International Journal of Multiphysics
Volume 18, No. 2, 2024
ISSN: 1750-9548

Prediction of Soybean Yield in Jilin Based on Diverse Machine Learning Algorithms and Meteorological Disaster Indices

Lifang Song[1], Ruixiang Song[2], Chengcheng Wang[2], Xinjing Zhang[2,*], Yunfeng Li[2], Lei Li[1], Mingyan Li[1], Meng Zhang[1]

[1]Jilin Meteorological Service Center, Changchun 130062, China

[2]Jilin Meteorological Information Network Center, Changchun 130062, China

*Corresponding Author.

Abstract

To enhance the precision and timeliness of soybean yield prediction in Jilin Province, this study leveraged data from 43 weather stations within the region. Employing methodologies such as Random Forest, Genetic Algorithm-assisted BP Neural Network, Support Vector Machine, and Convolutional Neural Network, predictive models for soybean yield were developed, with a specific focus on comparing the outcomes when including meteorological disaster index variables. The research findings highlight the paramount importance of certain feature variables closely linked to soybean yield, namely minimum temperature, accumulated temperature ≥10°C, mean temperature, frost-free days, maximum temperature, longitude, and precipitation. Notably, the integration of meteorological disaster index variables led to superior simulation results. Among the models utilizing these variables, the Random Forest model demonstrated the highest simulation accuracy, while the GA-BP and SVM models displayed relatively lower performance, with MAE values of 4.45 and 4.49 respectively. Conversely, the CNN model showcased the weakest performance in the context of this study. Ultimately, the collective models exhibited commendable accuracy levels.

Keywords: Jilin Province, soybean, yield prediction, meteorological disasters, machine learning.

1. Introduction

Currently, the primary soybean-producing regions in China are predominantly situated in the northeast region and the Huang-Huai-Hai Plain, with Jilin Province being a key soybean-producing province in the northeastern region. Research indicates that the Songnen Plain and Sanjiang Plain in the northeastern plains exhibit the greatest potential for increased soybean production [1, 2]. With the rapid development of the domestic economy, the domestic demand for soybeans has surpassed local production levels, leading to a widening gap between production and demand [3]. Therefore, for the local soybean industry's advancement, the timely and accurate prediction of soybean yields holds significant importance.

Nevertheless, soybean yields are influenced by a plethora of factors, including soil quality, seed varieties, pest infestations, and meteorological conditions, all of which significantly impact soybean production. Meteorological elements such as temperature, precipitation, sunlight, wind, drought, and flooding play crucial roles in the growth, development, and yield of soybeans [4]. Meteorological disasters represent a pivotal factor that cannot be overlooked in the soybean production process. Different types of meteorological disasters may detrimentally affect crop growth and development, potentially leading to substantial yield reductions [5, 6]. Currently, numerous scholars have investigated the impact of meteorological disasters on soybean yields. For instance, Wang and others [7] have highlighted the pronounced effects of floods and droughts on crop production in Heilongjiang Province, with hailstorms and low-temperature disasters demonstrating relatively stable impacts and causing minimal harm to agricultural production. Research by Yin et al. [8] reveals that soybeans encountering drought stress during the emergence stage post-sowing experience reduced yields, even mild droughts triggering unfavorable yield outcomes. Additionally, Gai and colleagues [9] note that frequent occurrences of unpredictable low-temperature damage pose obstacles to the stable production of soybeans, severely constraining the growth of soybean yields and the sustainable development of the soybean industry. Studies by Qu and others [10] indicate that low temperatures, excessive rainfall, and their combined effects can lead to decreased soybean yields, with significant variations in yield reductions. Furthermore, the Standardized Precipitation Evapotranspiration Index (SPEI), used to assess drought conditions by considering both precipitation and potential evapotranspiration effects, has been widely employed in drought evaluation within 1-6 month scales and can be utilized to forecast the yields of various crops [11, 12].

To ascertain the impact of meteorological factors and meteorological disasters on soybean yields and to quantify these effects, it is essential to establish a mathematical mapping relationship between meteorological factors and soybean yields. In recent years, with the continuous advancement of machine learning technologies, utilizing machine learning for crop yield prediction has become a prominent research focus [13, 14]. Models for yield simulation, such as those based on Random Forest, Genetic Algorithm-Backpropagation (GA-BP), Support Vector Machine (SVM), and Convolutional Neural Network (CNN) algorithms, have garnered widespread research attention and practical applications. Khanal et al. [15] employed remote sensing techniques to gather crop data under compacted conditions, subsequently integrating linear regression, Random Forest, and other machine learning algorithms to devise a remote sensing-based maize yield prediction model. Zhou et al. [16] utilized the GA-RF algorithm to establish a classification model for assessing the impact of soil compactness on soybean yield. Furthermore, Luo and colleagues [12] demonstrated that models based on Gradient Boosting Decision Trees (GBDT), Bidirectional Recurrent Backpropagation (BRBP), and SVM machine learning algorithms achieved significantly higher accuracy in simulating apple yields relative to models based on traditional linear regression algorithms. The majority of the aforementioned studies have been dedicated to investigating other crops or factors influencing soybean yields, with limited research examining the impact of meteorological disaster factors on soybean production.

This study is grounded in meteorological geographic spatial variables alongside augmented meteorological disaster variables. By employing Random Forest, Genetic Algorithm-enhanced Backpropagation Neural Networks, Support Vector Machines, and Convolutional Neural Networks, a predictive model for soybean yields was constructed. The precision of these models was compared, offering insights that could furnish decision-making information for large-scale agricultural management.

2. Data and Methodology

2.1 Study area and data sources

In this study, data from 43 meteorological stations in Jilin Province from 1978 to 2017 were utilized, encompassing daily meteorological variables such as maximum temperature (T_{max}), minimum temperature (T_{min}), mean temperature (T_{mean}), relative humidity (RH), precipitation (P), and sunshine hours (R), all sourced from the Jilin Provincial Meteorological Bureau. Additionally, geographical variables including longitude (Lon), latitude (Lat), and elevation (Dem) were sourced from respective observation sites, as depicted in Figure 1. Soybean yield data were sourced from the *Statistical Yearbook of Jilin Province*. Quality control measures were implemented, involving screening of meteorological and soybean yield data to eliminate missing data points, culminating in a total of 1,681 viable data points for model construction (comprising 1,200 samples for the training set and 481 samples for the test set). Prior to inputting data into the soybean yield prediction model, normalization procedures were undertaken to mitigate the impact of differing dimensions on model performance and enhance accuracy [17]. The study employed the min-max normalization method for data preprocessing, followed by inverse normalization to convert model predictions back to actual soybean yield values. The computational formula is as follows:

$$x_n = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{1}$$

Where,

x_n = the normalized input data;

x =the original input variable data;

$x_{\max}$ =the maximum input variable;

and $x_{\min}$ =the minimum input variable.

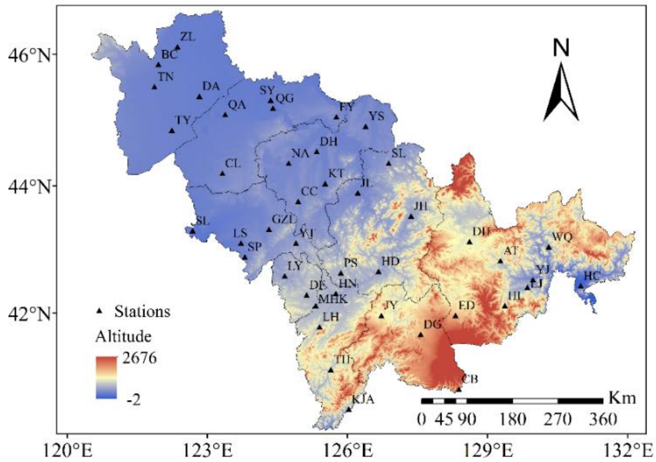

Figure 1 The spatial distribution of the 43 meteorological stations in Jilin Province

2.2 Characteristic variables for model input

The soybean growth period in Jilin Province spans from May to September [18]. Within this growth period, ten meteorological geographic spatial variables and two meteorological disaster indicators were selected as input variables for constructing the soybean yield prediction model, as outlined in Table 1.

Table 1 Characteristic variables for soybean yield prediction model input.

Characteristic Variables	Variables	Calculation methods
Meteorological geographic spatial variables	Mean temperature (T_{mean})	Mean temperature during the growing season
	minimum temperature (T_{min})	Average minimum temperature during the growing season
	maximum temperature (T_{max})	Average maximum temperature during the growing season
	Accumulated temperature ≥10°C (T_c)	Total sum of temperatures ≥ 10°C during the growing season
	Relative humidity (RH)	Arithmetic mean of relative humidity during the growing season
	Precipitation (P)	Total precipitation amount during the growing season
	Sunshine hours (R)	Total sunshine hours during the growing season
	Longitude	Longitude of the meteorological station
	Latitude	Latitude of the meteorological station
	Altitude	Altitude of the meteorological station
Meteorological disaster indicators	Frost-free days	Number of frost-free days per year
	Standardized precipitation evapotranspiration index	Thornthwaite method

Accumulated temperature refers to the total sum obtained by adding up the daily mean temperatures over a certain period, and it is utilized to investigate the relationship between temperature and the development rate of organisms. In Jilin Province, the accumulated temperature for active growth (≥10°C) between May and September ranges from 2100 to 3100 (°C·days), with the effective accumulated temperature for the entire growth period falling between 1700 and 2900 (°C·days). Most regions in Jilin Province meet the temperature requirements for soybean growth, yet certain western areas exhibit higher ≥10°C accumulated temperatures. Consequently, soybean cultivation in these regions may not yield optimal results.

This study employed the Standardized Precipitation Evapotranspiration Index (SPEI) to assess the supply and demand of moisture. SPEI is derived through the accumulation of the normalized difference sequence between precipitation and potential evapotranspiration. To calculate SPEI, monthly precipitation and monthly average temperature were utilized as input data to establish an accumulated sequence of water surplus or deficit based on the difference between precipitation and potential evaporation, encompassing various time scales. The Thornthwaite method was utilized to compute potential evapotranspiration (PET), which considers temperature variations comprehensively and requires minimal input variables, facilitating convenient calculations. This method effectively captures the characteristics of surface potential evaporation. The specific computational process referenced the approach outlined by Ma et al. [19].

Firstly, the calculation of potential evapotranspiration is conducted as follows:

$$PET = 16 \times (\frac{10T_i}{H})^A \tag{2}$$

$$H = \sum_{i=1}^{12} H_i = \sum_{i=1}^{12} (\frac{T_i}{5})^{1.514} \tag{3}$$

$A = 0.49 + 0.179H - 0.0000771H^2 + 0.000000675H^3$.

Where,

PET =potential evapotranspiration;

T_i=the monthly average temperature;

H= the annual heat index;

and A= a constant.

Subsequently, the difference between monthly precipitation and evapotranspiration is computed:

$$D_i = P_i - PET_i \tag{4}$$

Where,

D_i=the difference between precipitation and evapotranspiration;

P_i= the monthly precipitation;

and PET_i= the monthly evapotranspiration.

In the third step, as negative values may exist in the original data sequence, the difference sequence is standardized using a three-parameter log-logistic probability distribution to calculate the corresponding SPEI values:

$$F(x) = \left[1 + \left(\frac{\alpha}{x - \gamma}\right)^{\beta}\right]^{-1} \tag{5}$$

Where, parameters α , β and γ represent scale, shape, and location parameters, respectively, obtained through linear moment estimation:

$$\alpha = \frac{(\omega_0 - 2\omega_1)\beta}{\Gamma(1+1/\beta)\Gamma(1-1/\beta)} \tag{6}$$

$$\beta = \frac{2\omega_1 - \omega_0}{6\omega_1 - \omega_0 - 6\omega_2} \tag{7}$$

$$\gamma = \omega_0 - \alpha\Gamma(1+1/\beta)\Gamma(1-1/\beta) \tag{8}$$

Where: Γ= the factorial function; ω_0 , ω_1 and ω_2 are the probability weighted moments of the data sequence D_i :

$$\omega_s = \frac{1}{N}\sum_{i=1}^{N}(1-F_i)^s D_i \tag{9}$$

$$F_i = \frac{i-0.35}{N} \tag{10}$$

Where, N =the number of months.

Finally, the cumulative probability density was normalized to compute the SPEI:

$$P = 1 - F(x) \tag{11}$$

$$SPEI = w - \frac{c_0 + c_1 w + c_2 w^2}{1 - d_1 w + d_2 w^2 + d_3 w^3} \tag{12}$$

$$w = \sqrt{-2\ln(P)} \tag{13}$$

Where, the cumulative probability $P \leq 0.5$ and operated similarly to SPI calculations.

2.3 Machine learning algorithms

2.3.1 Random forest

Random Forest (RF) is a supervised ensemble learning algorithm belonging to the Bagging category, which enhances prediction accuracy by integrating multiple decision trees. This algorithm possesses robust anti-overfitting capabilities and high-precision prediction characteristics [14, 20]. Comprising multiple decision trees, each tree is generated by bootstrap sampling the dataset. During training, each decision tree learns features and patterns from different data subsets and predicts unknown data. Finally, the final prediction is derived by voting or averaging the predictions of all decision trees.

2.3.2 Genetic algorithm-backpropagation neural network (GA-BP)

Yang et al. [21] detailed the fundamental steps and principles of the GA-BP model. In this study, 70% of the overall effective data was randomly selected as training samples, while the remaining 30% was used for testing the GA-BP model's predictions. The average results of 50 modeling and prediction tests (including root mean square error and correlation coefficients) were calculated to quantitatively evaluate the accuracy and stability of the model constructed based on meteorological disaster indicator variables [12]. The parameters for the GA-BP model were set as follows: 5 hidden layer nodes, maximum genetic generations of 50, crossover probability of 0.7, mutation probability of 0.01, maximum training iterations of 1000, minimum target root mean square error of 0.001, and learning rate of 0.01.

2.3.3 Support vector machine (SVM)

Support Vector Machine (SVM) is a supervised learning algorithm aimed at finding the optimal hyperplane for best data classification [22, 23]. This hyperplane separates different classes in the dataset with the maximum margin, thus SVM classifies by seeking a hyperplane with the maximum margin. During training, SVM utilizes a set of labeled training data to construct a classification model by learning the features and patterns of the data. This classification model can then be used to predict the classification of new unknown data. SVM excels in handling high-dimensional data, addressing non-linear problems, and exhibiting strong generalization capabilities. It finds extensive application in pattern recognition, text classification, image recognition, and other fields.

2.3.4 Convolutional neural network (CNN)

Convolutional Neural Network (CNN) is a deep learning model particularly adept at processing data with grid structure, such as images and videos [24]. Its primary feature lies in extracting spatial structural features of data through convolution operations and reducing parameter volume and computational complexity through pooling operations. CNNs are constructed by stacking multiple convolutional and pooling layers, culminating in fully connected layers for classification or regression tasks. During training, CNN adjusts network parameters using a large labeled dataset through the backpropagation algorithm, enabling the network to learn features and patterns in the data.

2.4 Model evaluation

In this study, the evaluation of model simulation accuracy is conducted using three statistical metrics: the Pearson correlation coefficient (r), the Root Mean Square Error (RMSE), and the Mean Absolute Error (MAE).

3. Results and Analysis

3.1 Correlation analysis between soybean yield and variables

This study delves into the correlation between soybean yield and various variables (refer to Figure 2, significance level $P < 0.05$). The results reveal that temperature-related variables including average temperature (Tmean), maximum temperature (Tmax), minimum temperature (Tmin), and sunshine hours (Tc) exhibit a positive correlation with soybean yield. Specifically, the correlation between Tmin and soybean yield is most significant, with a coefficient reaching 0.25. Following closely are Tc and Tmean, showing relatively high correlation coefficients, while the correlation between Tmax and soybean yield is the weakest at a coefficient of only 0.16. Moreover, precipitation (P) demonstrates a positive correlation with soybean yield, whereas relative humidity (RH) and sunshine hours (R) exhibit lower correlation coefficients with soybean yield. Additionally, concerning meteorological disaster indicators, the number of frost-free days shows a positive correlation with soybean yield, with a coefficient of 0.11; the SPEI index also shows a positive correlation with soybean yield, albeit with a weaker correlation. Regarding geographical spatial variables, longitude demonstrates a negative correlation with soybean yield, while latitude and altitude exhibit a positive correlation with soybean yield. Notably, the correlation between longitude and soybean yield is the most significant, while the correlation with the other two variables is relatively low.

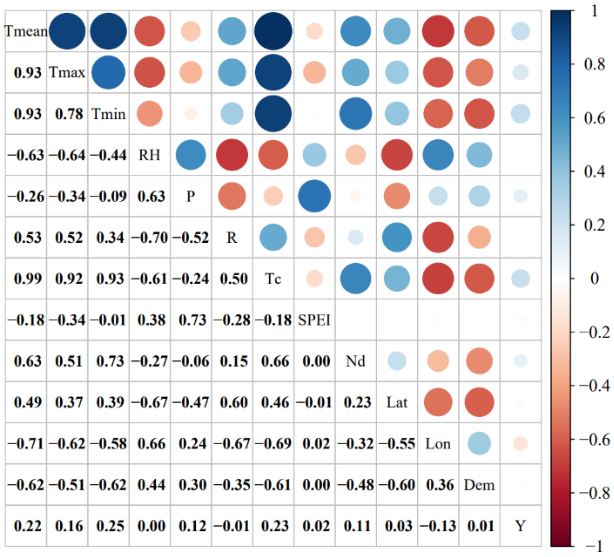

Figure 2 Heatmap illustrating the correlation between soybean yield and meteorological feature variables

3.2 Simulation accuracy analysis of different soybean yield models based on meteorological disaster index variables

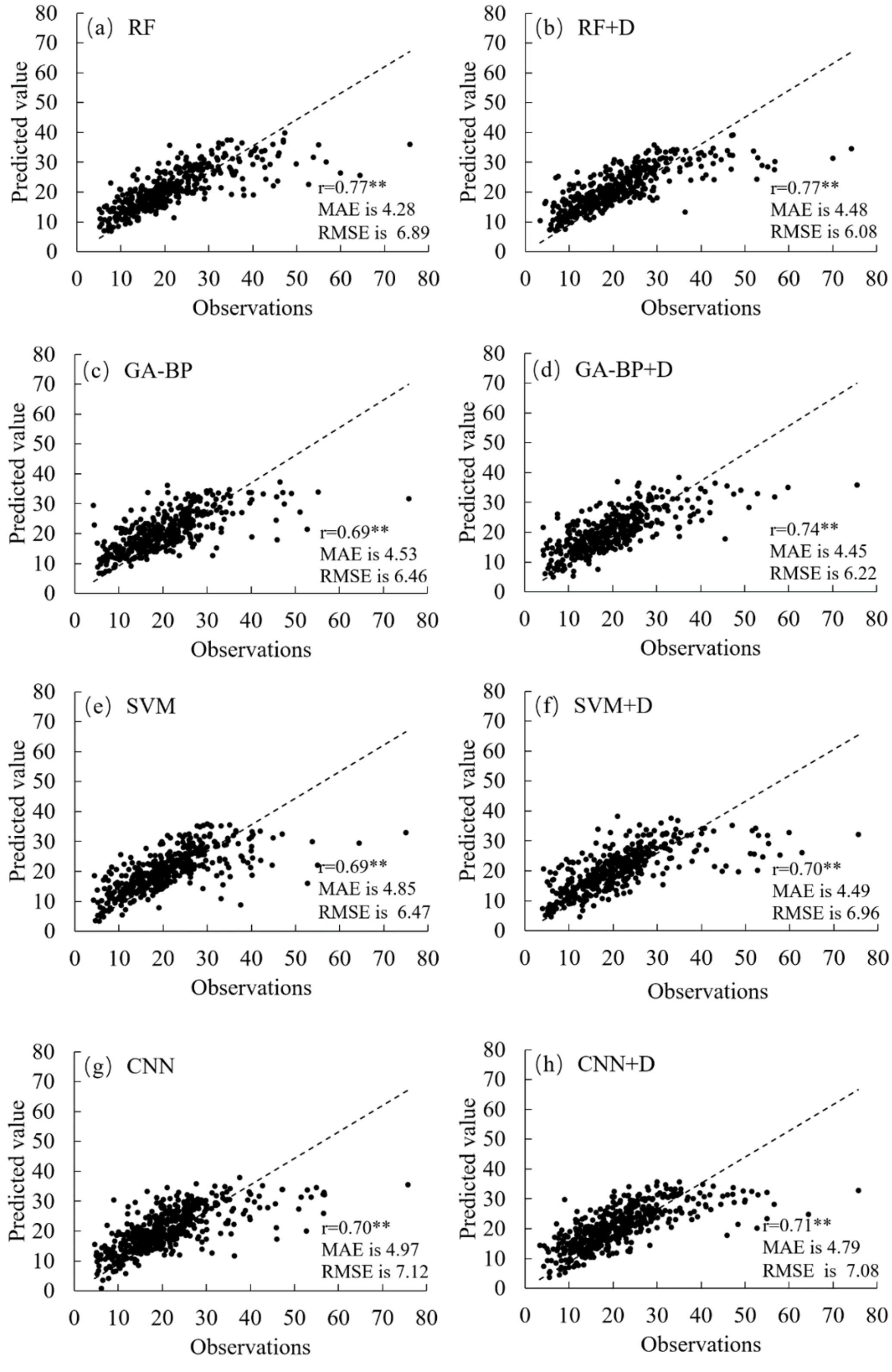

Figure 3 A comparison of the predictive accuracy of soybean yield simulation models constructed using four algorithms based on meteorological and geographical spatial variables ((a), (c), (e), (g)) versus models incorporating additional meteorological disaster variables ((b), (d), (f), (h)).

The fitting effects of soybean yield prediction models established using meteorological and geographical spatial factors via RF, GA-BP, SVM, and CNN algorithms are illustrated in Figure 3. As observed in Figure 3 ((a), (c), (e), and (g), where the dotted line represents the 1:1 line; asterisks denote $P < 0.01$), the CNN model, as depicted in Fig.3(g), exhibits the weakest predictive performance in the test set, with a correlation coefficient of 0.70 and the highest average absolute error and root mean square error at 4.97 and 7.12, respectively. The GA-BP neural network (Figure 3(c)) and SVM algorithm model (Figure 3(e)) yield similar results, with correlation coefficients of 0.69 each and RMSE values of 6.46 and 6.47, respectively. In terms of the mean absolute error, the GA-BP neural network model outperforms. The RF model (Figure 3(a)) distinctly outperforms the other models, with an r value of 0.77, MAE of 4.28, and RMSE of 6.89.

The soybean yield prediction models incorporating meteorological disaster factors demonstrate a significant enhancement in fitting effects. As shown in Figure 3 ((b), (d), (f), and (h)), compared to models solely considering meteorological and geographical spatial factors, the addition of meteorological disaster factors leads to improvements in the fitting effects of all models. Among these models, the soybean yield impact prediction model based on the Random Forest method (Figure 3(b)) performs the best, with the lowest root mean square error (RMSE). Compared to the model considering only meteorological and geographical spatial factors, the RMSE decreases by 11.8%. The correlation coefficient remains the same, while the MAE slightly increases. Additionally, by incorporating meteorological disaster factors, the simulation effects of the GA-BP neural network (Figure 3(d)) and SVM models (Figure 3(f)) also improve. The correlation coefficient increases by 7.2% and 1.4%, MAE decreases by 1.8% and 7.4%, and RMSE values are 6.22 and 6.96 respectively. Despite the CNN model's (Figure 3(h)) poorer fitting ability compared to other models, there is still some improvement. Its correlation coefficient r increases by 1.4%, MAE decreases by 3.6%, and RMSE increases by 1%.

In conclusion, considering the impact of meteorological disasters on soybean yield leads to enhanced model accuracy. Although CNN, as a deep learning algorithm, did not exhibit the best performance in this context, as it is more suited for processing grid-structured data like images and videos. In contrast, ensemble learning algorithms (such as Random Forest) are better at handling complex relationships between data, enhancing model predictive performance and thus demonstrating superior performance in soybean yield impact models. Meanwhile, the GA-BP neural network and SVM belong to genetic neural networks and supervised learning algorithms, respectively. Through comparing these three types of algorithms, we find that ensemble learning algorithms are more suitable for studying the impact of meteorological disasters on soybean yield.

4. Conclusions

This study, based on meteorological and geographical spatial variables as well as meteorological disaster variables, utilized Random Forest, Genetic Algorithm-based BP Neural Network, Support Vector Machine, and Convolutional Neural Network methods to establish predictive models for soybean yield, yielding favorable prediction outcomes. The following conclusions were drawn: 1. The feature variables with the highest correlation to soybean yield are, in descending order, minimum temperature, accumulated temperature above 10°C, average temperature, frost-free days, maximum temperature, longitude, and precipitation. Variables such as latitude, SPEI index, sunshine hours, elevation, and relative humidity exhibit lower correlations. 2. Incorporating meteorological disaster index variables improves the fitting effects of the models compared to models considering only meteorological and geographical spatial factors. 3. In terms of simulating with meteorological disaster index variables, the RF model demonstrates high precision, with a correlation coefficient r of 0.77, root mean square error (RMSE) of 6.08, and mean absolute error (MAE) of 4.48. The GA-BP and SVM models show slightly lower performance, with correlation coefficients r of 0.74 and 0.70, RMSE values of 6.22 and 6.96, and MAE values of 4.45 and 4.49, respectively. The CNN model exhibits the weakest results, with a correlation coefficient r of 0.71, RMSE of 7.08, and MAE of 4.79. By comparing these algorithms, it is evident that ensemble learning algorithms are better suited for investigating the impact of meteorological disasters on soybean yield.

References

[1] Liu Aimin, Feng Zhiming, Yan Lizhen, Yu Ge. A Study on the Throughput and Balance of Supply and Demand of Soybean in China. Chinese Journal of Agricultural Resources and Regional Planning, 2003, 24(4): 36-39.

[2] Liu Aimin, Feng Zhiming, Yan Lizhen, Yu Ge. Study on soybean throughput based on cultivated land resources restriction in China. Journal of Natural Resources, 2003, 18(4): 430-436.

[3] Wang Chen, Chu Lin, Yang Zhe, Yang Zhenhao, Zhang Xinya, et al. Spatial heterogeneity and determinants of soybean yield in Northeast China. Transactions of the Chinese Society of Agricultural Engineering, 2023, 39(21): 108-119.

[4] Yan Ping, Jiang Lixia, Wang Ping, Nan Rui, Wang Chenyi, et al. Analysis of the illumination effect on the soy output from May to July. Heilongjiang Meteorology, 2005, (4), 2.

[5] Zhang Ping. Analysis of Agriculture Natural Disasters in Heilongjiang Province. Journal of Agricultural Mechanization Research, 2011, 33(2): 249-252.

[6] ZHAO J, YANG X G, Lu S, Liu Z J, WANG J. Variability of available climate resources and disaster risks for different maturity types of spring maize in North-east China. Regional Environmental Change, 2014, 14(1): 17-26.

[7] Wang Qiujing, Li Xiufen, Yan Ping, Lyu Jiajia, Wang Liangliang, et al. Main Agro-meteorological Disasters Heilongjiang: Sequential Characteristics and Grey Correlation Analysis of Their Effects on Soybean Yield. Chinese Agricultural Science Bulletin, 2020, 36(3): 81-87.

[8] Yin Shiping, Zhu Haixia, Zhao Huiying, Sun Anhui, Gong Lijuan, et al. Breeding Traits and Yield Factors of Soybean: Response to Soil with Different Drought Grades. Chinese Agricultural Science Bulletin. 2020, 36(28), 111-117.

[9] Gai Zhijia, Liu Jingqi, Cai Lijun, Du Jiaxing, Gu Wei, et al. Cold Damage to Soybean and Its Prevention and Control Measures: Research Progress. Journal of Agriculture. 2021, 11(1): 7-10.

[10] Qu Huihui, Li Xiufen, Zhu Haixia, Wang Liangliang, Qu Bingyang, Wang Qiujing, Lyu Jiajia, Ji Yanghui, Jiang Lixia. Influence of Low-Temperature Rainfall in the Seedling Stage on Soybean Yield in High-Cold Areas. Journal of Ecology. [URL](https://kns.cnki.net/kcms2/detail/21.1148. Q.20230712.1900.026.html)

[11] SHI L J, FENG P Y, WANG B, LIU D L, YU Q. Quantifying future drought change and associated uncertainty in southeastern Australia with multiple potential evapotranspiration models. Journal of Hydrology, 2020, 590: 125394.

[12] Luo Qi, Ru Xiaoya, Jiang Yuan, Feng Hao, Yu Qiang, et al. Prediction of Apple Relative Meteorological Yields Based on Machine Learning and Meteorological Disaster Indices. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(9): 252-264.

[13] Tang Zijun, Zhang Wei, Huang Xiangyang, Xiang Youzhen, Zhang Fucang, Chen Junying. Soybean Seed Yield Estimation Model Based on Ground Hyperspectral Remote Sensing Technology. Transactions of the Chinese Society for Agricultural Machinery. [URL](https://link.cnki.net/urlid/11.1964.s.20231122.1547.005)

[14] Zhou Xiuli, Qin Na, Wang Kaiyu, Sun Hao, Wang Dawei, et al. Effect of Mechanical Compaction on Soybean Yield Based on Machine Learning. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(11): 139-147.

[15] Khanal S, Klopfenstein A, Kushal K C, Ramarao V, Fulton J, et al. Assessing the impact of agricultural field traffic on corn grain yield using remote sensing and machine learning. Soil and Tillage Research, 2021, 208: 104880.

[16] Zhou Xiuli, Zhang Pingping, Qin Na, Huo Dongxu, Qiao Jinyou. Effects of soil penetration resistance on soybean yield in black soil region based on GA-RF model. Journal of Northeast Agricultural University, 2022, 53(10): 67-75.

[17] BRAY M, HAN D. Identification of support vector machines for runoff modelling. Journal of Hydroinformatics, 2004, 6(4): 265-280.

[18] Cao Ling, Xu Shiqi, Yang Xueyan, Liu Baixin, Qiao Yike, et al. Climate Characteristics of Northeast Cold Vortex in Summer and the Effects on Growth Period of Soybean in Jilin Province. Desert and Oasis Meteorology, 2022, 16(15): 135-143. DOI: 10.12057/j.issn.1002-0799.2022.05.018.

[19] Ma Jingzhao, Hao Lu. Temporal and spatial variation of drought in the Xilingol grassland based on the standardized precipitation index and standardized precipitation evapotranspiration index. Pratacultural Science, 2021, 38(12): 2327-2339.

[20] Feng Jianying, Su Yunhui, Gong Shaoqi, Wang Zhi, Mu Weisong. Evaluation Method of Agricultural Production Technology Efficiency based on Ensemble Learning. Journal of Agricultural Machinery, 2021, 52(Supp.): 148-155.

[21] Yang Dong, Zhu Jiamin, Ding Yeyi, Jin Zhifeng. Responses of phenology, quality and yield of juicy peach to climate change in Zhejiang Province. Chinese Journal of Ecology, 2019, 38(11): 3366-3375.

[22] PONTIL M, VERRI A. Properties of support vector machines. Neural Computation, 1998, 10(4): 955-974.

[23] SMOLA A J, SCHÖLKOPF B. A tutorial on support vector regression. Statistics and Computing, 2004, 14(3):199-222.

[24] Fu Hao, Zhao Xueguan, Zhai Changyuan, Zheng Kang, Zheng Shenyu, et al. Research progress on weed recognition method based on deep learning technology. Journal of Chinese Agricultural Mechanization, 2023, 44(5): 198-207.