Turkish Journal of Agriculture and Forestry

Volume 47 | Number 6 Article 15

12-1-2023

Energy crop yield simulation and prediction system based on machinelearning algorithm

JIE ZHANG

ZHIDONG LIU

Recommended Citation

ZHANG, JIE and LIU, ZHIDONG (2023) "Energy crop yield simulation and prediction system based on machinelearning algorithm," *Turkish Journal of Agriculture and Forestry*: Vol. 47: No. 6, Article 15. https://doi.org/10.55730/1300-011X.3142
Available at: https://journals.tubitak.gov.tr/agriculture/vol47/iss6/15

Turkish Journal of Agriculture and Forestry
http://journals.tubitak.gov.tr/agriculture/
Research Article
Turk J Agric For
(2023) 47: 972-982
© TÜBİTAK
doi:10.55730/1300-011X.3142

Energy crop yield simulation and prediction system based on machine learning algorithm

Jie ZHANG, Zhidong LIU*
[1]Electrical and Information Engineering College, Jilin Agricultural Science and Technology University, Jilin, China

Received: 27.03.2023 • **Accepted/Published Online:** 11.07.2023 • **Final Version:** 01.12.2023

Abstract: The yield of energy crops has been widely questioned by the public, but few researchers have analyzed the yield prediction of these crops, which has greatly limited their distribution and use. Based on this, in this study, a machine learning algorithm was used to design an energy crop yield simulation and prediction system, and this system was used to analyze the yields of four common energy crops. First the factors that affect the yield of energy crops are discussed, and then, the analysis of the machine learning algorithm and its application in the yield prediction of energy crops is presented, followed by an introduction to the framework of the energy crop yield simulation and prediction system. At the end of this paper, the effect of the energy crop yield simulation and prediction system was analyzed, and the feasibility conclusion was finally reached. The energy crop yield simulation and prediction system designed in this paper can greatly improve the accuracy of yield prediction when forecasting and analyzing energy crop yields. The sugarcane yield measured by the back propagation (BP) neural network model was 8694 kg/hm^2, with a relative deviation of 0.86% from the actual. The BP neural network algorithm is widely used in energy crop yield prediction, and can greatly improve the accuracy. In the future, with the maturity and development of a BP neural network algorithm, it would promote the development of crop yield prediction.

Key words: Energy crops, machine learning algorithms, BP neural network, yield prediction

1. Introduction

Predicting crop trends is an important prerequisite for ensuring future food security and responding to climate change. Crop models are important tools for simulating and predicting large-scale yield, but they contain great uncertainty. Previous studies have focused on the sources of uncertainty of crop model inputs and parameters, but less attention has been paid to the uncertainty caused by the model structure, especially the systematic study of the differences and similarities of the performance of different types of models. Because few researchers have analyzed crop yields, there is currently a lack of data on energy crop yield prediction, and measures need to be taken to further analyze and study it.

Crop yield prediction is not a new topic, and many scholars have made contributions in this regard. Shahhosseini (2021) investigated whether the coupling of crop modeling and machine learning improved corn yield prediction in the US corn belt. The main purpose was to explore whether the hybrid method would bring better prediction results. Kamath (2021) provided a quick check of agricultural yield prediction using the random forest method. Peng (2018) evaluated the benefits of using seasonal climate prediction and satellite remote sensing data to predict US corn yield at the national and county levels. Zhu et al. (2019) proposed an improved reinsurance pricing framework, including a crop yield forecasting model. The model used a new credibility estimator and closed reinsurance pricing formula to integrate weather variables and crop production information from different geographically relevant regions. Heino (2020) made use of the crop yield data of the global grid crop model set for a series of crop management scenarios simulation. Devika and Ananthi (2018) used data mining technology to predict the annual yield of main crops. However, few researchers have used machine learning algorithms to analyze and study crop yields.

The machine learning algorithm has a wide range of applications in forecasting data, including not only crop yield, but also electricity price forecasting. Palanivel and Chellammal (2019) analyzed the method of using machine learning and big data technology to predict crop yield. Garanayak (2021) used crop recommendation systems with different machine learning regression methods to calculate crop yield recommendations by accurately comparing many machine learning regression methods. Andrianov

* Correspondence: liuzhidong@jlnku.edu.cn

(2018) proved that the memory cycle artificial networks can not only accurately estimate the multi-phase rate of the current time (acting as a virtual flowmeter), but can also predict the rate of a series of time moments in the future. Naumzik and Stefan (2021) used machine learning to predict electricity prices. They analyzed the sensitivity of the predictor. However, there are currently no studies on the design of an energy crop yield simulation and prediction system in the research of machine learning prediction data.

In order to improve the accuracy of energy crop yield prediction, the back propagation (BP) neural network method was used to design the energy crop yield simulation and prediction system. In this paper, four kinds of energy products were analyzed by a remote sensing yield estimation model, multiple regression and factor analysis model, grey theory model, and BP neural network model. Finally, it was determined that the output of the sorghum, cassava, corn, and sugarcane measured by the energy crop yield simulation and prediction system using the BP neural network was the closest to the actual value, and its relative accuracy was the highest. Then, a simulation and prediction system for energy crop yield was designed to analyze crop yield, which has certain reference value.

2. Factors affecting the output of energy crops

The environmental impact of energy crops comes from the environmental impact of the planting, processing, and use of energy crops, resulting in changes in the structure and function of ecosystems, such as natural, chemical, and physical environmental impacts (Vera, 2021). The development and use of energy crops have improved the structure and function of ecosystems in many ways. In terms of natural environmental impact, it can reduce the input and discharge of pollutants in agricultural production, increase the recycling of waste, improve the ecological environment, and reduce energy pressure. It promotes urban development and resource efficiency. In terms of physical and chemical environmental impacts, it can provide clean energy, reduce dependence on fossil fuels, reduce acid rain and greenhouse gas emissions, etc. Sustainable energy crops can also have a positive impact on the environment. At the same time, the development of energy crops would have some negative impacts on the environment. The growth of energy crops would affect the surrounding vegetation and groundwater, the processing of energy crops would consume fossil energy and cause secondary pollution, the competition of energy crops for agricultural fertilizers, and the freshwater crisis. The more energy crops are produced, the greater the impact on human survival and development. In this paper, the factors affecting the yield of energy crops were analyzed at two levels: natural factors and human factors, as shown in Figure 1.

2.1. Impact of natural factors on energy crop yields

The main natural factors affecting the output of energy crops are heat, light, rainfall, and precipitation, which have a significant impact on agricultural production and affect the expansion and development of agriculture (Fahlbusch, 2018). Different crops need different growth environments and climatic conditions, and there are obvious geographi-

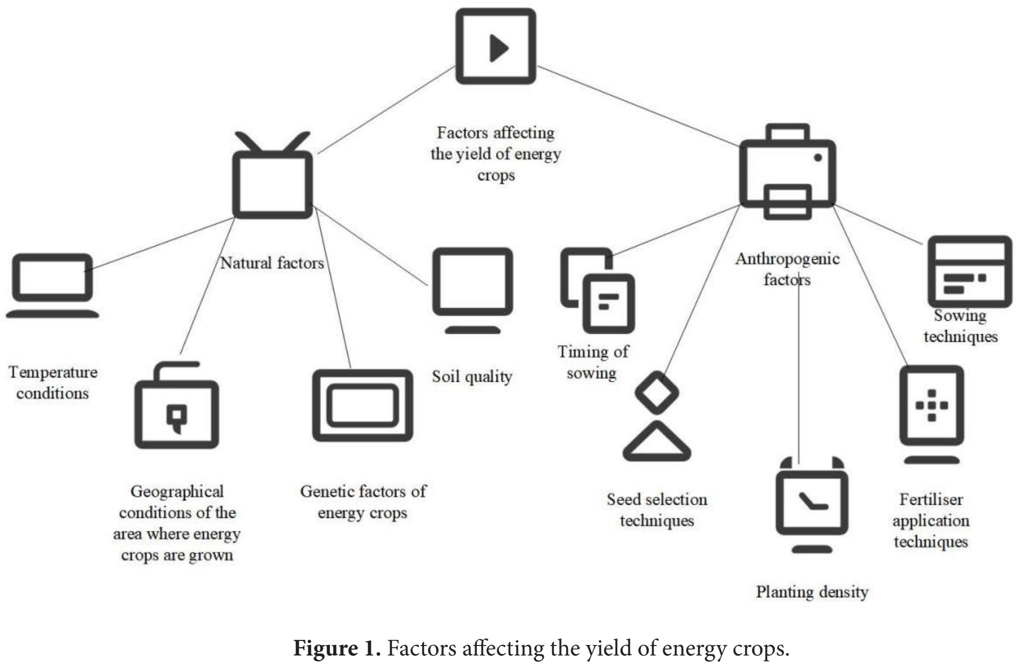

Figure 1. Factors affecting the yield of energy crops.

cal differences in climate distribution. Therefore, the selection of energy crops in specific areas must be coordinated with climate factors. The demand for heat and sufficient light is determined by the plant's own genes, which determine whether they lack light or heat. Light intensity affects the growth and development of energy crops. Different plants need different light intensities. Most energy crops are more suitable for growing in full sunlight, but some plants can grow normally in less light. Rainfall and actual rainfall: Rainfall depends on soil texture, soil characteristics, rainfall intensity, soil moisture at that time, water demand of energy crops and other factors.

2.1.1. Influence of temperature on energy crop yields

Temperature is the most important factor affecting the yield of energy crops (Paschalidou et al., 2018). The high temperature and sunshine in the daytime promote the photosynthesis of energy crops and increase the nutrient yield. The low temperature at night reduces the respiration of energy crops and energy intake, which is beneficial to the accumulation of nutrients.

2.1.2. Influence of the soil quality on energy crop yields

Soil is an important basis for the healthy growth and yield improvement of energy crops, so it is very important to improve soil quality. Soil nutrients and pH have a direct impact on the growth of energy crops, so scientific planning should be carried out at the initial stage of energy crop planting. Moreover, energy crops need a certain degree of soil permeability, so fields should be ploughed and fertilized regularly to ensure that the soil is conducive to the healthy growth of these crops. In order to improve the soil quality, organic fertilizer should be used as well as modern irrigation methods to irrigate the soil. The addition and supply of water and nutrients provide a solid foundation for the production of high-yield energy crops.

2.1.3. Influence of genetic factors on energy crops

The genetic determinants of energy crops have a great impact on the yield. The taste and internal components of energy crops are controlled by plant genetics. If the genetics of energy crops do not include sufficient high-yield genes, it is difficult to achieve the total yield in agricultural production.

2.1.4. Influence of the geography of the energy crop growing area

Geographic location is an important factor in determining the yield of energy crops. When the same types of energy crops are planted under different geographical conditions, the yields of these crops are different. In high-latitude and high-altitude areas, the sugar content of crops is high due to the low temperature, long duration of sunshine, and large diurnal variation. The yield and quality of energy crops planted in different seasons also have significant differences.

2.2. Influence of human factors on energy crop yields

2.2.1. Influence of the sowing time on energy crop yields

The seeding of energy crops is affected not only by the soil quality, but also by the seeding time. Different types of energy crops have different requirements in regard to sowing time, temperature, and humidity. It is also important to select the right seeding method for energy crop varieties. Scientific seeding enables energy crops to withstand the challenges of temperature, humidity, and adverse weather conditions in the subsequent growth process. The sowing date has a direct impact on the germination and growth rate of energy crops. Only by adjusting the sowing date according to the field environmental conditions can the seed germination rate be increased, and thus, the yield be increased.

2.2.2. Influence of seed selection techniques on energy crop yields

Before formally planting energy crops, farmers often have to go through the proper seed selection process. The quality of the screening process has a significant impact on the overall crop yield. When making the selection, the grower must take into account the impact of the natural environment on the growth of energy crops and try to select seeds that are compatible with the planting site and natural environment. Scientific methods are also needed to ensure the quality of energy crop seeds. For example, some seeds can be tested before production to check the quality and yield of the energy crop and eliminate seeds that do not meet the actual needs of crop production. In addition, this process removes as many diseased seeds as possible to prepare for future crop production.

2.2.3. Influence of sowing techniques on energy crop yields

Sowing methods used in agricultural production also have a great impact on the yield of agricultural products. When planting agricultural products, producers need to ensure that the planting density and time are appropriate. The most suitable sowing time should take into account the geographical conditions of the planting area and the current situation of energy crops. With regard to the planting density, it is necessary to check the leaf density between energy crops, taking into account the specific varieties of the energy crops to be planted.

2.2.4. Impact of fertilizer application techniques on energy crop yields

Fertilization also affects the yield and quality of energy crops. The specific amount and interval of fertilization varies for different energy crops. The farmers' choice of fertilizer also leads to differences in the yield and quality of energy crops.

2.2.5. Effect of the planting density on energy crop yields

Energy crop varieties have different requirements for plant density. The plant density is also closely related to their ability to better carry out photosynthesis and whether they are well ventilated. Therefore, in order to ensure high yield, attention should also be paid to controlling the plant density of energy crops to ensure optimal photosynthesis. Taking sorghum as an example, high plant density would lead to insufficient nutrition and serious consumption of soil resources. Insufficient planting density would also affect the yield of sugarcane. If the planting density is too high, the sunlight would be insufficient. If the planting density is too low, the yield per plant may be relatively high, but the total amount of sugarcane planted in the same area would decrease, which would ultimately affect the total yield of sugarcane.

3. Machine learning algorithm and its application in energy crop yield prediction

3.1. Introduction to machine learning algorithms

Machine learning allows computers to learn and process problems in the same way as humans, so people must pay attention to how they use machine learning algorithms to solve real-world problems (Greener, 2022). Machine learning has a wide range of research and application scenarios. It is an interdisciplinary field, including the results of artificial intelligence, information theory, probability and statistics, cybernetics, and other disciplines.

The neural network algorithm is an improved BP algorithm used in crop yield prediction, which is a self-learning algorithm used to build a three-line BP neural network (a-b-c) (Zhao, 2022). A is the input neuron, b is the hidden layer neuron, and c is the output neuron. The more layers, the easier it is to solve more complex and nonlinear problems. The yield of crops is affected by various uncertain factors, which is a complex nonlinear system. As for the artificial neural network analysis method, BP is a nonlinear system with only one input and one output.

The BP neural network is a multilayer feedforward neural network. The BP neural network is a basic network that contains the best and most complete neural network content. Each neuron in the input layer is responsible for receiving information from the outside world and sending it to the neurons in the hidden layer. The hidden layer is the internal information processing layer, responsible for transforming information. According to the needs of the information processing capability, the hidden layer can be designed as one or more hidden layers. Finally, the hidden layer is transferred to the output layer to complete the forward learning process. If the actual expected value is not equal to the expected value in the output, it adjusts the weight of each layer according to the reduction of the error gradient and completes the return of the hidden layer and the input layer.

The number of nodes in the hidden layer of the three-layer BP network determines the success or failure of the model. In this paper, the samples were divided into training samples and test samples, which were used to test and improve the network. The BP network can gradually increase or decrease the number of nodes in the hidden layer according to the results to adjust the network. Once the hidden layer node is determined, the network is trained and tested. If the error is within a reasonable range, it can be used for predictive testing. In a word, BP neural network algorithm is widely used in the analysis of energy crop yield.

The BP neural network algorithm consists of two processes: forward signal propagation and backward error propagation. The process of forecasting energy crop yield is as follows: The incoming energy crop yield information is first transferred to the hidden layer for layer-by-layer processing, and then the output information is transferred from the hidden layer to the output node. If the output layer cannot achieve the expected result, it will conduct backscattering to minimize the error signal generated. This results in a final satisfactory weight distribution result. The weight adjustment process is also the learning and training process of the network. This process continues until the output error of the network is reduced to an acceptable level, or until the scheduled training times are completed.

The target of input energy crops is set as in Eq. (1) below:

$$B_z = (b_1, b_2, \cdots b_m)(z = 1,2, \cdots, n). \tag{1}$$

Here, n is the logarithm of learning mode and m is the number of input layer units.

The expected output vector for the input energy crop target is as shown in Eq. (2):

$$X_z = \left(x_1, x_2, \cdots x_p\right)(z = 1,2 \cdots, n) \tag{2}$$

Here, p is the number of output layer elements.

According to the idea of mode forward propagation, the input and output of each unit in the output layer are as shown below in Eqs. (3) and (4):

$$O_k = \sum_{i=1}^{n} w_{ik} c_i - x_k\,, \tag{3}$$

$$D_k = f(O_k)(k = 1,2, \cdots p) \tag{4}$$

The weight vector and threshold of energy crop analysis are modified using the gradient method, as given in Eq. (5) below:

$$\Delta w(n+1) = -\eta \frac{\alpha F}{\alpha w(n)} + \beta \Delta w(n) = w(n+1) - w(n) \tag{5}$$

Here, η is the learning rate or learning factor of the BP neural network, and β is the momentum factor.

The general error given in Eq. (6):

$$F_t = \sum_{i=1}^{p} (x_k^t - D_k^t)^2/2 \tag{6}$$

The error of BP neural network is as given in Eq. (7):

$$F = \sum_{t=1}^{n} F_t = \sum_{t=1}^{n}\sum_{k=1}^{p} (x_k^t - D_k^t)^2/2 \qquad (7)$$

3.2 Application of the BP neural network algorithm in energy crop yield prediction

The generalization ability of the BP neural network means that the design of the pattern classifier must consider whether the network can correctly classify the required classification objects, and whether the network can correctly classify the invisible or noise-polluted patterns after training. In other words, the BP neural network can apply the learning results to new information. If the local or partial neurons of the BP neural network are damaged, it would have little impact on the global learning results; that is, even if the local neurons are damaged, the system can continue to operate normally. In other words, the BP neural network has a certain degree of fault-tolerance.

The BP neural network model has important advantages in complex nonlinear systems. It does not need to generate complex mathematical models but can generate complex relationships that occur after training. It has high adaptability and fault-tolerance and allows data input with certain noise. It has a distributed and parallel storage mode. It has the adaptability of nonprogramming, self-organization, and computation. Therefore, the neural network is particularly suitable for data processing that is difficult to express by traditional calculation methods, and its application in energy efficiency prediction is very effective. Its advantages are summarized in Figure 2.

4. Framework of the energy crop yield simulation and prediction system

4.1. Design of the energy crop yield simulation and forecasting system

The energy crop yield prediction system is an information system based on many agrometeorological crop prediction models in China. The system combines energy output prediction and graphical management, and presents the prediction results in graphical form, which is highly scalable and inheritable.

4.2. General structure of the energy crop yield simulation and forecasting system

The energy crop yield simulation and prediction system is mainly composed of a data layer, model layer, and application layer. The data layer mainly manages the data used for model-based application, assimilation, or prediction. The model layer includes application interfaces of different crop models, powerful model calculation methods, and spatial analysis tools. The application layer mainly deals

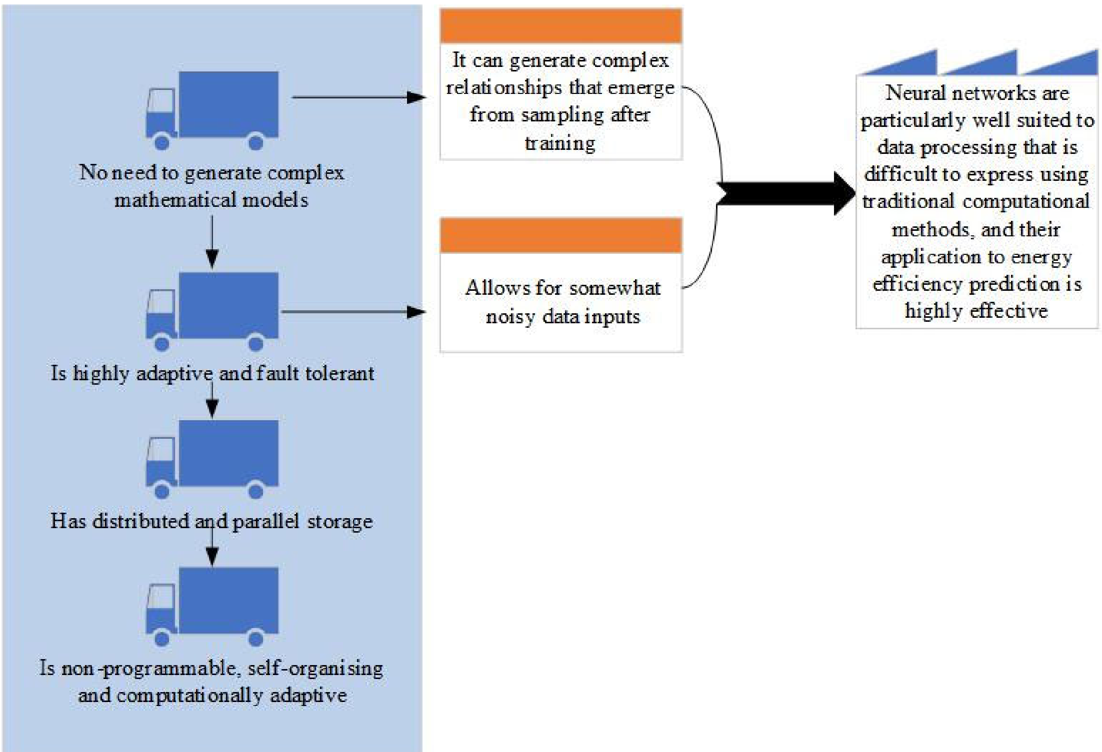

Figure 2. Advantages of BP neural network algorithms for energy crop yield prediction.

with tools used to run crop models in operational applications. The overall framework of the system is given in Figure 3.

4.2.1. The data layer

The data layer is mainly used to manage and pre-process data from different sources. The meteorological data are interpolated to the points in the spatial grid using spatial interpolation technology. The soil parameters in the field are mainly from agrometeorological stations, while the soil parameters in space are from a national database that provides very accurate soil parameters. The main soil parameters include the saturated water content, wilting coefficient, and field water holding capacity. Once the region parameters are obtained, they are allocated to a region using the Tyson polygon expansion method.

4.2.2. The model layer

The calculation of the model requires the use of powerful computing resources and spatial analysis. The model layer is an open interface, which can easily integrate other crop models to merge several models. The model relies on daily weather data as the driving force and combines soil parameters and farm management information to simulate crop growth dynamics.

4.2.3. The application layer

The model results are used in the application layer of the system for crop growth monitoring, agrometeorological risk early warning and impact assessment, yield prediction and other agrometeorological services.

4.2.4. Data management and system integration

The energy crop simulation and prediction system include a database, a crop modeling module, an algorithm module for commercial agrometeorological applications, an interactive graphical user interface, and data summary software file. The database mainly includes meteorological station and network database, spatial soil database, agrometeorological database, earth observation database, remote sensing database and weather forecast database. All of the databases are based on the agrometeorological system database of the National Meteorological Center. In the crop model module, the application programming interface design method can be used to design the data generation, parameter setting, model operation, and model result processing of different crop models in the system. Due to the flexible application programming interface model, it can flexibly integrate other crop models into the system. A plug-in-based network is used for spatial search, analysis, and display of the model results.

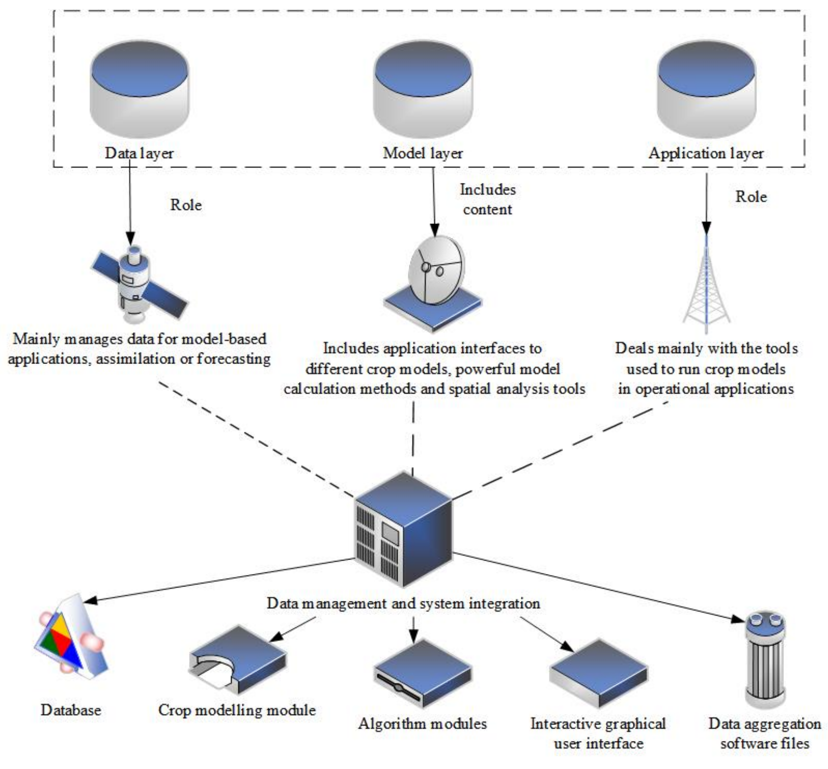

Figure 3. General framework of the energy crop yield simulation and prediction system.

5. Effect of the energy crop yield simulation and prediction system

The yield prediction of energy crops plays an extremely important role in the distribution and utilization of these crops. Only by accurately predicting the output of energy crops can people achieve accurate distribution and utilization in crop analysis and utilization. Therefore, the yield prediction of energy crops was analyzed herein, and four main energy crops were selected to predict the main energy substances. At the same time, four different yield simulation and prediction systems were selected for the comparative analysis. The experimental materials and methods that were selected are summarized in the Table.

5.1. Forecasting the sorghum yield

Sorghum with an actual yield of 4454 kg/hm^2 was selected as the test sample. Four models were used for the analysis, namely, A refers to the remote sensing yield estimation model, B refers to the multiple regression and factor analysis model, C refers to the grey theory model, and D refers to the BP neural network model. The results are presented in Figure 4.

Figure 4 shows the predicted value of the sorghum yield predicted by the different methods (left), and the relative deviation of the sorghum yield predicted by the different methods (right). The sorghum yield measured by the remote sensing yield estimation model was 4355.2 kg/hm^2, with a relative deviation of –2.22% from the actual. The sorghum yield measured by the multiple regression and factor analysis model was 4635.5 kg/hm^2, with a relative deviation of 4.07% from the actual. The sorghum yield measured by the grey theory model was 5135.4 kg/hm^2, with a relative deviation of 15.3%. The sorghum yield measured by the BP neural network model was 4462.8 kg/hm^2, with a relative deviation of 0.2% from the actual. In comparison, the sorghum yield measured by the energy crop yield simulation and prediction system using the BP neural network was the closest to the actual value and the highest accuracy compared with the sorghum yield measured by the other models.

Table. Materials and methods selected for the experiments.

Experimental materials	Experimental methods
Sorghum	Remote sensing yield estimation model
Cassava	Multiple regression and factor analysis models
Maize	Grey theory models
Sugar cane	BP neural network models

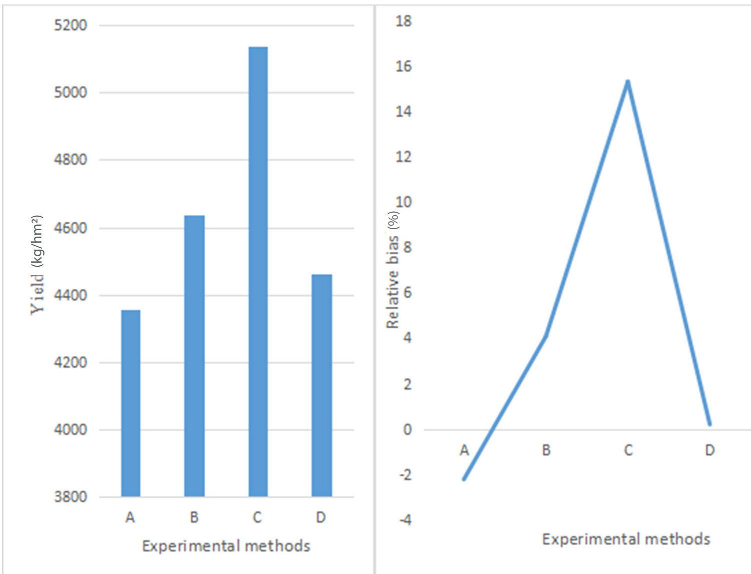

Figure 4. Results of the different methods for predicting the sorghum yield. Predicted values of the sorghum yield by the different methods (left) and relative deviations of the sorghum yields predicted by the different methods (right).

5.2. Forecasting the cassava production

Cassava with actual output of 5300 kg/hm^2 was selected as the test sample, and the measured cassava output results are presented in Figure 5.

Figure 5 shows the predicted value of the cassava yield predicted by the different methods (left), and the relative deviation of the cassava yield predicted by the different methods (right). The cassava yield measured by the remote sensing yield estimation model was 5631 kg/hm^2, with a relative deviation of 6.25%. The cassava yield measured by the multiple regression and factor analysis model was 6324 kg/hm^2, with a relative deviation of 19.32% from the actual. The cassava yield measured by the grey theory model was 4925 kg/hm^2, with a relative deviation of –7.08%. The cassava yield measured by the BP neural network model was 5442 kg/hm^2, with a relative deviation of 2.68%.

5.3. Forecasting the corn yield

Corn with actual yield of 13,500 kg/hm^2 was selected as the test sample, and the measured corn yield results are presented in Figure 6.

Figure 6 shows the predicted value of the corn yield predicted by the different methods (left), and the relative deviation of the corn yield predicted by the different methods (right). The corn yield measured by the remote sensing yield estimation model was 20,154 kg/hm^2, with a relative deviation of 49.29% from the actual. The corn yield measured by the multiple regression and factor analysis model was 14,206 kg/hm^2, with a relative deviation of 5.23% from the actual. The corn yield measured by the grey theory model was 11,254 kg/hm^2, with a relative deviation of –16.64% from the actual. The corn yield measured by the BP neural network model was 13,467 kg/hm^2, with a relative deviation of –0.24% from the actual. The corn yield measured by the energy crop yield simulation and prediction system using the BP neural network was the closest to the actual value, but slightly lower than the actual value.

5.4. Forecasting the sugarcane yield

Sugarcane with an actual yield of 8620 kg/hm^2 was selected as the test sample and the measured sugarcane yield results are presented in Figure 7.

Figure 7 shows the predicted value of the sugarcane yield predicted by the different methods (left), and the relative deviation of the sugarcane yield predicted by the different methods (right). The sugarcane yield measured by the remote sensing yield estimation model was 10,024 kg/hm^2, with a relative deviation of 16.29% from the actual yield. The sugarcane yield measured by the multiple regression and factor analysis model was 8935 kg/hm^2, with a relative deviation of 3.65% from the actual. The sugarcane yield measured by the grey theory model was 6954

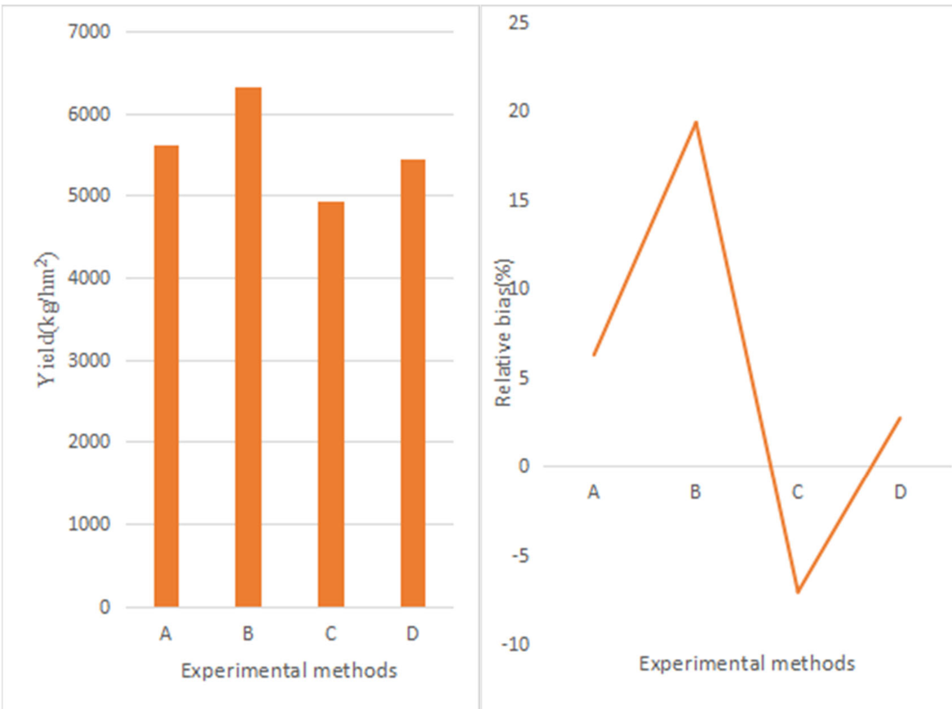

Figure 5. Results of the different methods for predicting the cassava yield. Predicted values of the cassava yield by the different methods (left) and relative deviations of the cassava yields predicted by the different methods (right).

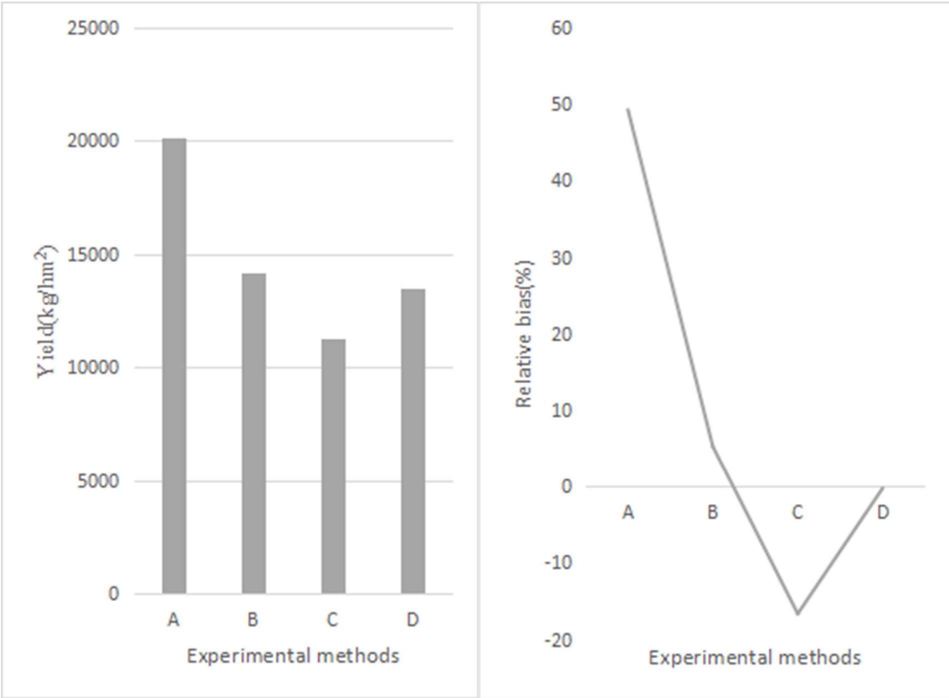

Figure 6. Results of the different methods for predicting the maize yield. Predicted values of the maize yield by the different methods (left) and relative deviations of the maize yields predicted by the different methods (right).

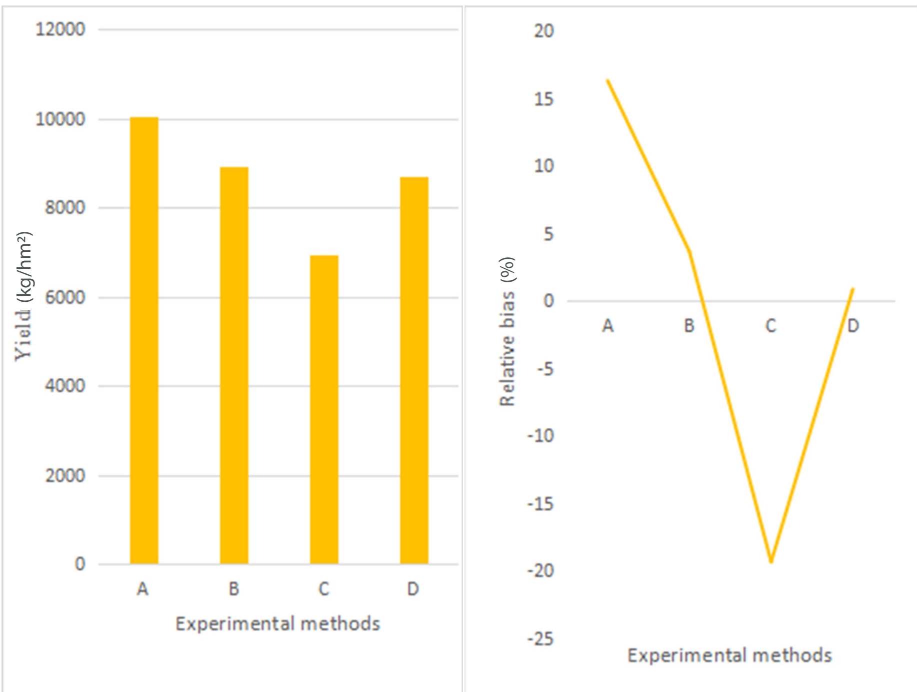

Figure 7. Results of different methods for predicting sugarcane yield. Predicted values of the sugarcane yield by the different methods (left) and relative deviations of the sugarcane yields predicted by the different methods (right).

kg/hm^2, with a relative deviation of −19.33%. The sugarcane yield measured by the BP neural network model was 8694 kg/hm^2, with a relative deviation of 0.86% from the actual. The relative deviation between the sugarcane yield value measured by the energy crop yield simulation and prediction system using the BP neural network and the actual value was less than 1%, which can basically be ignored.

6. Discussion

6.1. Problems with prediction in the energy crop yield model

The growth, development, and yield processes of energy crops are very complex, and are affected by natural and human factors such as climate and soil. Socioeconomic factors, such as politics and markets, also amplify the changes in local food production. At present, there are many methods to predict and forecast the medium and long-term yield of energy crops. Some models have high prediction accuracy, but the above analysis shows that the yield prediction model needs to be improved in the following aspects.

6.1.1. Forecast accuracy

Some models have high accuracy in regional crop yield prediction, such as remote sensing models and technical data models. Most models cannot be used for medium and long-term prediction, and the accuracy of yield prediction for small single crops is also not high. The mechanism model can be used for simulation of different scales, but the future medium and long-term prediction would depend on the accuracy of climate change prediction. At present, it is more likely to estimate the yield trend under different climate change and policy scenarios in the future.

6.1.2. The model's utility

In statistical and mechanical models, many parameters are related to speed in specific application fields. Therefore, it is difficult to use them in other fields or on a larger spatial scale. Expanding these models to a larger spatial scale requires more detailed information. As the number of parameters increases, the accuracy and applicability of these models would also decline.

6.1.3. Practicality of the model

It may be difficult to choose the correct model because different models use different principles and methods for different purposes and application fields. Many models are very complex, including many factors that affect the performance of energy crops. Therefore, users need to determine which factors affect the simulation of the model and the role they play. The lack of practicality of these models limits their dissemination and use because it is difficult for ordinary users to define parameters, calibrate models and use them.

6.2. Strategies for improving the yield of energy crops

6.2.1. Selecting high-quality varieties for planting

With the improvement of science and technology, there are more and more types of energy crops, and their yield and quality are different under different conditions. Therefore, it is important for growers to select the most suitable energy crop varieties for the region before planting. When selecting seeds, growers should also be aware of the impact of regional climate conditions on the growth and development of energy crops.

6.2.2. Proper land management

Soil conditions have a great impact on the yield and quality of energy crops, so growers need to prepare soil at the planting site, which is a part of daily production operations. When sowing, the seeds of energy crops should be evenly sown in soil that is as moist and flat as possible. This ensures that sufficient nutrients are available for the growth and development of the crops. If any unevenness is found in the soil during crop growth, it should be corrected immediately. Finally, if the soil is too dry, timely irrigation is needed to ensure that there is enough water for the growth and development of the crops.

6.2.3. Formulation of a scientific and reasonable fertilization plan

Before applying mineral fertilizer, it is important to first select the fertilizer type that is most suitable for the actual demand of the energy crops. As part of this process, growers should fully understand the characteristics of the energy crops, and then use these data to select fertilizers that meet the actual needs before formal fertilization. Growers need to choose different fertilizers for their production activities, taking into account the specific conditions of the different planting stages. This should maximize the role of chemical fertilizers in the growth and development of energy crops to effectively improve the yield of energy crops.

6.2.4. Controlling the growth density of energy crops

Growers should check the density of the energy crops before sowing. According to the actual needs of the energy crops, appropriate planting plans must be selected to scientifically adjust the plant density. If the plant density is too high during the cultivation of energy crops, the number of plants should be reduced in time to allow the normal growth of the energy crops.

7. Conclusions

In order to improve the prediction accuracy of energy crop yield, this study used the machine learning algorithm to design a simulation and prediction system for energy crop yield and used different energy crop analysis systems to analyze the yield of the different energy crops. The accuracy results of four models: the remote sensing yield estimation model, multiple regression and factor analysis model, grey theory model, and BP neural network model were compared. It was found that the yield of sorghum, cassava, corn, and sugarcane measured by the energy crop yield simulation and prediction system using the BP neural network was the closest to the actual value, with the highest relative accuracy. Machine learning algorithms play an extremely important role in the prediction of energy crop yields, which can be used to predict more data.

Funding
Science Research Foundation of Education Department of Jilin Province (JJKH20230430KJ)

References

Shahhosseini M (2021). Coupling machine learning and crop modeling improves crop yield prediction in the us corn belt. Scientific Reports 11 (1): 1-15. https://doi.org/10.1038/s41598-020-80820-1

Kamath P (2021). Crop yield forecasting using data mining. Global Transitions Proceedings 2.2: 402-407. https://doi.org/10.1016/j.gltp.2021.08.008

Peng B (2018). Benefits of seasonal climate prediction and satellite data for forecasting US maize yield. Geophysical Research Letters 45 (18): 9662-9671. https://doi.org/10.1029/2018GL079291

Zhu WJ, Lysa P, Ken ST (2019). A credibility-based yield forecasting model for crop reinsurance pricing and weather risk management. Agricultural Finance Review 79 (1): 2-26. https://doi.org/10.1108/AFR-08-2017-0064

Heino M (2020). A multi-model analysis of teleconnected crop yield variability in a range of cropping systems. Earth System Dynamics 11 (1): 113-128. https://doi.org/10.5194/esd-11-113-2020

Devika B, Ananthi B (2018). Analysis of crop yield prediction using data mining technique to predict annual yield of major crops. International Research Journal of Engineering and Technology 5 (12): 1460-1465.

Palanivel K, Chellammal S (2019). An approach for prediction of crop yield using machine learning and big data techniques. International Journal of Computer Engineering and Technology 10 (3): 110-118. https://doi.org/10.34218/IJCET.10.3.2019.013

Garanayak M (2021). Agricultural recommendation system for crops using different machine learning regression methods. International Journal of Agricultural and Environmental Information Systems 12 (1): 1-20. https://doi.org/10.4018/IJAEIS.20210101.oa1

Andrianov N (2018). A machine learning approach for virtual flow metering and forecasting. IFAC-PapersOnLine 51 (8): 191-196. https://doi.org/10.1016/j.ifacol.2018.06.376

Naumzik C, Stefan F (2021). Forecasting electricity prices with machine learning: predictor sensitivity. International Journal of Energy Sector Management 15 (1): 157-172. https://doi.org/10.1108/IJESM-01-2020-0001

Vera I (2021). Supply potential of lignocellulosic energy crops grown on marginal land and greenhouse gas footprint of advanced biofuels—A spatially explicit assessment under the sustainability criteria of the Renewable Energy Directive Recast. GCB Bioenergy 13 (9): 1425-1447. https://doi.org/10.1111/gcbb.12867

Fahlbusch W (2018). Trace element delivery for biogas production enhanced by alternative energy crops: results from two-year field trials. Energy, Sustainability and Society 8 (1): 1-11. https://doi.org/10.1186/s13705-018-0180-1

Paschalidou A, Michael T, Kiriaki K (2018). Perennial vs annual energy crops-SWOT analysis (case study: Greece). International Refereed Journal of Engineering and Science 7 (9): 01-24.

Greener JG (2022). A guide to machine learning for biologists. Nature Reviews Molecular Cell Biology 23 (1): 40-55. https://doi.org/10.1038/s41580-021-00407-0

Zhao Y (2022). Predicting delays in prefabricated projects: SD-BP neural network to define effects of risk disruption. Engineering, Construction and Architectural Management 29 (4): 1753-1776. https://doi.org/10.1108/ECAM-12-2020-1050