



## Prediction of Steam Consumption in the Tobacco Industry's Primary Processing Workshop

Wei Gao<sup>1</sup>, Kai Zhang<sup>2</sup>, Feng Zheng<sup>1</sup>, Xiaolei Liu<sup>1</sup>, Mingqi Nan<sup>1</sup>, Lei Yan<sup>1</sup>, Yongchun Wen<sup>1</sup>,  
Yuxuan Han<sup>2</sup>, Hongxi Zhang<sup>2,3\*</sup>

<sup>1</sup> Zhangjiakou Cigarette Factory Co., Ltd., Zhangjiakou 075000, China

<sup>2</sup> Department of Energy Engineering, Hebei University of Architecture, Zhangjiakou 075000, China

<sup>3</sup> Hebei Energy Storage Heating Technology Innovation Centre, Zhangjiakou 075000, China

Corresponding Author Email: [zhx1996@hebiace.edu.cn](mailto:zhx1996@hebiace.edu.cn)

Copyright: ©2024 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/ijht.420538>

### ABSTRACT

**Received:** 7 April 2024

**Revised:** 20 September 2024

**Accepted:** 3 October 2024

**Available online:** 31 October 2024

#### **Keywords:**

*steam, primary processing, Random Forest, prediction model*

To achieve precise energy supply and energy-efficient operation of the steam system in the primary processing workshop of tobacco enterprises, a Random Forest prediction model is introduced in this study to forecast steam consumption. A correlation analysis is conducted to identify the factors that significantly impact steam load. These key features are then used to develop the Random Forest prediction model, with appropriate hyper parameters selected to enhance prediction accuracy. The results indicate that the Random Forest model offers high prediction precision and strong practical applicability, providing a basis for optimizing the scheduling of the steam system.

## 1. INTRODUCTION

With the continuous maturation of emerging technologies such as artificial intelligence and machine learning, the intelligent application of the primary processing production process has gradually become a major research focus in the industry [1-2]. Steam, as an essential secondary energy source in the tobacco industry, relies heavily on non-renewable energy sources, including natural gas, coal, and petroleum. Rational prediction and decision-making regarding steam production and consumption are of great significance for enterprises to save resources and improve production efficiency [3-7]. Currently, in the steam systems of tobacco enterprises, the primary processing workshop consumes a large amount of steam. When production begins in the primary processing workshop, gas boilers in the power workshop operate at high loads to provide steam to the system. To ensure the production quality of cut tobacco, the steam supply to the primary processing workshop often exceeds its actual demand. This unrefined management of steam supply results in significant steam dispersion, falling short of achieving precise supply and demand management, and causing energy waste due to excessive steam production.

In recent years, increasing attention has been paid to research on energy conservation, consumption reduction, and refined management. Commonly used energy consumption prediction methods include parametric regression, time series analysis, and artificial neural networks [8-13]. For example, Chen et al. [14] proposed an energy consumption prediction model for air conditioning systems based on a deep learning gated recurrent unit (GRU) neural network, predicting energy consumption data for the air conditioning system of a tobacco factory's storage workshop. Zhao et al. [15] developed a

prediction model for steam production and consumption in the steel industry, achieving accurate steam consumption predictions and reducing energy consumption per ton of steel. Song et al. [16] used six machine learning algorithms, including decision trees, Bayesian classifiers, neural networks, and logistic regression, to predict personal annual income. Yu et al. [17] analyzed educational big data using five major machine learning algorithms: logistic regression, decision trees, Bayesian algorithms, Random Forest, and others. While the scope of research on load prediction is broad, studies on steam load prediction in the primary processing workshops of tobacco enterprises are limited. Multiple production devices interact within the workshop, with five steam pipelines supplying high-pressure steam to each piece of equipment, greatly increasing the complexity of the steam load fluctuation patterns. The steam load in the primary processing workshop of tobacco enterprises fluctuates significantly with the start and stop of production equipment. Relying on workers' experience to measure and adjust the steam supply often leads to energy waste. If an accurate prediction model is available, it would facilitate source-side control, achieving the goal of energy conservation and emission reduction.

The steam load in the primary processing workshop is affected by various factors, including production processes, equipment, and tasks, resulting in nonlinear changes in steam load data with no obvious patterns and relatively small data volumes. Due to its ensemble nature and feature randomness, Random Forest can effectively reduce the problem of model overfitting. Moreover, the OOB Predictor Importance function embedded in the TreeBagger function of the Random Forest model provides feature importance evaluation, helping identify the key factors influencing steam load. Based on the above analysis, this paper examines the production processes

of tobacco enterprises and identifies the factors affecting steam load. It analyzes the relationship between production plans and the required steam load, selecting highly correlated factors as input variables for the model. The Random Forest learning method is employed to predict steam consumption in the primary processing workshop, and the prediction results are compared with those from other traditional models. This research offers a refined prediction method for forecasting steam consumption in the primary processing workshop, providing data support for the rational scheduling of the steam system in tobacco enterprises.

2. PRODUCTION PROCESS AND STEAM SYSTEM OF THE PRIMARY PROCESSING WORKSHOP IN A TOBACCO ENTERPRISE

The primary processing stage is a crucial part of the production process in tobacco enterprises. Its production processes and equipment significantly impact the overall production efficiency and product quality of the entire factory. The steam consumption of the primary processing workshop is a top priority for the whole factory, and the stability of the workshop's steam system directly affects the productivity and annual economic performance of the enterprise.

2.1 Production process of the primary processing workshop

Steam plays multiple roles in the production process, primarily used for humidifying and heating tobacco leaves and stems, generating hot air to dry tobacco, and injecting flavors to enhance tobacco quality. The detailed production process of the primary processing workshop is shown in Figure 1. When assigning production tasks, the enterprise distributes them across the following eight areas: the vacuum area, leaf area, Phase I leaf strip area, Phase II leaf strip area, Phase I blending

area, Phase II blending area, stem area (before stem storage), and stem area (after stem storage).

In the vacuum area and leaf area, the vacuum rehumidifiers and loose rehumidifiers consume large amounts of steam. When these devices are in operation, the steam pipeline load fluctuates significantly. In the blending area, the production process mixes stems, leaf strips, expanded strips, and recycled tobacco strips according to a set ratio based on the flow of leaf strips. The blended tobacco strips are then transferred to the flavoring machine in the blending area, where they are flavored in a specific ratio to produce qualified cut tobacco, which is stored in cabinets for later use. This process involves minimal steam-consuming equipment, with only the flavoring machine requiring a small amount of steam for injecting flavors. The steam-consuming equipment in different process areas is listed in Table 1.

2.2 Steam system of the primary processing workshop

During production, the daily steam consumption of the primary processing workshop accounts for approximately 65% to 90% of the factory's total steam consumption, making it the largest steam consumer in the plant. The power workshop of the tobacco factory provides steam to the internal steam system. As the production conditions change, the boiler system load in the power workshop increases rapidly within a short period, resulting in significant fluctuations in the steam load. The steam supply to the primary processing workshop is distributed through five steam pipelines: the Phase I vacuum rehumidification steam pipeline, the Phase II vacuum rehumidification steam pipeline, the Phase I leaf line steam pipeline, the Phase II leaf line steam pipeline, and the stem line steam pipeline. Each pipeline provides steam to different steam-consuming equipment, and the steam quality requirements vary across the equipment. The required steam quality in the primary processing workshop ranges from 0.2 MPa to 1.1 MPa.

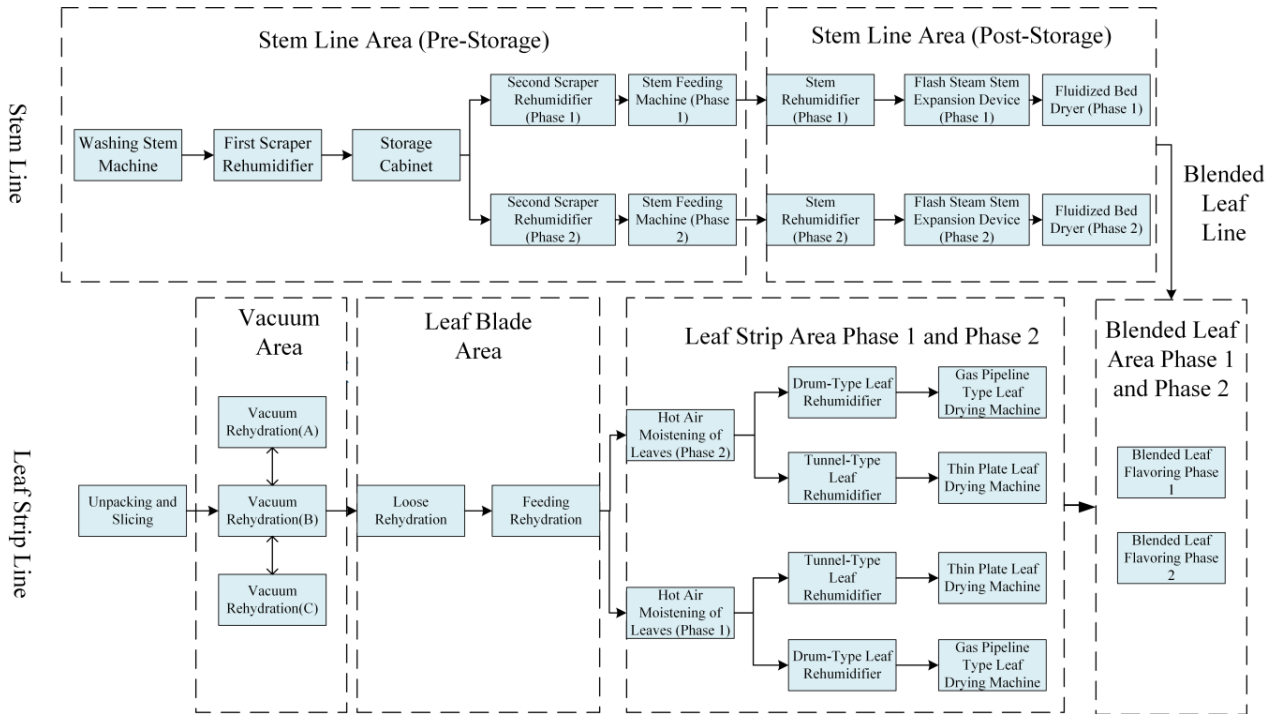


Figure 1. Production process of the tobacco manufacturing workshop

**Table 1.** Steam-consuming equipment in different production process areas of the primary processing workshop

Process Area of the Primary Processing Workshop	Steam-Consuming Equipment
Vacuum Area	Vacuum Rehumidifier (A), Vacuum Rehumidifier (B), Vacuum Rehumidifier (C)
Leaf Blade Area	Loose Rehumidifier, Feeding Rehumidifier
Leaf Strip Area (Phase I)	Hot Air Leaf Conditioner (Phase I), Drum-Type Leaf Strip Rehumidifier (Phase I), Tunnel-Type Leaf Strip Rehumidifier (Phase I), Plate-Type Leaf Dryer (Phase I), Gas-Pipeline-Type Leaf Dryer (Phase I)
Leaf Strip Area (Phase II)	Hot Air Leaf Conditioner (Phase II), Drum-Type Leaf Strip Rehumidifier (Phase II), Tunnel-Type Leaf Strip Rehumidifier (Phase II), Plate-Type Leaf Dryer (Phase II), Gas-Pipeline-Type Leaf Dryer (Phase II)
Stem Area (Before Stem Storage)	Stem Washer, Primary Scraper Conveyor, Secondary Scraper Conveyor (Phase I), Secondary Scraper Conveyor (Phase II), Stem Feeder (Phase I), Stem Feeder (Phase II)
Stem Area (After Stem Storage)	Stem Rehumidifier (Phase I), Stem Rehumidifier (Phase II), Flash Steam Stem Expansion Unit (Phase I), Flash Steam Stem Expansion Unit (Phase II), Fluidized Bed Dryer (Phase I), Fluidized Bed Dryer (Phase II)
Blending Area (Phase I)	Flavoring Machine
Blending Area (Phase II)	

### 2.3 Current steam usage in the primary processing workshop

The total daily steam consumption of the primary processing workshop is the sum of the steam used by the five steam pipelines in the workshop's steam system. A comprehensive analysis of the workshop reveals a strong coupling between the fluctuations in daily steam load and the production tasks assigned by the enterprise. These tasks are distributed according to the following process areas: vacuum area, leaf blade area, Phase I leaf strip area, Phase II leaf strip area, Phase I blending area, Phase II blending area, stem area (before stem storage), and stem area (after stem storage). Since there are no steam meters installed in the workshop to measure the steam consumption of each process area, it is not possible to collect data on the steam used by each area. Therefore, this paper analyzes the production tasks assigned to the eight process areas and the total daily steam consumption of the workshop to forecast the daily steam load.

## 3. STEAM CONSUMPTION PREDICTION MODEL FOR THE PRIMARY PROCESSING WORKSHOP

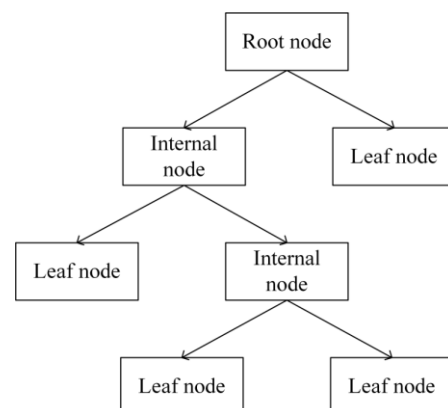
### 3.1 Data processing

The data used in this study were collected from the primary processing workshop of a tobacco factory over a period of seven months, from March 2023 to September 2023. Due to factory shutdowns or production scheduling issues, data for 138 days of production tasks and steam energy consumption were obtained during this period. First, outlier detection was performed on the dataset, and noisy data with high fluctuations in steam load were preprocessed by removing faulty data. There were no missing values in the dataset. Afterward, normalization was applied to the dataset, and the processed data were fed into the model. After preprocessing, a total of 131 data samples were obtained, which were used in the simulation tests of the prediction model described later.

### 3.2 Decision tree model

The decision tree is a commonly used machine learning algorithm that performs classification and regression analysis by simulating the human decision-making process. Its core idea is to partition the dataset based on features to build a tree-structured model. As a supervised learning algorithm, the decision tree is easy to understand and interpret. In this study, the decision tree extracts decision rules from the factors

affecting the steam load in the primary processing workshop and the corresponding steam load to make predictions. The structure of the decision tree is similar to that of a real tree, consisting of internal nodes, a root node, and leaf nodes. The classification or partitioning of the steam load-related factors in the workshop determines the number of nodes in the decision tree. The goal of this study is to use the influencing factors of the steam load in the workshop to predict the steam load for the next time step. Since this is a regression problem, a regression decision tree has been chosen for steam load prediction. Figure 2 shows the topology of the decision tree.

**Figure 2.** Principle of decision tree

The primary algorithm used in regression decision trees is Classification and Regression Trees (CART). The CART algorithm can be applied to both classification and regression problems. For regression problems, CART predicts continuous target variables by constructing decision trees. The CART algorithm follows these steps: the model input consists of data from the training set, and it continuously calculates the size of different values under various influencing factors. The optimal splitting point is found using the least squares method to minimize the squared error, which solves the following equation:

$$\min_{j,s} [\min_{c_1} \sum_{x_i \in R_1(j,s)} (y_i - c_1)^2 + \min_{c_2} \sum_{x_i \in R_2(j,s)} (y_i - c_2)^2]$$

where,  $j$  represents the  $j$ -th feature variable in the dataset;  $s$  represents value of the  $j$ -th feature variable; and there are:  $R_1(j,s) = \{x | x^j \leq s\}$ , and  $R_2(j,s) = \{x | x^j \geq s\}$ .

Through this process, the calculation values for different regions are obtained, and the above steps are repeated for subregions until the squared error reaches its minimum. The

final result is a division into several regions, forming a regression decision tree.

3.3 Random Forest model

The Random Forest is an ensemble learning method that primarily uses decision trees as base learners. By constructing multiple decision trees and aggregating their prediction results, the model enhances both its accuracy and stability. The Random Forest model employs bootstrap sampling to train and generate multiple decision trees, which together form the Random Forest. The outputs of these trees are then aggregated using voting or weighted averaging, improving the model's robustness and generalization capability. This approach embodies the concept of “collective wisdom”—while individual models may be prone to errors, the collective prediction of multiple models is more reliable. Random Forests perform exceptionally well across various datasets, especially when the feature space is large. Compared to traditional BP neural networks and some deep learning models, Random Forests achieve greater robustness through collective predictions by building multiple decision trees. This makes them highly effective in situations involving large fluctuations in steam load in the workshop and relatively small datasets, ensuring high-performance predictions.

3.4 Correlation analysis

The Random Forest model not only solves classification or regression problems but also evaluates the importance of each factor in the dataset. The Out-of-Bag (OOB) dataset is used in the Random Forest algorithm to assess model performance and feature importance. OOB refers to the data that are not included in the training set due to the bootstrap sampling process, where approximately one-third of the data are left out. These OOB samples serve as a built-in validation set, allowing model evaluation without additional computational costs. The concept of OOBPredictorImportance is introduced in the model to assess the importance of each predictor (feature) using the OOB dataset. Each decision tree in the Random Forest is independently built, with the splitting features at each node selected from a randomly chosen subset of features. This ensures that each feature is evaluated a varying number of times across different trees, providing a natural way to compare feature importance. The importance of a feature is determined by calculating the impurity reduction (for classification) or error reduction (for regression) it achieves when splitting nodes across all trees. The cumulative contribution of each feature is averaged or weighted to obtain an overall importance score. The higher the importance score, the greater the impact of that feature on the model's predictions. Using this method, the importance of various factors affecting the steam load in the primary processing workshop is determined, with the results presented in Table 2.

Analysis of the Table 2 shows that the factors significantly affecting the steam load in the primary processing workshop are the Vacuum Area, Leaf Blade Area, Stem Area (Before stem storage), and Stem Area (After stem storage), with correlation coefficients all above 0.5, indicating strong correlations. The Leaf Strip Areas (Phase I and Phase II) exhibit weak correlations with steam load, with coefficients around 0.4. The Blending Areas (Phase I and Phase II) have correlation coefficients below 0.2, indicating very low correlation. The low correlation in the blending areas can be

explained by the fact that only the blending and flavoring machines use steam in this area, and the amount of steam used is minimal, making it negligible. When building the predictive model, it is important to select highly correlated factors as input features. The more comprehensive and relevant the input features, the higher the model's prediction accuracy. However, including too many features may lead to overfitting, negatively affecting the model's performance. The correlation analysis provides a reference for selecting input features [18]. For this study, factors with correlation coefficients above 0.5—namely, the Vacuum Area, Leaf Blade Area, Stem Area (Before stem storage), and Stem Area (After stem storage)—are chosen as input features for the prediction model.

Table 2. Correlation between influencing factors and steam load in the primary processing workshop

Influencing Factor	Correlation Coefficient
Vacuum Area	0.68
Leaf Blade Area	0.69
Leaf Strip Area (Phase I)	0.40
Leaf Strip Area (Phase II)	0.44
Blending Area (Phase I)	0.16
Blending Area (Phase II)	0.13
Stem Area (Pre-Storage)	0.73
Stem Area (Post-Storage)	0.65

3.5 Hyperparameter selection

For the Random Forest model, the tree depth (dp) and number of trees (es) are generally considered the two most important hyperparameters. In this paper, the primary focus is on selecting the maximum tree depth and the number of regression trees, while other parameters are kept at their default values. The maximum depth of the tree is a crucial hyperparameter for both decision trees and Random Forest models. If the maximum depth is set too shallow, the model may fail to capture the complex relationships in the data, leading to underfitting. Conversely, if the depth is too large, the model might overfit the noise in the training data, reducing its generalization ability. Common methods for determining the optimal tree depth include cross-validation and grid search. In this study, cross-validation was employed to determine the optimal maximum depth, which was found to be 4.

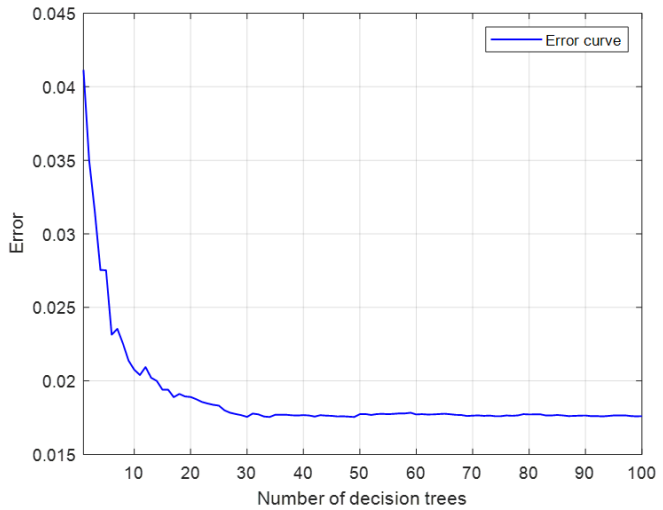


Figure 3. Impact of the number of trees on prediction accuracy



The number of trees in a Random Forest refers to the number of decision trees within the model. Increasing the number of trees typically improves the model's performance, as more trees can better capture the features of the data and reduce variance. However, the performance gains are not linear, and after a certain point, additional trees offer diminishing returns. Too many trees may also lead to overfitting, especially when the dataset is relatively small. The Figure 3 shows the change in prediction error with different numbers of decision trees, using the OOB error as the evaluation metric. As illustrated in Figure 3 and Table 3, the prediction error decreases as the number of trees increases. The error reaches 0.01755 when the number of trees reaches 30, and further increases in the number of trees provide no significant improvement in prediction accuracy. To avoid overfitting, the optimal number of trees was set to 30.

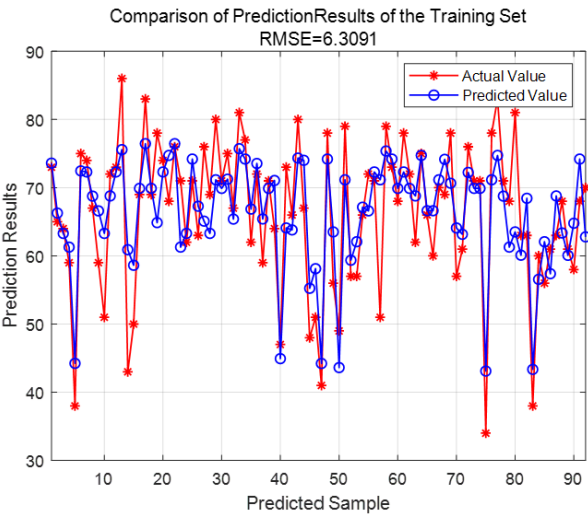
**Table 3.** Errors for different numbers of decision trees

Number of Trees	Error	Number of Trees	Error
10	0.02076	22	0.01857
11	0.02040	23	0.01846
12	0.02094	24	0.01838
13	0.02021	25	0.01832
14	0.01999	26	0.01800
15	0.01940	27	0.01784
16	0.01941	28	0.01775
17	0.01890	29	0.01767
18	0.01912	30	0.01755
19	0.01894	31	0.01778
20	0.01891	32	0.01773
21	0.01876	33	0.01758

#### 4. STEAM LOAD FORECAST

In this study, a prediction model was developed using MATLAB R2022a. The simulation focused on a tobacco manufacturing plant’s primary processing workshop, using seven months of data (from March to September 2023) with 131 sets of steam load data and corresponding influencing factors. The dataset was split into 91 training samples and 40 test samples.

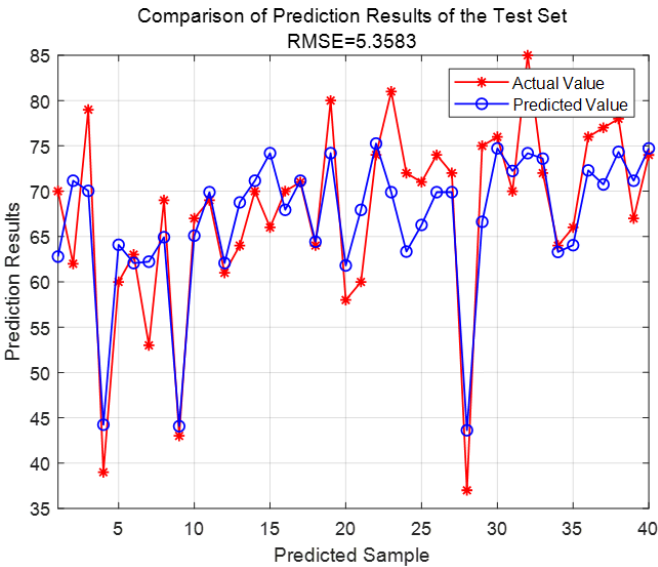
##### 4.1 Analysis of the Random Forest model



**Figure 4.** Random Forest model prediction results for the training set

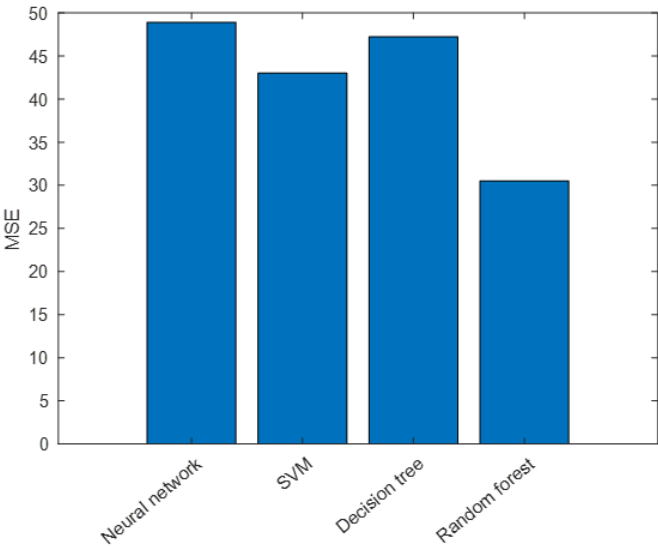
A Random Forest model was used to predict the steam load. The Root Mean Square Error (RMSE) was used to evaluate the prediction accuracy. The results showed that the RMSE for the training set was 6.3091, while the RMSE for the test set was 5.3583. The prediction results for both sets are shown in Figures 4 and 5, respectively. The RMSE calculation formula is as follows:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$



**Figure 5.** Random Forest model prediction results for the test set

##### 4.2 Analysis of different prediction models



**Figure 6.** MSE of different prediction models

In this paper, a Random Forest model was used for simulation testing. Its prediction accuracy was compared with that of a Decision Tree model, a Support Vector Machine (SVM) model, and a traditional BP neural network model [19-23]. The Mean Squared Error (MSE) was chosen as the evaluation metric, which measures the average squared difference between the predicted and actual values. The

comparison of MSE values for different models is illustrated in Figure 6. The MSE calculation formula is as follows:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

As shown in Figure 6, the Random Forest model exhibits higher prediction accuracy compared to the other three models. In particular, the BP neural network showed relatively poor performance on this dataset. To further validate the predictive performance of the proposed model, the following metrics were selected as evaluation criteria: Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and R-squared (R²). The calculation results are detailed in Table 4, with the formulas as follows:

$$MAE = \frac{1}{n} \sum_{i=1}^n |(y_i - \hat{y}_i)|$$

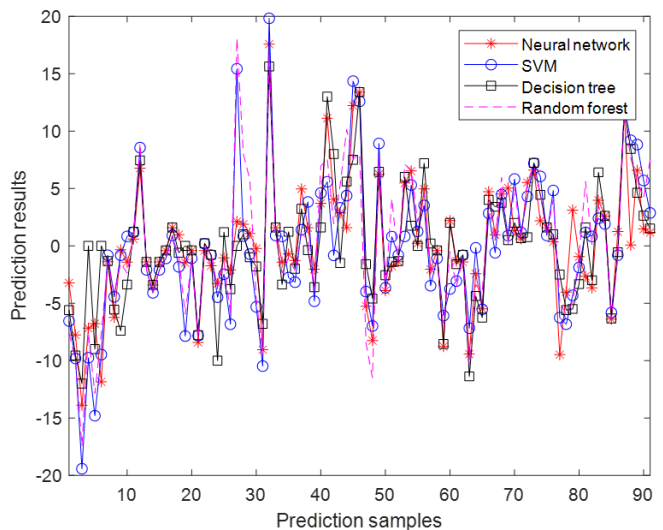
$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\%$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

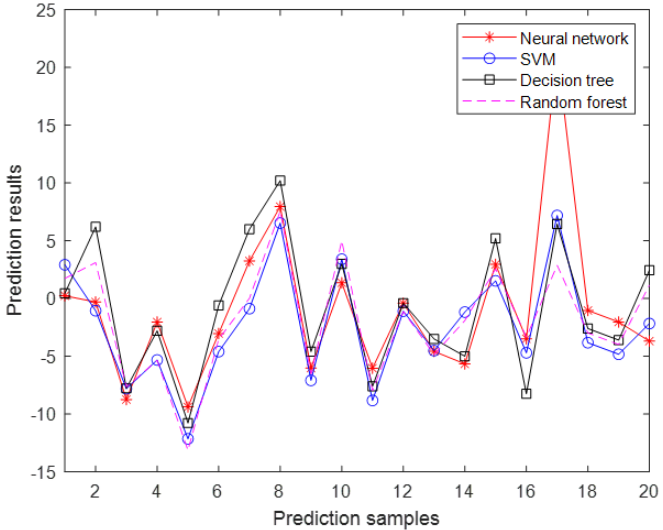
**Table 4.** Comparison of parameters for different models

Prediction Model	MAE	MAPE	R²
BP neural network	4.67	6.72%	0.271
SVM	4.58	6.29%	0.498
Decision tree	4.87	6.76%	0.392
Random Forest	4.32	5.87%	0.492

From Table 4, it can be seen that the Random Forest model has the lowest MAE at 4.32 and the lowest MAPE at 5.87%, outperforming the other prediction methods. The R² value is 0.492, which is the closest to 1 among the models, indicating that the Random Forest model has a better fitting effect. Figures 7 and 8 compare the prediction result errors for different models in the training and test sets, respectively, demonstrating that the BP neural network has a larger prediction error in this dataset, while the Random Forest model shows the smallest prediction error.



**Figure 7.** Comparison of prediction result errors for different models (training set)



**Figure 8.** Comparison of prediction result errors for different models (test set)

### 5. CONCLUSION

This paper analyzed the energy consumption of the steam system in a tobacco manufacturing plant's primary processing workshop, applying emerging technologies such as artificial intelligence and machine learning to predict daily steam consumption. Initially, the OOBPredictorImportance concept was introduced in the prediction model to assess the importance of each feature using the OOB dataset. Factors that significantly affect the steam load in the processing workshop were identified, including the vacuum area, leaf blade area, stem area (before storage), and stem area (after storage). Subsequently, the Random Forest method was employed to predict the daily steam load and compared with traditional BP neural network, SVM, and decision tree. The results demonstrate that the Random Forest model outperformed the others in both prediction accuracy and fitting effectiveness, providing a reliable basis for optimizing the scheduling of the enterprise's steam system. While the prediction results of the model generally achieved satisfactory outcomes, it is noted that the data exhibited significant fluctuations and the sample size was relatively small, which may limit the feature selection process. Therefore, future research should focus on exploring additional influencing factors and enriching the dataset to optimize the model and further enhance prediction accuracy, facilitating the regulation of steam supply and achieving energy-saving and emission-reduction goals.

### ACKNOWLEDGEMENTS

This paper was supported by Zhangjiakou Cigarette Factory Co., Ltd. Technology Project "Research on energy diagnosis and energy saving measures for boiler and heating system" (Grant No.: ZY012022F015).

### REFERENCES

[1] Baur, L., Ditschuneit, K., Schambach, M., Kaymakci, C., Wollmann, T., Sauer, A. (2024). Explainability and interpretability in electric load forecasting using machine

- learning techniques—A review. *Energy and AI*, 16: 100358. <https://doi.org/10.1016/j.egyai.2024.100358>
- [2] Cordeiro-Costas, M., Villanueva, D., Eguía-Oller, P., Martínez-Comesaña, M., Ramos, S. (2023). Load forecasting with machine learning and deep learning methods. *Applied Sciences*, 13(13): 7933. <https://doi.org/10.3390/app13137933>
- [3] Duan, N., Huang, Y. (2011). Measures for energy saving and emission reduction in Chinese tobacco industry. *Energy Procedia*, 5: 818-823. <https://doi.org/10.1016/j.egypro.2011.03.144>
- [4] Liu, Y., Liu, Q., Wang, W., Zhao, J., Leung, H. (2012). Data-driven based model for flow prediction of steam system in steel industry. *Information Sciences*, 193: 104-114. <https://doi.org/10.1016/j.ins.2011.12.031>
- [5] Wang, Q., Hu, Y.J., Hao, J., Lv, N., Li, T.Y., Tang, B.J. (2019). Exploring the influences of green industrial building on the energy consumption of industrial enterprises: A case study of Chinese cigarette manufactures. *Journal of Cleaner Production*, 231: 370-385. <https://doi.org/10.1016/j.jclepro.2019.05.136>
- [6] Aguilar Madrid, E., Antonio, N. (2021). Short-term electricity load forecasting with machine learning. *Information*, 12(2): 50. <https://doi.org/10.3390/info12020050>
- [7] Kusiak, A., Li, M., Zhang, Z. (2010). A data-driven approach for steam load prediction in buildings. *Applied Energy*, 87(3): 925-933. <https://doi.org/10.1016/j.apenergy.2009.09.004>
- [8] Fan, Z., Ren, Z., Chen, A. (2023). Multi-objective predictive control based on the cutting tobacco outlet moisture priority. *Scientific Reports*, 13(1): 199. <https://doi.org/10.1038/s41598-022-26694-x>
- [9] Alotaibi, M.A. (2022). Machine learning approach for short-term load forecasting using deep neural network. *Energies*, 15(17): 6261. <https://doi.org/10.3390/en15176261>
- [10] Han, M., Zhong, J., Sang, P., Liao, H., Tan, A. (2022). A combined model incorporating improved SSA and LSTM algorithms for short-term load forecasting. *Electronics*, 11(12): 1835. <https://doi.org/10.3390/electronics11121835>
- [11] Zhang, J., Qu, S., Zhang, Z., Cheng, S. (2022). Improved genetic algorithm optimized LSTM model and its application in short-term traffic flow prediction. *PeerJ Computer Science*, 8: e1048. <https://doi.org/10.7717/peerj-cs.1048>
- [12] Panda, S.K., Ray, P. (2022). Analysis and evaluation of two short-term load forecasting techniques. *International Journal of Emerging Electric Power Systems*, 23(2): 183-196. <https://doi.org/10.1515/ijeeps-2021-0051>
- [13] Al Mamun, A., Sohel, M., Mohammad, N., Sunny, M.S.H., Dipta, D.R., Hossain, E. (2020). A comprehensive review of the load forecasting techniques using single and hybrid predictive models. *IEEE Access*, 8: 134911-134939. <https://doi.org/10.1109/ACCESS.2020.3010702>
- [14] Chen, L., Cao, Z., Luo, J., He, X., Gao, J.L., Chen, H.X. (2022). Research on energy consumption prediction of air-conditioning system in silk storage workshop of a tobacco factory. *Refrigeration and Air Conditioning*, 22(8): 22-25. <https://doi.org/10.3969/j.issn.1009-8402.2022.08.005>
- [15] Zhao, L., Wu, S., Zhang, Q. (2022). Simulation modeling of steam system and production and consumption prediction in iron and steel enterprises. *Metallurgical Energy*, 41(4): 7-12. <https://doi.org/10.3969/j.issn.1001-1617.2022.04.002>
- [16] Song, X., Wang, Y., Liu, M. (2021). Application of machine learning algorithms in predicting individual annual income. *Information Technology and Informatization*, 8: 163-166. <https://doi.org/10.3969/j.issn.1672-9528.2021.08.048>
- [17] Yu, J., Bai, S., Wu, D. (2021). Research on predicting student performance in online teaching based on machine learning. *Computer Programming Skills and Maintenance*, 8: 118-119+154. <https://doi.org/10.16184/j.cnki.comprg.2021.08.047>
- [18] Bouktif, S., Fiaz, A., Ouni, A., Serhani, M.A. (2018). Optimal deep learning lstm model for electric load forecasting using feature selection and genetic algorithm: Comparison with machine learning approaches. *Energies*, 11(7): 1636. <https://doi.org/10.3390/en11071636>
- [19] Ahmad, W., Ayub, N., Ali, T., Irfan, M., Awais, M., Shiraz, M., Glowacz, A. (2020). Towards short term electricity load forecasting using improved support vector machine and extreme learning machine. *Energies*, 13(11): 2907. <https://doi.org/10.3390/en13112907>
- [20] Nagabushanam, P., Thomas George, S., Radha, S. (2020). EEG signal classification using LSTM and improved neural network algorithms. *Soft Computing*, 24(13): 9981-10003. <https://doi.org/10.1007/s00500-019-04515-0>
- [21] Dominiczak, K., Rządkowski, R., Radulski, W., Szczepanik, R. (2016). Online prediction of temperature and stress in steam turbine components using neural networks. *Journal of Engineering for Gas Turbines and Power*, 138(5): 052606. <https://doi.org/10.1115/1.4031626>
- [22] Hanmandlu, M., Chauhan, B.K. (2010). Load forecasting using hybrid models. *IEEE Transactions on Power Systems*, 26(1): 20-29. <https://doi.org/10.1109/TPWRS.2010.2048585>
- [23] Wang, W., Men, C., Lu, W. (2008). Online prediction model based on support vector machine. *Neurocomputing*, 71(4-6): 550-558. <https://doi.org/10.1016/j.neucom.2007.07.020>