data is obtained from ISO New England [5] and the noisy temperature is simulated by adding Gaussian noise of zero mean and standard deviation of 0.6 °C to actual temperature. The forecast results are shown in Figure 5.
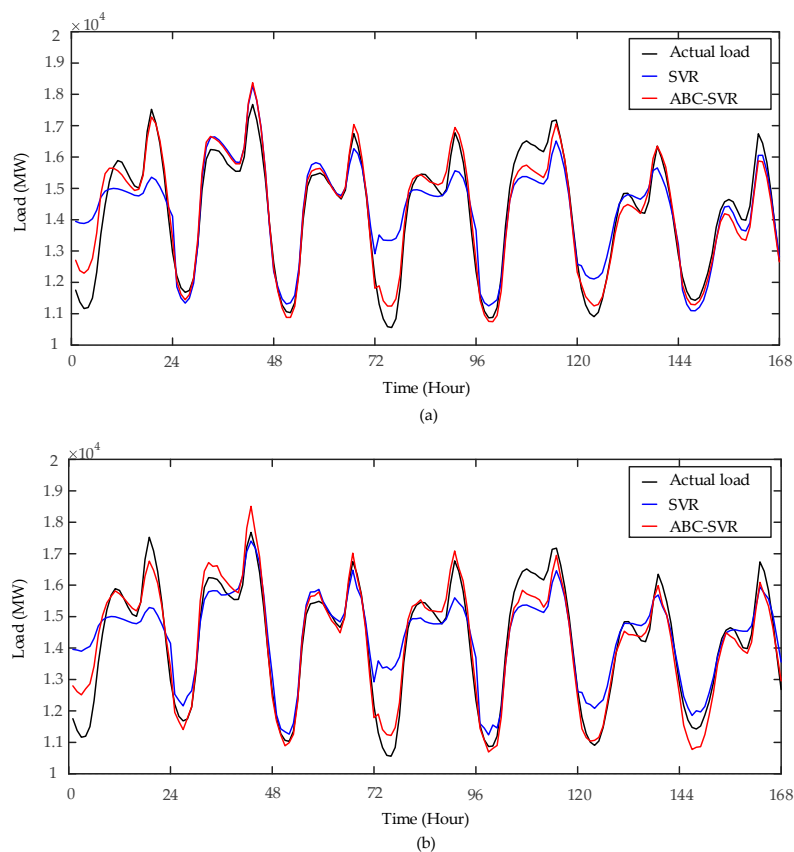
Figure 5. Load curves predicted by SVR and ABC-SVR. (**a**) Actual temperature; and (**b**) noisy temperature.

Figure 5a shows that, the accuracy of SVR is obviously improved by using ABC algorithm to optimize the SVR parameters in the actual temperature scenario. The average *MAPE* of SVR and ABC-SVR are 4.59% and 2.43%, respectively. Figure 5b shows that the prediction accuracy of ABC-SVR is still higher than that of SVR when the noisy temperature data is used as the temperature variables of models. The average *MAPE* of SVR and ABC-SVR are 4.67% and 2.51%, respectively. Compared with the results achieved in scenario of actual temperature, the forecast accuracy of SVR and ABC-SVR are all decreased. The average *MAPE* of ABC-SVR increases from 2.52% to 2.59% during working days and increases from 2.20% to 2.31% during weekends. The increase of the forecast error in ABC-SVR is less than that in SVR. Therefore, ABC-SVR can effectively improve the prediction accuracy of SVR, but its forecast accuracy is still affected by the NWP error.

3.4. Modifying Results of SARIMA by Forecasting Residuals Using ABC-SVR

To improve the forecast accuracy of SARIMA, it is necessary to enhance the response ability of the model responses to exogenous variables. The AS-SARIMA models are constructed by using ABC-SVR to modify the results of the SARIMA. The steps of building the AS-SARIMA models are described as follows.

(1)　Establish reasonable SARIMA models for STLF, then obtain the historical residuals and load values of forecasted day from SARIMA.

(2)　Build the ABC-SVR models to forecast the residuals achieved from Step 1, then obtain the residuals of the forecasted day. The same variables including day of week, day type, time index and temperature variables as described in Section 3.1 are selected as the inputs of the model. Particularly, the historical load variables are replaced by historical residuals variables when construct the feature set of ABC-SVR.

(3)　By adding the output values of ABC-SVR to the output values of SARIMA, the final load forecasting values are obtained.

　　SARIMA and AS-SARIMA models are used to forecast the load from 20 to 26 February 2012 in the scenarios of actual temperature and noisy temperature. The Gaussian noise of zero mean and standard deviation of 0.6 °C is added to the measured temperature to simulated noisy temperature. The above data are achieved from ISO New England [5]. The load forecast results are shown in Figure 6. Figure 6a shows that, the forecast accuracy of AS-SARIMA is significantly higher than that of SARIMA in the scenario of actual temperature. The *MAPE* of SARIMA and AS-SARIMA are 4.03% and 2.34%, respectively. Figure 6b shows that by considering the temperature variables, the *MAPE* of AS-SARIMA increases from 2.34% to 2.36%, but the overall prediction accuracy of AS-SARIMA is still higher than that of SARIMA. Therefore, although the prediction accuracy of the AS-SARIMA is affected by the NWP error, it is still superior to the forecast accuracy of the SARIMA.
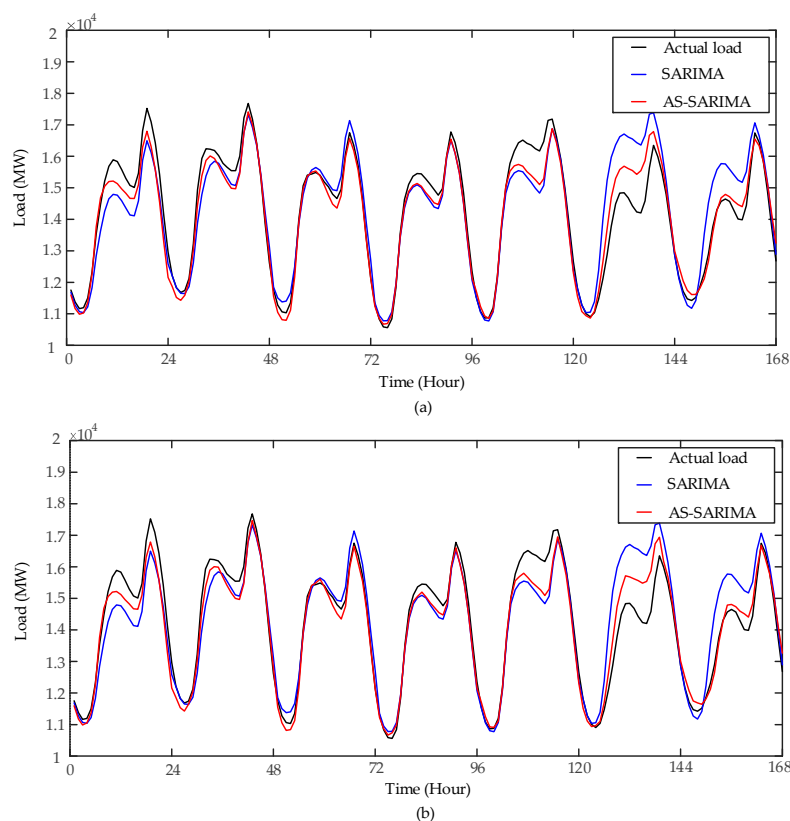
Figure 6. Load curves predicted by SARIMA and AS-SARIMA. (**a**) Actual temperature; and (**b**) noisy temperature.

3.5. Comparison of Forecast Accuracy of ABC-SVR and AS-SARIMA and Construction of the Proposed Method

　　In order to compare the prediction accuracy of optimal approaches and analyze the forecasting performance affected by NWP errors of the two models, ABC-SVR and AS-SARIMA are used to forecast the load from 20 to 26 February in the scenarios of actual temperature and noisy temperature. The Gaussian noise of zero mean and standard deviation of 0.6 °C is added to actual temperature. The forecast curves of ABC-SVR and AS-SARIMA are shown in Figure 7, and the *MAPE* of the two models are listed in Table 3 (where AT denotes actual temperature, NT denotes noisy temperature). In Table 3, shadowed areas are results of forecasting load in working days and weekends correspondingly generated by AS-SARIMA and ABC-SVR.
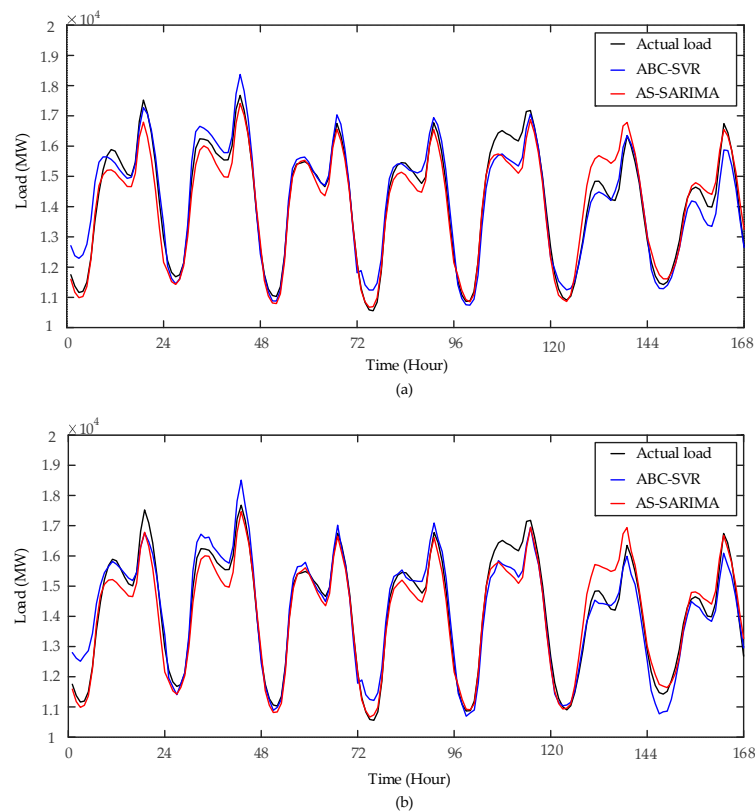
Figure 7. Load curves predicted by ABC-SVR and AS-SARIMA. (**a**) Actual temperature; and (**b**) noisy temperature.

Table 3. *MAPE* (%) of ABC-SVR and AS-SARIMA.

| Day	ABC-SVR		AS-SARIMA	
	AT	NT	AT	NT
Monday	4.14	4.22	3.05	3.04
Tuesday	2.07	2.12	2.05	2.01
Wednesday	0.97	1.03	1.14	1.13
Thursday	2.62	2.68	1.56	1.57
Friday	2.79	2.90	2.75	2.76
Saturday	1.58	1.70	4.09	4.15
Sunday	2.82	2.92	1.75	1.84
Average *MAPE*	2.43	2.51	2.34	2.36

Figure 7 and Table 3 show that, when forecasting the load during weekends in the scenario of actual temperature, the average *MAPE* generated by ABC-SVR and AS-SARIMA are 2.20% and 2.92%, respectively. With considering the NWP errors, the average of the two models are 2.31% and 3.00%. Therefore, the ABC-SVR has higher prediction accuracy than AS-SARIMA for weekend load forecasting. When forecasting the load during working days, the average *MAPE* of ABC-SVR and AS-SARIMA are 2.52% and 2.11% in the actual temperature scenario. After considering the NWP errors, the average *MAPE* of the two models are 2.59% and 2.10%. Therefore, the forecasting performance of AS-SARIMA is better than that of ABC-SVR during working days.

3.6. The Establishment of the Proposed Method

According to the above analysis, it is concluded that the ABC-SVR model is suitable for forecasting load in weekends and the AS-SARIMA model is suitable for forecasting the load in the working days,

so a novel hybrid forecast method can be constructed by using ABC-SVR and AS-SARIMA to forecast the weekends and working days load values, respectively. The efficiency of the proposed method in the scenarios of actual temperature and noisy temperature with the Gaussian noise of standard deviations of 0.6 °C has been preliminarily proved through the above comparative experiments.

When forecasting the load in working days, the load data of first 20 working days (without weekends) are used to construct historical working days load series. By build the SARIMA models, the residuals of the first 20 working days are achieved. Then, the ABC-SVR model is established to forecast the residuals from SARIMA. Finally, the predicted load in working days is obtained by using ABC-SVR to modify the results from SARIMA. When ABC-SVR models are constructed to forecast the load during weekends, data of the first 20 days (including working days and weekends) are used as training samples. The flowchart of the proposed method is shown in Figure 8.
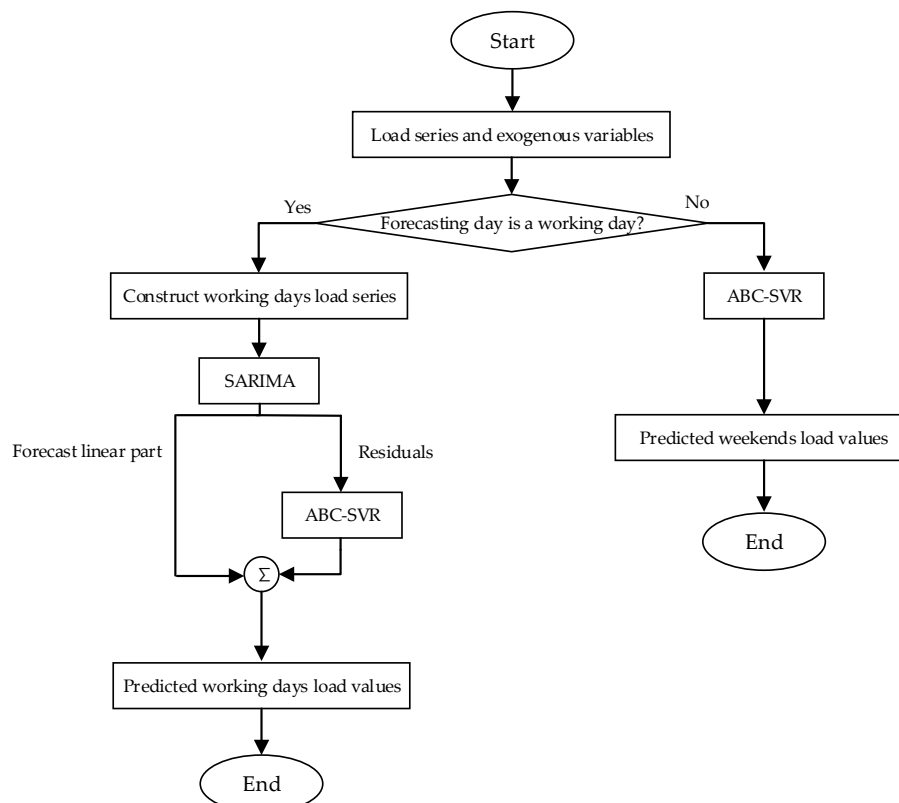

Figure 8. The flowchart of the proposed method.

4. Experimental Results and Analysis

To verify the validity of proposed model, experiments using the data from 1 December 2011 to 31 December 2012 are performed.

4.1. Forecasting Results of the Proposed Method

The data of the first 20 days are selected as training samples and the load values in the next day are selected as testing samples. In order to analyze the influence of the NWP errors on prediction results, the Gaussian noise of zero mean and standard deviation of 0.6 °C is added to actual temperature to simulate the forecasting temperature data. The proposed method is used to forecast the load in four weeks of the four seasons. The experimental results are shown in Figure 9 and the *MAPE* of the proposed method are presented in Table 4 (where AT denotes the actual temperature and NT denotes a noisy temperature). Figure 9 and Table 4 show that the prediction accuracy of the proposed is high and less affected by the NWP errors.
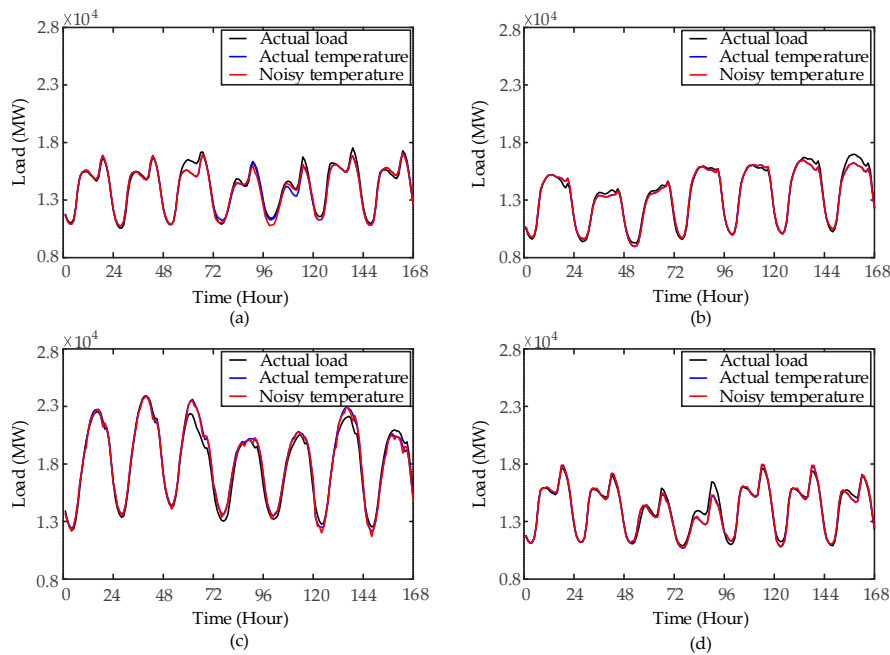
Figure 9. Load curves predicted by the proposed method in 2012. (**a**) Prediction result from 22 to 28 February; (**b**) prediction result from 18 to 24 May; (**c**) prediction result from 8 to 14 August; and (**d**) prediction result from 15 to 21 November.

Table 4. *MAPE* (%) of the proposed method.

Day	Winter 22–28 February		Spring 18–24 May	
	AT	NT	AT	NT
Monday	2.13	2.17	1.16	1.12
Tuesday	1.35	1.35	0.75	0.77
Wednesday	0.78	0.79	1.48	1.51
Thursday	1.17	1.19	2.80	2.84
Friday	2.73	2.76	1.14	1.14
Saturday	1.58	1.70	1.55	1.64
Sunday	2.82	2.92	1.80	1.84
Average *MAPE*	1.79	1.84	1.53	1.55
Day	Summer 8–14 August		Fall 15–21 November	
	AT	NT	AT	NT
Monday	3.11	3.24	1.33	1.36
Tuesday	2.54	2.89	1.53	1.53
Wednesday	1.08	1.12	1.60	1.60
Thursday	1.13	1.19	0.73	0.76
Friday	3.52	3.29	0.62	0.63
Saturday	3.54	3.34	1.89	1.99
Sunday	2.68	2.42	4.08	4.07
Average *MAPE*	2.51	2.50	1.68	1.70

4.2. Comparison and Discussion

4.2.1. Comparison of Forecast Accuracy of Different Models

In order to fully verify the effectiveness of the proposed method, ABC-SVR_WT (without considering the temperature variables, the features including load variables, day of week, day type and time index as described in Section 3.1 are selected as inputs of ABC-SVR_WT), ABC-SVR (the complete feature variables as described in Section 3.1 are selected as inputs of ABC-SVR) and AS-SARIMA are

used to forecast the load in four weeks of four seasons in the scenarios of actual temperature and noisy temperature. The data from ISO New England [5] are used and the Gaussian noise of zero mean and standard deviation of 0.6 °C is added to actual temperature to simulate the noisy temperature. The MAPE of the four models are listed in Tables 5–8 (AT denotes actual temperature, NT denotes noisy temperature).

Table 5. Comparison of *MAPE* (%) from 22 to 28 February 2012.

Day	ABC-SVR_WT		ABC-SVR		AS-SARIMA		Proposed Method	
	AT	NT	AT	NT	AT	NT	AT	NT
22 February	5.25	5.25	0.97	1.03	1.14	1.13	0.78	0.79
23 February	2.59	2.59	2.62	2.68	1.56	1.57	1.17	1.19
24 February	4.60	4.60	2.79	2.90	2.75	2.76	2.73	2.76
25 February	5.34	5.34	1.58	1.70	4.09	4.15	1.58	1.70
26 February	4.78	4.78	2.82	2.92	1.75	1.84	2.82	2.92
27 February	6.07	6.07	5.46	5.52	2.92	2.88	2.13	2.17
28 February	3.41	3.41	2.10	2.15	1.91	1.91	1.35	1.35
Average *MAPE*	4.58	4.58	2.62	2.70	2.30	2.32	1.79	1.84

Table 6. Comparison of *MAPE* (%) from 18 to 24 May 2012.

Day	ABC-SVR_WT		ABC-SVR		AS-SARIMA		Proposed Method	
	AT	NT	AT	NT	AT	NT	AT	NT
18 May	2.95	2.95	2.05	2.09	1.27	1.31	1.14	1.14
19 May	1.99	1.99	1.55	1.64	1.92	1.99	1.55	1.64
20 May	2.56	2.56	1.80	1.84	1.97	1.99	1.80	1.84
21 May	2.40	2.40	1.67	1.72	1.21	1.22	1.16	1.12
22 May	5.99	5.99	1.41	1.46	1.31	1.29	0.75	0.77
23 May	2.82	2.82	1.95	2.01	1.69	1.70	1.48	1.51
24 May	3.89	3.89	3.66	3.77	2.85	2.88	2.80	2.84
Average *MAPE*	3.23	3.23	2.01	2.08	1.75	1.77	1.53	1.55

Table 7. Comparison of *MAPE* (%) from 8 to 14 August 2012.

Day	ABC-SVR_WT		ABC-SVR		AS-SARIMA		Proposed Method	
	AT	NT	AT	NT	AT	NT	AT	NT
8 August	5.15	5.15	2.39	2.44	1.49	1.52	1.08	1.12
9 August	6.50	6.50	1.48	1.52	1.30	1.37	1.13	1.19
10 August	4.82	4.82	4.04	4.11	4.90	4.91	3.52	3.29
11 August	4.03	4.03	3.54	3.34	3.63	3.62	3.54	3.34
12 August	3.39	3.39	2.68	2.42	2.81	2.54	2.68	2.42
13 August	3.36	3.36	2.89	2.91	2.69	2.66	3.11	3.24
14 August	3.83	3.83	2.98	3.40	1.96	2.29	2.54	2.89
Average *MAPE*	4.44	4.44	2.86	2.88	2.68	2.70	2.51	2.50

Table 8. Comparison of *MAPE* (%) from 15 to 21 November 2012.

Day	ABC-SVR_WT		ABC-SVR		AS-SARIMA		Proposed Method	
	AT	NT	AT	NT	AT	NT	AT	NT
15 November	4.39	4.39	2.82	2.86	1.34	1.37	0.73	0.76
16 November	2.00	2.00	1.52	1.54	0.93	0.94	0.62	0.63
17 November	2.49	2.49	1.89	1.99	4.81	4.89	1.89	1.99
18 November	5.88	5.88	4.08	4.07	2.27	2.26	4.08	4.07
19 November	5.24	5.24	4.78	4.81	4.74	4.73	1.33	1.36
20 November	5.41	5.41	1.90	1.98	1.60	1.61	1.53	1.53
21 November	1.85	1.85	1.58	1.66	1.42	1.41	1.60	1.59
Average *MAPE*	3.89	3.89	2.65	2.70	2.44	2.46	1.68	1.70

Tables 5–8 show that, without considering the NWP error, the average *MAPE* of the four weeks generated by ABC-SVR_WT, ABC-SVR, AS-SARIMA and the proposed method is 4.04%, 2.54%, 2.29% and 1.88%, respectively. The average *MAPE* of ABC-SVR_WT is larger than that of ABC-SVR. After adding the Gaussian noise to the actual temperature data, the *MAPE* of ABC-SVR_WT is still 4.04% and the corresponding *MAPE*s of ABC-SVR, AS-SARIMA and the proposed method are 2.59%, 2.31% and 1.90%, respectively. The error of ABC-SVR_WT is still greater than that of ABC-SVR. Therefore, it is necessary to add the temperature variables to the feature set as inputs of the models. Meanwhile, compared with ABC-SVR_WT, ABC-SVR and AS-SARIMA, the proposed method has the best performance in different scenarios.

4.2.2. Comparison of Experimental Results with Different Numerical Weather Prediction Errors

To further prove the efficiency of the proposed method with various temperature errors, the Gaussian noises of zero mean and different standard deviations are added to the actual temperature data in this case. The standard deviations of 0.6 °C, 0.9 °C and 1.2 °C are selected. The ABC-SVR, AS-SARIMA and the proposed method are used to forecast the load in the four weeks of four seasons in 2012. The data are obtained from ISO New England [5]. The experimental results are shown in Figure 10. AT denotes actual temperature, NT_1, NT_2 and NT_3 denotes the noisy temperature with the Gaussian noise of standard deviations of 0.6 °C, 0.9 °C and 1.2 °C.
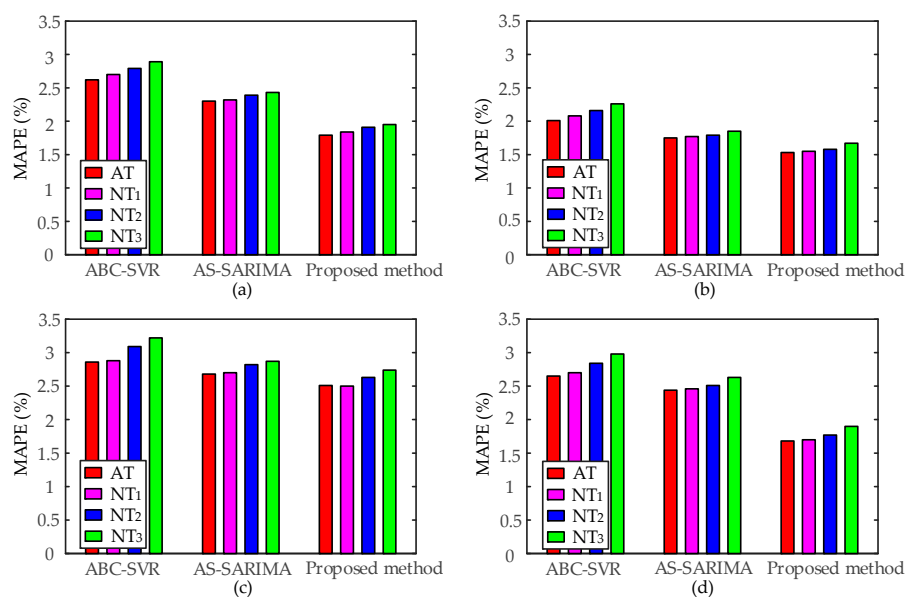
Figure 10. The average *MAPE* of different methods in 2012. (**a**) Prediction from 22 to 28 February; (**b**) prediction from 18 to 24 May; (**c**) prediction from 8 to 14 August; and (**d**) prediction from 15 to 21 November.

Figure 10 shows that the increases in *MAPE* of the three models are not large when the Gaussian noise with standard deviation of 0.6 °C. But when the standard deviations are 0.9 °C and 1.2 °C, the forecasting errors of ABC-SVR are obviously larger than those obtained in the scenario of actual temperature. Compared to ABC-SVR, the forecasting accuracy of the AS-SARIMA and the proposed method is less affected by the Gaussian noise with standard deviations of 0.9 °C and 1.2 °C. The experimental results are shown in Figure 9.

Table 9 indicates that, after adding the Gaussian noise with standard deviation of 0.6 °C to the actual temperature, the temperature error varies in the interval [−2.2 °C, 2.3 °C]. Considering the improvement of NWP accuracy, the temperature prediction error in the interval is reasonable [5]. In addition, when the standard deviations are 0.9 °C and 1.2 °C, the temperature errors vary in the

intervals [−3.3 °C, 3.5 °C] and [−4.4 °C, 4.4 °C], respectively. Meanwhile, the forecast results of different models show that the prediction accuracy of the ABC-SVR is the lowest and the most affected by temperature errors at any level. When the standard deviations of Gaussian noises are 0.6 °C and 0.9 °C, the rises of *MAPE* generated by AS-SARIMA and the proposed method are same. When the standard deviation of the noise is 1.2 °C, the rise of *MAPE* generated by the proposed method is slightly greater than that generated by AS-SARIMA. However, the proposed method always has the highest forecast accuracy.

Table 9. The *MAPE* (%) for different models with Gaussian noises of zero mean.

Standard Deviation (°C)	Error Range of Temperature (°C)	Models		
		ABC-SVR	**AS-SARIMA**	**Proposed Method**
0	n.a.	2.54	2.29	1.88
0.6	[−2.2, 2.3]	2.59	2.31	1.90
0.9	[−3.3, 3.5]	2.72	2.38	1.97
1.2	[−4.4, 4.4]	2.84	2.45	2.07

5. Conclusions

By comparing the experimental results for different models in forecasting load in various day types and analyzing the effects of temperature errors on prediction accuracy of each model, a new method based on ABC-SVR and AS-SARIMA is proposed in this paper. The advantages of the proposed method are as follows:

(1) Through using the ABC algorithm to optimize the parameters of SVR, the ABC-SVR model is constructed. It could improve the forecast accuracy of SVR by avoiding the selection of unreasonable parameters in the model.

(2) Considering the fluctuation of the load in weekends and comparing the prediction accuracy of different models, the ABC-SVR is used to forecast the load on weekends. When an ABC-SVR model is established, the exogenous variables such as temperature and day types are selected in addition to the historical load as the input variables of the model, so the forecast accuracy of the ABC-SVR approach is satisfactory.

(3) Considering the stability of the working days load and the influence of the exogenous variables on load, AS-SARIMA is used to forecast the load on working days. In the AS-SARIMA model, SARIMA is used to forecast original load values and ABC-SVR is used to.

(4) Modify the results of SARIMA by forecasting the residuals. Therefore, the prediction accuracy of the AS-SARIMA is high and little affected by NWP errors for working days load forecasting.

The simulation results based on real load data demonstrate that the proposed method has nice performance considering the NWP errors in STLF. To further improve the accuracy of STLF, more effective forecasting models and feature selection methods will be considered in future research.

Author Contributions: Guowei Cai put forward to the main idea and design the whole venation of this paper. Wenjin Wang did the experiments and prepared the manuscript. Junhai Lu guided the experiments and paper writing.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Raza, M.Q.; Khosravi, A. A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings. *Renew. Sustain. Energy Rev.* **2015**, *50*, 1352–1372. [CrossRef]

2. Bozic, M.; Stojanovic, M.; Stajic, Z.; Floranovic, N. Mutual information-based inputs selection for electric load time series forecasting. *Entropy* **2013**, *15*, 926–942. [CrossRef]

3. Kouhi, S.; Keynia, F. A new cascade NN based method to short-term load forecast in deregulated electricity market. *Energy Convers. Manag.* **2013**, *71*, 76–83. [CrossRef]

4. Kulkarni, S.; Simon, S.P.; Sundareswaran, K. A spiking neural network (SNN) forecast engine for short-term electrical load forecasting. *Appl. Soft Comput.* **2013**, *13*, 3628–3635. [CrossRef]

5. Li, S.; Wang, P.; Goel, L. Short-term load forecasting by wavelet transform and evolutionary extreme learning machine. *Electr. Power Syst. Res.* **2015**, *122*, 96–103. [CrossRef]

6. Reis, A.J.R.; da Silva, A.P.A. Feature extraction via multiresolution analysis for short-term load forecasting. *IEEE Trans. Power Syst.* **2005**, *20*, 189–198.

7. Sudheer, G.; Suseelatha, A. Short term load forecasting using wavelet transform combined with Holt–Winters and weighted nearest neighbor models. *Int. J. Electr. Power* **2015**, *64*, 340–346. [CrossRef]

8. Wang, J.J.; Wang, J.Z.; Li, Y.N.; Zhu, S.L.; Zhao, J. Techniques of applying wavelet de-noising into a combined model for short-term load forecasting. *Int. J. Electr. Power* **2014**, *62*, 816–824. [CrossRef]

9. Li, S.; Wang, P.; Goel, L. A novel wavelet-based ensemble method for short-term load forecasting with hybrid neural networks and feature selection. *IEEE Trans. Power Syst.* **2016**, *31*, 1788–1798. [CrossRef]

10. Gu, C.J.; Yang, D.Z.; Jirutitijaroen, P.; Walsh, W.M.; Reindl, T. Spatial load forecasting with communication failure using time-forward kriging. *IEEE Trans. Power Syst.* **2014**, *29*, 2875–2882.

11. Lee, W.J.; Hong, J. A hybrid dynamic and fuzzy time series model for mid-term power load forecasting. *Int. J. Electr. Power* **2015**, *64*, 1057–1062. [CrossRef]

12. Niu, D.X.; Shi, H.F.; Wu, D.D. Short-term load forecasting using bayesian neural networks learned by Hybrid Monte Carlo algorithm. *Appl. Soft Comput.* **2012**, *12*, 1822–1827. [CrossRef]

13. Hooshmand, R.A.; Amooshahi, H.; Parastegari, M. A hybrid intelligent algorithm based short-term load forecasting approach. *Int. J. Electr. Power* **2013**, *45*, 313–324. [CrossRef]

14. Chaturvedi, D.K.; Sinha, A.P.; Malik, O.P. Short term load forecast using fuzzy logic and wavelet transform integrated generalized neural network. *Int. J. Electr. Power* **2015**, *67*, 230–237. [CrossRef]

15. Kouhi, S.; Keynia, F.; Ravadanegh, S.N. A new short-term load forecast method based on neuro-evolutionary algorithm and chaotic feature selection. *Int. J. Electr. Power* **2014**, *62*, 862–867. [CrossRef]

16. Ko, C.N.; Lee, C.M. Short-term load forecasting using SVR (support vector regression)-based radial basis function neural network with dual extended Kalman filter. *Energy* **2013**, *49*, 413–422. [CrossRef]

17. Ceperic, E.; Ceperic, V.; Baric, A. A strategy for short-term load forecasting by support vector regression machines. *IEEE Trans. Power Syst.* **2013**, *28*, 4356–4364. [CrossRef]

18. Kavousi-Fard, A.; Samet, H.; Marzbani, F. A new hybrid modified firefly algorithm and support vector regression model for accurate short term load forecasting. *Expert Syst. Appl.* **2014**, *41*, 6047–6056. [CrossRef]

19. Che, J.X.; Wang, J.Z. Short-term load forecasting using a kernel-based support vector regression combination model. *Appl. Energy* **2014**, *132*, 602–609. [CrossRef]

20. Hu, Z.Y.; Bao, Y.K.; Xiong, T. Comprehensive learning particle swarm optimization based memetic algorithm for model selection in short-term load forecasting using support vector regression. *Appl. Soft Comput.* **2014**, *25*, 15–25. [CrossRef]

21. Lin, C.T.; Chou, L.D.; Chen, Y.M.; Tseng, L.M. A hybrid economic indices based short-term load forecasting system. *Int. J. Electr. Power* **2014**, *54*, 293–305. [CrossRef]

22. Moazzami, M.; Khodabakhshian, A.; Hooshmand, R. A new hybrid day-ahead peak load forecasting method for Iran's National Grid. *Appl. Energy* **2013**, *101*, 489–501. [CrossRef]

23. Wang, J.J.; Li, L.; Niu, D.X.; Tan, Z.F. An annual load forecasting model based on support vector regression with differential evolution algorithm. *Appl. Energy* **2012**, *94*, 65–70. [CrossRef]

24. Quan, H.; Srinivasan, D.; Khosravi, A. Uncertainty handling using neural network-based prediction intervals for electrical load forecasting. *Energy* **2014**, *73*, 916–925. [CrossRef]

25. Che, J.X. A novel hybrid model for bi-objective short-term electric load forecasting. *Int. J. Electr. Power* **2014**, *61*, 259–266. [CrossRef]

26. Hong, W.C.; Dong, Y.C.; Zhang, W.Y.; Chen, L.Y.; Panigrahi, B.K. Cyclic electric load forecasting by seasonal SVR with chaotic genetic algorithm. *Int. J. Electr. Power* **2013**, *44*, 604–614. [CrossRef]

27. Chen, Y.H.; Yang, Y.; Liu, C.Q.; Li, C.H.; Li, L. A hybrid application algorithm based on the support vector machine and artificial intelligence: An example of electric load forecasting. *Appl. Math. Model.* **2015**, *39*, 2617–2632. [CrossRef]

28. Yan, X.; Chowdhury, N.A. Mid-term electricity market clearing price forecasting utilizing hybrid support vector machine and auto-regressive moving average with external input. *Int. J. Electr. Power* **2014**, *63*, 64–70. [CrossRef]

29. Zhang, W.Y.; Hong, W.C.; Dong, Y.C.; Tsai, G.; Sung, J.T.; Fan, G.F. Application of SVR with chaotic GASA algorithm in cyclic electric load forecasting. *Energy* **2012**, *45*, 850–858. [CrossRef]

30. Hong, W.C. Chaotic particle swarm optimization algorithm in a support vector regression electric load forecasting model. *Energy Convers. Manag.* **2009**, *50*, 105–117. [CrossRef]

31. Yang, D.; Liu, Y.; Li, S.; Li, X.; Ma, L. Gear fault diagnosis based on support vector machine optimized by artificial bee colony algorithm. *Mech. Mach. Theory* **2015**, *90*, 219–229. [CrossRef]

32. Mandal, S.K.; Chan, F.T.S.; Tiwari, M.K. Leak detection of pipeline: An integrated approach of rough set theory and artificial bee colony trained SVM. *Expert Syst. Appl.* **2012**, *39*, 3071–3080. [CrossRef]

33. Kang, F.; Li, J. Artificial bee colony algorithm optimized support vector regression for system reliability analysis of slopes. *J. Comput. Civ. Eng.* **2015**, *30*. [CrossRef]

34. Karaboga, D.; Ozturk, C.; Karaboga, N.; Gorkemli, B. Artificial bee colony programming for symbolic regression. *Inf. Sci.* **2012**, *209*, 1–15. [CrossRef]

35. Mernik, M.; Liu, S.H.; Karaboga, D.; Crepinsek, M. On clarifying misconceptions when comparing variants of the Artificial Bee Colony Algorithm by offering a new implementation. *Inf. Sci.* **2015**, *291*, 115–127. [CrossRef]

36. El-Fergany, A.A.; Abdelaziz, A.Y. Capacitor placement for net saving maximization and system stability enhancement in distribution networks using artificial bee colony-based approach. *Int. J. Electr. Power* **2014**, *54*, 235–243. [CrossRef]

37. Karaboga, D.; Gorkemli, B.; Ozturk, C.; Karaboga, N. A comprehensive survey: Artificial bee colony (ABC) algorithm and applications. *Artif. Intell. Rev.* **2014**, *42*, 21–57. [CrossRef]

38. Areekul, P.; Senjyu, T.; Toyama, H.; Yona, A. Notice of violation of IEEE publication principles a hybrid ARIMA and neural network model for short-term price forecasting in deregulated market. *IEEE Trans. Power Syst.* **2010**, *25*, 524–530. [CrossRef]

39. Hu, Z.Y.; Bao, Y.K.; Xiong, T.; Chiong, R. Hybrid filter-wrapper feature selection for short-term load forecasting. *Eng. Appl. Artif. Intell.* **2015**, *40*, 17–27. [CrossRef]

40. Ding, N.; Benoit, C.; Foggia, G.; Besanger, Y.; Wurtz, F. Neural network-based model design for short-term load forecast in distribution systems. *IEEE Trans. Power Syst.* **2016**, *31*, 72–81. [CrossRef]

41. Che, J.X.; Wang, J.Z.; Tang, Y.J. Optimal training subset in a support vector regression electric load forecasting model. *Appl. Soft Comput.* **2012**, *12*, 1523–1531. [CrossRef]