Computer Life (CPL) ISSN: 1819-4818
DELIVERING QUALITY SCIENCE TO THE WORLD

Based on GBDT Algorithm of the Heating Load Forecast

Weihua Pan [a], Wenqian Zhang [b]

School of Control and Computer Engineering, North China Electric Power University, Baoding, China;

[a]23680118@qq.com, [b]505390676@qq.com

Abstract: Heating load forecasting can solve the problem of uneven load and short-term overload and load shortage. It can arrange heating work economically and reasonably to ensure normal production and life of society. The traditional prediction method is based on mathematical statistics, svm, etc. The cost of prediction is large and the time is long. This paper improves the parameters of load forecasting, adopts GBDT prediction model, and improves the prediction efficiency. Combined with dropout technology, the DGBDT model is proposed to make the prediction result. More precise, improving the model's resistance to over-fitting.
Keywords: Heating load; Load forecasting; GBDT; Dropout.

---

1. INTRODUCTION

With the continuous development of the global economy, energy consumption has become one of the most important problems in the world today. Energy consumption includes building energy consumption, transportation energy consumption and industrial energy consumption. Building energy consumption is the most important consumption of energy consumption. As China's urban central heating system continues to develop, the demand for heating quality by hot users continues to increase, and heating and heating for winter residents is the main factor for building energy consumption. Therefore, the energy saving of the central heating system is particularly important, which can effectively reduce energy consumption.

The heating load forecasting is based on the full consideration of the operating characteristics of the heating system, natural conditions and social development. The scientific analysis method is used to objectively reflect the mutual laws between various factors, so as to construct a model that can provide The trend of the heat load is predicted. In recent years, a large number of scholars have been engaged in research on heating load forecasting. Literature [1] uses the advantages of cross-validation in model performance evaluation and selection, combined with the global optimization ability of genetic algorithm, proposes a SVR heating load prediction model based on genetic algorithm optimization. In Literature [2] , aiming at the shortcomings of particle swarm optimization (PSO), which is easy to fall into local optimum and slow convergence in the late stage of evolution, a heating load forecasting based on improved particle swarm optimization and combined prediction is

proposed. In Literature [3] , a strong predictive model combining extended ant colony algorithm and wavelet neural network is proposed for the problem that the heating load is highly nonlinear and the wavelet neural network is easy to fall into the local minimum and the convergence speed is slow.

This paper systematically introduces the influencing factors of heating load, and proposes new parameters based on its thermal inertia and periodicity. In addition, it introduces the principle, advantages and disadvantages of GBDT algorithm, and adopts Shrinkage regularization technology in combating over-fitting. The DGBDT algorithm is obtained by means of Dropout technology, which further improves the over-fitting performance of the GBDT algorithm.

2. INFLUENCING FACTORS OF HEATING LOAD FORECASTING

The main factors affecting the thermal load of the central heating system include meteorological factors and internal factors of the system. The meteorological factors have a great influence on the heat load of the central heating system, but most of the meteorological factors at this stage mainly consider the influence of outdoor temperature, ignoring the indirect effects of other factors. In 2004, Westphal et al. used different meteorological parameters and different cycles to influence the heating load. Through experiments, they found that the difference between different standard meteorological data can reach 18%. In 2006, Nielsen et al. used the gray box method to study the heat consumption of the district heating system, and established the relationship between heat consumption and outdoor temperature, wind speed, solar radiation and so on. In 2009, Tolga et al. selected three parameters of temperature, solar radiation and wind speed to study the regression equations of energy consumption and these meteorological factors.

The outdoor temperature has the greatest influence on the central heating load. This paper selects the experimental data of a heat exchange station from January to June of 2014, and performs a one-way regression fitting by least squares method to obtain the corresponding outdoor temperature and The fitting equation for the heat load is:

$$Q_l = 1.0221 - 0.0147T$$

Q_l among them is the daily average load of the heat exchange station, the unit position MW, T is the outdoor temperature, the unit position °C, and the square coefficient of the one-way regression straight line fitting is 0.8011, which shows that the relationship between the outdoor temperature and the heating load is approximately linear. relationship. The influence of the increase of wind speed on the surrounding environment is the decrease of temperature. With this idea, this paper considers the influence of wind speed on outdoor temperature. When the wind speed increases, the corresponding outdoor temperature can be converted into an equivalent temperature decrease, and the proposed wind speed cooling is equivalent to the temperature calculation formula:

$$\Delta T_w = 0.0246(\lg(7.23W_w))^3 - 0.4525(\lg(7.23W_w))^2 + 3.2398\lg(7.23W_w)$$

W_w is the wind speed (m/s) of the outside world; ΔT_w is the equivalent temperature value (°C) for the wind speed reduction. After that, the converted equivalent temperature of the wind speed is added to the actual outdoor temperature, and the outdoor equivalent temperature affected by the wind speed is obtained. The first-order regression fitting is also performed by the least squares method to obtain the heat exchange station under the influence of the wind speed. The fitted equation of the load:

$$Q_l = 0.9556 - 0.198T_{ot+w}$$

Through the above equations, it can be seen that the outdoor temperature and the heat load of the heat exchange station have an approximate linear relationship, and the coefficient is larger than the value considering the outdoor temperature alone.

The influence of the same sunshine on the heating load can also be indirectly converted into the influence of temperature on the heating load. The conversion process will not be described again, and the one-way regression equation between the outdoor temperature and the heat load of the heat exchange station can also be obtained:

$$Q_l = 1.0374 - 0.0183T_{o-s}$$

According to the above analysis, it can be concluded that the outdoor temperature and the heat load of the central heating system have an approximate linear relationship. By affecting the influence of wind speed or sunshine as the temperature, the influence of wind speed and sunshine on the outdoor temperature can be obtained. In summary, this paper chooses outdoor temperature, sunshine and wind speed as the influencing factors of the heat load of the central heating system. Considering the thermal inertia and periodicity of the thermal load, it is necessary to consider the previous thermal load and the influence of time on the current thermal load. Here, the previous thermal load selects the value of the previous day, and the periodicity of time is mainly based on the year. The unit, so the time is taken in the month, based on the above, the input and output variables of the selected thermal load prediction model are summarized as shown in the table.

Table.1 Thermal load input and output variables

Input variable	Output variable
Daily average outdoor temperature T/°C	Daily average heat load forecast Q/MW
Daily average wind speed W/（m/s）	
Daily average sunshine $S/(W/m^2)$	
Yesterday average heat load Q/MW	
Month	

3. BASED ON GBDT- THERMAL LOAD PREDICTION MODEL

The GBDT algorithm was proposed by Leo Breiman et al. in 1997 [4] and was used in regression analysis in 1999. Later, some scholars have improved and optimized the GBDT algorithm to make it more mature. The principle of GBDT is to use the gradient descent learning method to learn a number of weak learning periods, and then combine to produce a strong learner. In the iteration of GBDT, assuming that the strong learner obtained in the previous round of iteration is $f_{t-1}(x)$ and the loss function is $L(y, f_{t-1}(x))$, then the goal of this round of iteration is to find a CART regression tree model weak learner $h_t(x)$, which minimizes the loss of the current round of learning. In other words, the purpose of this round of falls is to find a decision tree that minimizes sample losses.

3.1 GBDT regression algorithm

Input: Training sample $T = \{(x_1, y_1), (x_2, y_2), ..., (x_m, y_m)\}$, the maximum number of iterations is T, and the loss function is L.

Output: Strong Learner $f(x)$

1) Initialize the weak learner

$$f_0(x) = \arg\min \sum_{i=1}^{m} L(y_i, c)$$

2) For the number of iterations t

$$r_{ti} = -[\frac{\partial L(y_i, f(x_i))}{\partial f(x_i)}]_{f(x)=f_{t-1}(x)}$$

b) Using (
J

$$c_{tj} = \arg\min \sum_{x_i \in R_{tj}} L(y_i, f_{t-1}(x_i) + c)$$

$$f_t(x) = f_{t-1}(x) + \sum_{j=1}^{J} c_{tj} I(x \in R_{tj})$$

3) Get the expression of the strong learner f(x)

$$f(x) = f_T(x) = f_0(x) + \sum_{t=1}^{T} \sum_{j=1}^{J} c_{tj} I(x \in R_{tj})$$

GBDT can handle various types of data flexibly, and it can achieve higher prediction accuracy with fewer parameters than SVM (Support Vector Machine). The difference between Gradient Boosting and the previous Boost is that the purpose of each training is to reduce the last residual, and in order to reduce the residual, it is necessary to train a new model in the gradient direction of reducing the residual, namely Gradient Boosting. Each new model is trained to reduce the residuals of the previous model. GBDT integrates the idea of collective learning, which is reflected in each iteration to train the information on each sample with a new model, and finally becomes a strong classifier through the fusion of N models.

3.2 Shrinkage Thought

Over-fitting a training sample leads to a decline in the generalization ability of the model. Shrinkage as a regularization method can reduce the influence of over-fitting in model construction. The meaning of Shrinkage is to reduce, that is, Shrinkage believes that the excessive pace of the model training leads to the lack of sample information. Compared with the next step, the desired result can better balance the global information, thus preventing over-fitting. effect.

The gradient descent algorithm has a part that is regularized by reduction. The purpose of the step is to prevent overfitting, as shown in the formula:

$$F_m(x) = F_{m-1}(x) + v \cdot r_m h_m(x), 0 < v < 1$$

The parameter v in the expression is called the learning rate. The general learning rate takes a smaller number, so the generalization ability of the model will be much improved. However, controlling the learning rate to increase the generalization ability requires increasing the number of iterations of the model, so that the lower the learning rate, the more iterations.

3.3 Combination of GBDT and Dropout

The Shrinkage idea of the previous chapter is a general idea of GBDT to prevent overfitting. This method can achieve better results under normal circumstances, but the effect is limited. This chapter proposes a new method to prevent overfitting, which is to apply the Dropout technology to GBDT and get the DGBDT algorithm. Dropout originally refers to the fact that during the training of the

deep learning network, the neurons do not work according to a certain probability, so for the random gradient drop, the random guarantee is that each batch is training different networks.

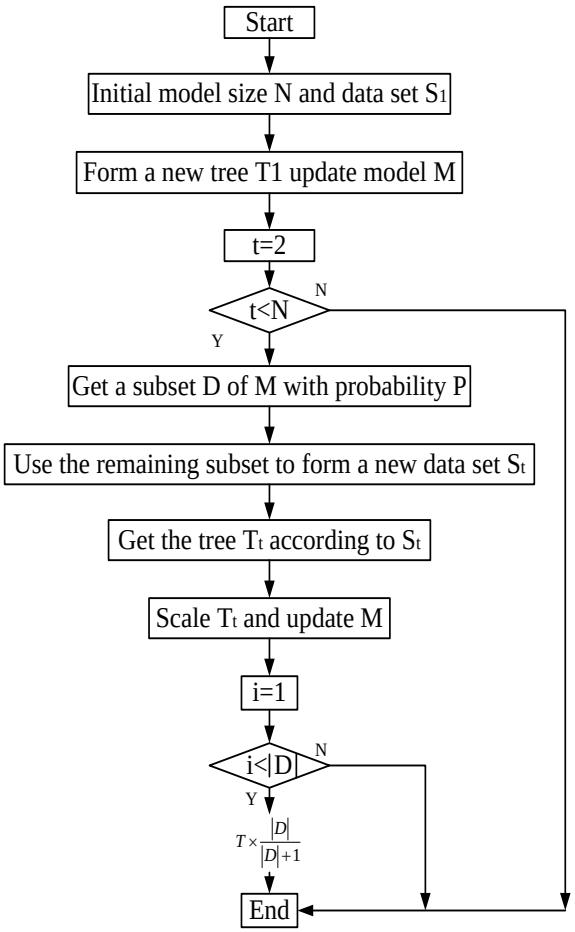

Fig. 1 DGBDT process

The DGBDT algorithm, the GBDT algorithm that incorporates Dropout, mainly improves both aspects of GBDT. The first is to calculate the current time random gradient, the model used is not the accumulation of the previous time model in GBDT, but a part of it is randomly selected. The second is that the change is when merging the new tree, because the new tree T is to narrow the error of $\hat{M}$ (the sum of randomly selected trees) and the optimal prediction model, but the tree abandoned in the random process is also to narrow the error. Therefore, the introduction of the new tree T and the abandoned tree will cause the result to exceed the target value. This requires a reduction operation on the new tree T. See the algorithm for the specific operation.

DGBDT algorithm:

1) Suppose N is the total number of all trees added to the final model

2) $S_1 \leftarrow \{x, -L_x^{'}(0)\}$ forms the data set S

3) Generate a new tree through data set S1

4) $M \leftarrow \{T_1\}$ by updating the model

5) For t=2,...,N do

6) A subset of $D \leftarrow$M, each tree T in M is added to D with a certain probability p

7) If D is not empty then D is the random element obtained from M and

8) $\hat{M} \leftarrow M \setminus D$ get the sum of the remaining models

9) $S_t \leftarrow \{x, -L_x^{'}(\hat{M}(x))\}$ get the new data set at time t

10) Get the new tree TT obtained according to the set ST training

11) 77 update M value

12) For the element T in D

13) $T \times \frac{|D|}{|D|+1}$

14) End

15) End

16) Output M

The DGBDT algorithm can effectively over-fitting, so it can be seen as a means of regularization to remove some elements. When you don't abandon any tree, the DGBDT is the same as the GBDT. From the other extreme, when all the trees are not considered, the DGBDT becomes a random forest algorithm. Therefore, the size of the deleted collection affects the performance of the algorithm. There are many ways to choose a tree to be abandoned. This article uses a two-plus-one technique. In this technique, each existing tree is discarded with a certain probability p. However, when no tree is abandoned, the above principle is used, that is, a random number is discarded. Therefore, only one number is discarded for each round of operations.

4. EXPERIMENT AND ANALYSIS

The experimental data used more than 40,000 pieces of data from 2010 to 2014 as a training set. In 2015, more than 8,000 pieces of data were used as test sets. The data set was too small, and the root mean square error RMSE was larger overall. In the course of this experiment, the data processing mainly includes the following aspects: 1) removing invalid data, that is, deleting all the data in the format error, the data value is too large or too small, etc.; 2) adding the yesterday load and each line of data Date; 3) Disrupt the data set.

The experimental links were compared with random forests, GBDT and DGBDT using Shrinkage mechanism, and parameters were tuned for each model. The parameters included the Shrinkage value of GBDT, the Dropout rate of DGBDT, the number of trees, and the leaves of each tree. The number and the maximum depth of the tree, where the black is the optimal value for the specific parameters of the model.

Table.2 Comparison of various model parameters

Parameter	Random forest algorithm	GBDT	DGBDT
Shrinkage		0.1,0.2,0.3,0.4	
Dropout rate			0.015,0.03,0.045
Number of trees	100,250,500,1000	100,250,500,1000	100,250,500,1000
Number of leaves per tree	100,250,500,1000	40	40

Maximum depth of number	1,2,3,4,5,6,7,8,	1,2,3,4,5,6	1,2,3,4,5,6

Table 3 Comparison of RMSE values of each model

Total model size	50	100	250	500	1000
Random forest	65.58	64.47	65.29	64.97	63.89
GBDT	52.20	52.87	52.07	51.94	52.36
DGBDT	52.09	51.13	51.67	50.97	51.62

It can be seen from Table 2 that the model of DGBDT is slightly better than the GBDT based on Shrinkage regularization, and the optimal value is obtained when the total number of iterations is 500, and the RMSE value of the random forest is significantly worse than GBDT and DBGDT. It is caused by the model itself, and on the other hand, it is mainly reflected in the small data set. Figure 2 is a heat supply load prediction chart under three models. It can be seen from the figure that all three models have better simulation prediction trends, and the DGBDT fitting actual value is the best.

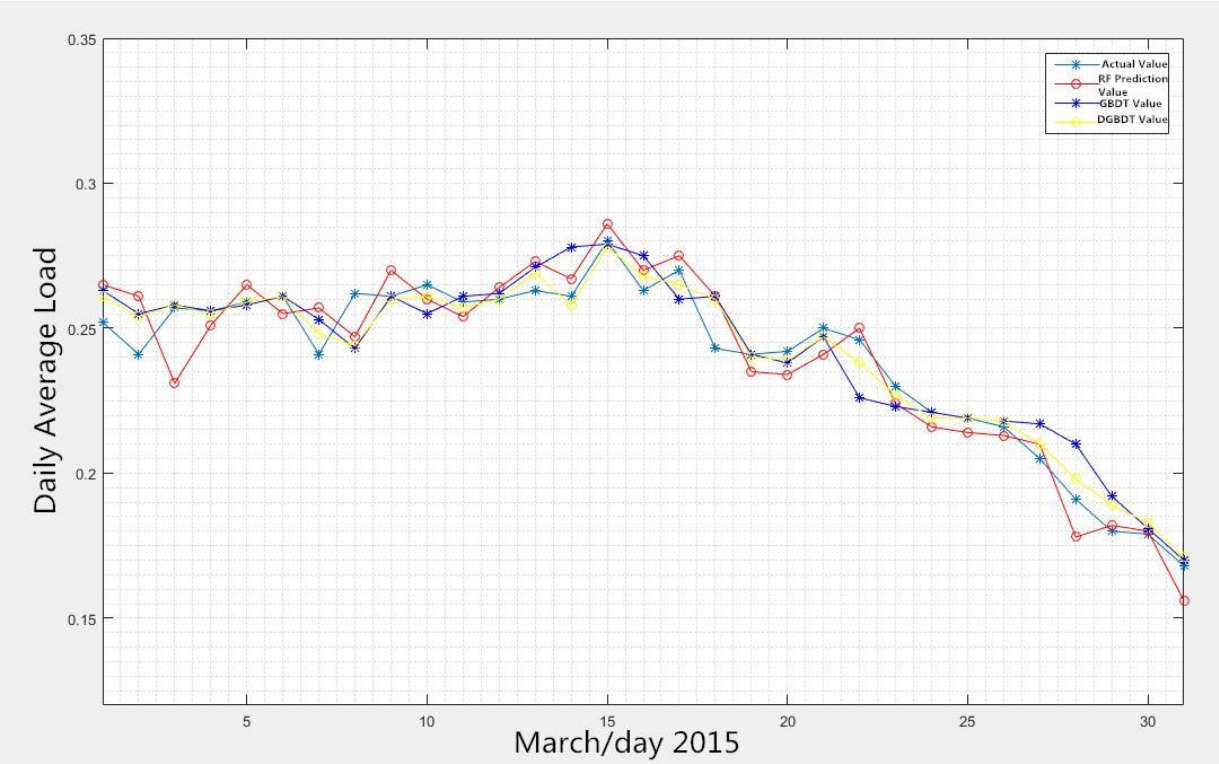

Fig. 2 Heat load system heat load prediction map (unit MW)

5. CONCLUSION

This paper systematically introduces the whole process of heating load forecasting based on GBDT algorithm. Firstly, the relationship between meteorological factors and heating load is introduced in the first chapter, and the regression is performed by the one-way regression equation, and the wind speed and sunshine are converted into temperature pairs. The influence of the heating load, and based on the thermal inertia of the load prediction, periodically predicts the heating load. On the other hand, this paper systematically introduces the algorithm principle of GBDT and the general Shrinkage method against over-fitting, and with the successful application of Dee-pout technology in deep learning, it is fused with GBDT to obtain DGBDT algorithm. After experimental comparison, The

algorithm can effectively combat over-fitting. In addition, the GBDT algorithm cannot be paralleled like a random forest, which is also a direction for future research.

REFERENCES

[1]. Zhang Wei, Tian Qi, Wang Meiping. Heating load forecasting based on genetic algorithm optimization of support vector regression parameters,Vol.2 HVAC (2017) No. 47, p.104-108.
[2]. Gao Bingkun, Li Yang, Xu Mingzi. Application of optimized particle swarm optimization algorithm in combined heating load forecasting. Journal of Terahertz Science and Electronics, s,Vol.5 (2011) No. 09, p.655-659.
[3]. Wei Yujie, Yang Jieming. Heating load forecasting based on wavelet neural network and ant colony algorithm. Coal Technologys,Vol.8 (2015) No. 34.
[4]. Breiman, L. "Arcing The Edge". 1997
[5]. Ke Guolin. Research on parallel learning algorithm of gradient lifting decision tree (GBDT) . Xiamen University, 2016.
[6]. Li Shengtao. Research on thermal load forecasting method for central heating system . Chang'an University, 2014.
[7]. Wang Wenbiao, Cai Wei, Wang Siyuan. Research on Thermal Load Forecasting of Central Heating System Based on Meteorological Factors. Journal of Computer Measurement and Controls,Vol.2 (2016) No. 24, p.22-26.