Plant Gene and Trait 2024, Vol.15, No.1, 1-7
http://genbreedpublisher.com/index.php/pgt

Research Article Open Access

Predicting Wheat Response to Drought Using Machine Learning Algorithms

Wu Weichang
Jiugu MolBreed SciTech Ltd., Zhuji, 311800, China
Corresponding author email: 3397575099@qq.com
Plant Gene and Trait, 2024, Vol.15, No.1 doi: 10.5376/pgt.2024.15.0001
Received: 10 Dec., 2023
Accepted: 25 Jan., 2024
Published: 15 Feb., 2024

Preferred citation for this article:
Wu W.C., 2024, Predicting wheat response to drought using machine learning algorithms, Plant Gene and Trait, 15(1): 1-7 (doi: 10.5376/pgt.2024.15.0001)

Abstract With the intensification of global climate change, drought poses a serious threat to agricultural output, so it is essential to find accurate forecasting methods. Machine learning algorithms such as support vector machines, neural networks and random forests have been widely used in modeling and forecasting wheat drought response. By analyzing multidimensional data during plant growth, these algorithms are able to identify key growth indicators and drought response factors, providing a powerful tool to improve the cultivation and management of drought resistance in wheat. This review summarizes the research progress in using machine learning algorithms to predict wheat crop response to drought, highlights the potential of machine learning in predicting wheat drought response, and suggests directions for future research to further improve the prediction accuracy and applicability of wheat drought resistance.
Keywords Wheat; Drought response; Machine learning algorithms; Growth index; Drought disaster

As one of the world's most important food crops, wheat plays an indispensable role in maintaining food security and safeguarding human survival, and its high-yield and high-quality production is essential to meet the needs of the world's growing population. However, wheat production faces challenges from a variety of environmental pressures, the most significant of which is drought, which not only directly affects the growth and development of wheat, but also leads to a sharp decline in production, and triggers a global food crisis, which in turn threatens global food security (Zhang et al., 2023).

In recent years, with the rapid development and wide application of machine learning technology, the agricultural field has gradually begun to use these advanced algorithms and models to solve many problems in wheat production, especially in predicting the response of wheat crops to drought. Machine learning models have shown great potential (Ding et al., 2020, IT Manager World, 23(6): 188-189). However, despite some progress in research, there are still some challenges and problems in practical applications, such as limitations in data acquisition, bottlenecks in model accuracy, and so on. Therefore, further research is needed on how to better use machine learning algorithms to predict wheat's response to drought.

The purpose of this study is to systematically investigate the prediction of wheat response to drought, and to analyze its application in agricultural research from the perspective of machine learning model. By summarizing the existing research results, evaluating the advantages and limitations of machine learning models in predicting the effects of drought on wheat yield and quality, and looking forward to future research, the aim is to provide theoretical support and guidance for improving wheat drought resistance and ensuring food security.

1 Response Mechanism of Wheat to Drought

1.1 Changes in physiological processes

Wheat showed a variety of physiological process changes in arid environment to adapt to water restriction stress. In the face of drought stress, the physiological processes of wheat plants have been adjusted and changed in many aspects. In response to water stress, wheat adopted a series of water regulation strategies. Plants reduce transpiration by regulating stomatal opening and closing to reduce water loss (Zhang et al., 2019). At the same time, root morphology and structure change to enhance water absorption and utilization, including the deep penetration of roots into the soil and the increase of capillary roots (Figure 1).

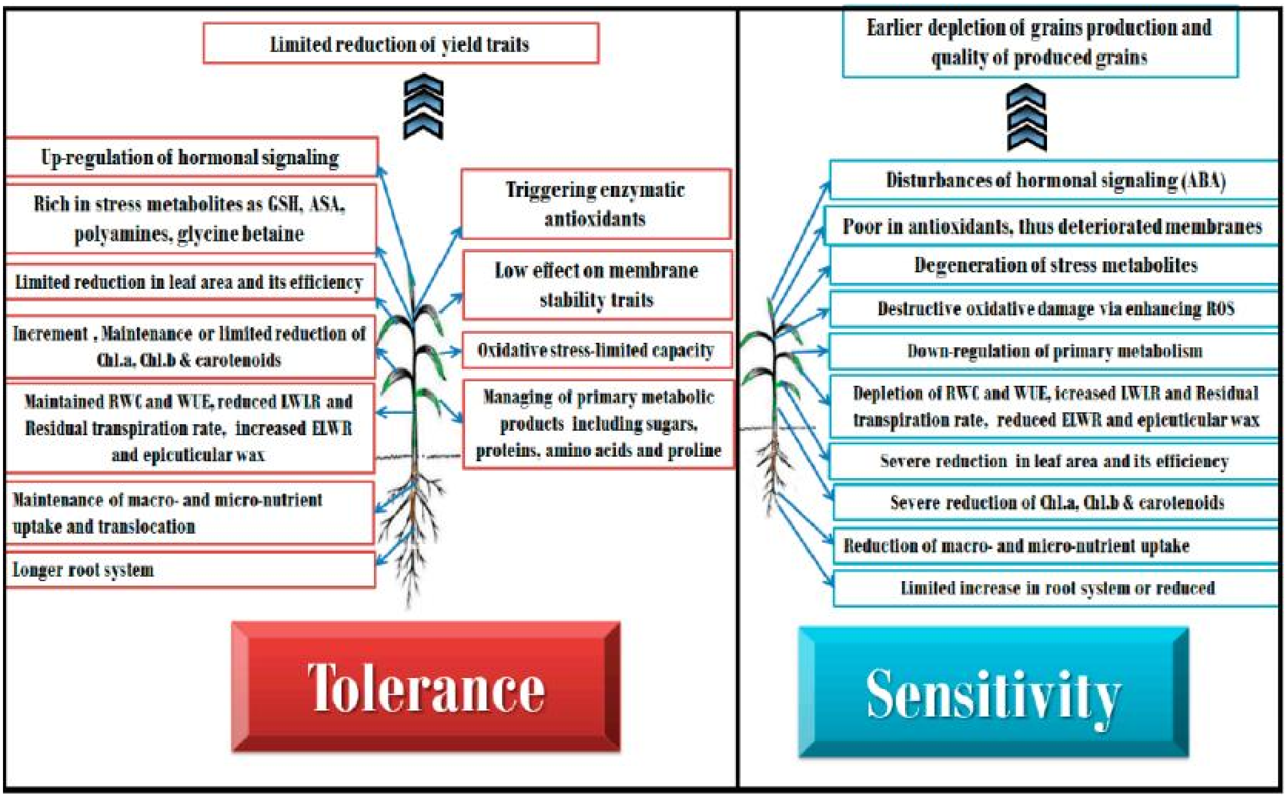

Figure 1 Physiological changes in wheat and barley genotypes in response to drought stress (Sallam et al., 2019)

The response of wheat to drought also involves the regulation of plant growth and development. In a water-constrained environment, wheat plants grow at a slower rate, including reduced leaf growth, stem elongation, and root development, to conserve water and adjust the allocation of carbon resources. At this time, plants may selectively retain or cut back some metabolic processes related to growth and development in order to adapt to drought stress.

The metabolic pathways of plants are also adjusted under drought conditions. Wheat may have increased the activity of the antioxidant enzyme system in response to oxidative damage due to drought. At the same time, hormone levels also change, for example, the accumulation of abscisic acid may affect plant growth and development and stress response. The adjustment and change of these physiological processes work together to coordinate plant adaptation and help wheat better cope with the challenges of arid environments.

1.2 Reaction at the molecular level

The molecular responses and regulatory changes of wheat under drought stress are very complex. In order to adapt to drought environment, plants regulate gene expression and signal transduction networks through a variety of molecular mechanisms, which involve the regulation of expression of many key genes, especially genes regulating stress response and antioxidant response (Rijal et al., 2021).

At the molecular level, wheat can regulate multiple pathways to adapt to drought stress, and ABA (abscisic acid) pathway is a key signaling mechanism in stress response pathway, which can regulate the expression of multiple stress response genes. By regulating gene expression, ABA pathway activates stress-responsive genes such as LEA protein gene family, which encode proteins that help maintain cell stability and improve wheat tolerance to drought. The ABA signaling pathway also triggers stomatal closure, which is critical for reducing water evaporation. Under drought conditions, wheat can limit water loss in this way and improve leaf water use efficiency. Wheat also increases the surface area of the root system by regulating the growth pattern of the root system and secretes root secretions, thereby improving the water absorption capacity in response to drought stress. Drought conditions can lead to oxidative stress, producing excessive reactive oxygen species, causing damage to plant cells, and wheat will enhance the antioxidant system by activating the ABA pathway to reduce oxidative damage.

Drought response of wheat also involves the regulation of expression of many stress-related genes. These genes may encode protective proteins, such as proline aminolyase and antioxidant enzymes, to alleviate oxidative damage caused by stress. In addition, the regulation of several transcription factors and signal transduction elements is also key, which can initiate or inhibit gene expression in specific pathways and regulate stress response networks within cells.

1.3 Restrictions and limitations of traditional research methods

The traditional research methods often have some restrictions and limitations in analyzing the mechanism of wheat response to drought. For the complex biological processes related to drought stress, traditional research methods are difficult to analyze comprehensively and efficiently. Traditional biological and physiological research is limited to the study of specific biological processes or biomolecules, which leads to a potentially incomplete understanding of the overall mechanisms of drought stress response (Mwadzingeni et al., 2016). Because drought stress involves complex interactions at the molecular, cellular, and tissue levels, traditional biological approaches alone may not provide a full insight into the complexity of these interactions.

Traditional experimental operations are also limited by time and space. Drought is a gradual process, and it may take a long time to monitor the physiological and molecular changes of wheat under different drought degrees. In addition, due to the limitation of experimental environment, it is difficult for traditional methods to completely simulate the complex natural drought environment. Traditional methods may also have the problem of insufficient detection sensitivity. Some molecular changes or interactions may require more sensitive instruments or techniques to accurately capture, and traditional methods may not be able to meet this need, resulting in some subtle but critical molecular changes being overlooked or masked.

2 Machine Learning Model

2.1 Selection and application of typical machine learning models

In studies exploring wheat's response to drought, typical machine learning models are widely used to predict and explain its response mechanism, and selecting an appropriate machine learning model is crucial to understanding wheat's drought response mechanism. In addition to wheat, machine learning models are also applied to other plants, taking corn as an example. In the field of genomics and epigenomics, machine learning models can help analyze the genomic data of corn and identify gene functions, regulatory networks and biological pathways. The application of these models provides important clues for gene editing and breeding of corn. Machine learning also plays a key role in the prediction and control of maize diseases. By analyzing disease data, the model can quickly identify disease types and provide corresponding prevention and control suggestions to help farmers prevent and control diseases in time. For the analysis of corn growth and ecological environment, machine learning models can also predict corn yield and adaptability according to various factors such as climate, soil and growth conditions, providing an important reference for agricultural production. In corn crop management and precision agriculture, machine learning techniques can optimize agricultural decisions based on real-time data, such as providing guidance on water use and fertilization, to improve corn growth quality and agricultural yield.

Common machine learning models include decision trees, support vector machines, random forests, neural networks, regression models, etc. (Cai et al., 2021). Decision tree model has attracted much attention because it is easy to understand and interpret. It gradually generates decision rules by branch selection of data set. Support vector machines (SVMS) classify and regression data by constructing hyperplanes and are suitable for complex and nonlinear data sets. Random forest is an integrated model based on multiple decision trees, which can efficiently process a large number of features and data sets. Neural networks mimic the connection patterns of human brain neurons and are suitable for processing complex and large-scale data, but require more data volume and computational resources. Regression models are often used to predict the response of continuous variables such as wheat growth or yield.

In the application of these models, it is necessary to consider the selection and preprocessing of data features, the optimization of model parameters, the problems of overfitting and underfitting, and the interpretability of models. In addition, for wheat drought response prediction, it is usually necessary to integrate multiple machine learning models to improve the accuracy and robustness of the prediction.

2.2 The results and enlightenment of model experiment

The model experiment has obtained many achievements and enlightenment in exploring the response mechanism of wheat to drought. These experiments provide an opportunity to gain insight into the mechanisms and characteristics of wheat drought response. Through model experiments, researchers are able to identify and understand the physiological, molecular level changes in wheat under drought conditions and its response to environmental stress. This provides us with methods and strategies for improving drought resistance of wheat in agricultural production (Ahmed and Hussain, 2022).

The model experiments also revealed the effects of drought on wheat growth and yield. By simulating and predicting the growth status and yield changes of wheat under different drought conditions, we can better assess the impact of drought on wheat planting yield, and provide scientific basis and advice for wheat planting under drought conditions.

The model experiments also provide tools and methods for predicting and evaluating wheat response to drought. By building machine learning models, the researchers were able to predict wheat growth, yield changes and its response to drought stress in different drought scenarios. This provides an important reference for wheat variety improvement and agricultural management in the future.

In the field of soybean research, model experiments have helped to identify the factors that affect the yield and quality of soybean under different growth conditions. Based on the analysis of climate, soil, plant characteristics and other data, the model can accurately predict soybean growth and yield changes, which helps farmers optimize land management and planting methods, and improve soybean yield and quality. The model experiments also provided insights into soybean diseases and pests. The model can identify common soybean diseases and insect pests, and predict their spread path and impact degree. This prediction facilitates the early implementation of necessary control measures to protect soybean crops from diseases and pests. Model experiments also play an important role in soybean breeding. Through the analysis of genomics and epigenomics, the model can more comprehensively understand the genetic characteristics and growth patterns of soybean varieties, and provide more accurate data support for seed selection and breeding.

2.3 Application of machine learning model to wheat drought response

Machine learning model plays an important role in the study of wheat drought response. A team of researchers built a convolutional neural network model by collecting physiological data and environmental parameters (such as soil moisture, air temperature, humidity, etc.) of wheat at different growth stages during the experiment. The model can predict the growth state of wheat under different drought levels. The research team processed and labeled the data set, and then designed a deep convolutional neural network to optimize the model through training and validation to improve the prediction accuracy.

Through the application of convolutional neural network, the research team can more accurately predict the growth of wheat under drought conditions. This case shows us the application prospect of advanced machine learning technology in the agricultural field. Using deep learning models such as convolutional neural network to solve agricultural problems not only improves the efficiency of agricultural production, but also improves the efficiency of agricultural production. It also promotes the innovative application of science and technology in the field of agriculture.

By processing a large amount of data, the machine learning model can accurately identify and predict the growth situation and yield changes of wheat under drought conditions. Through the analysis of multi-dimensional data such as environmental data, genetic information and growth indicators, the machine learning model can identify the key factors affecting wheat resistance to drought. This provides beneficial decision support for agricultural production (Ji and Li, 2019, Journal of Tonghua Normal University, 40(6): 73-77). Through the predictive power of the model, agricultural practitioners can better plan planting strategies, select wheat varieties adapted to drought conditions, and develop effective agricultural management practices to minimize the impact of drought on wheat yields.

Machine learning models have also played a role in exploring and explaining the mechanism of wheat drought response. Through the analysis and pattern recognition of a large number of data, these models are helpful to understand the ways and key factors of wheat response to drought, and provide new ideas and methods for further study of wheat drought resistance mechanism.

3 Challenges and Opportunities

3.1 Data quality and availability

The success of machine learning algorithms depends heavily on the quality and availability of data. In predicting wheat response to drought, data quality directly affects the accuracy and reliability of the algorithm, and high-quality data can provide the diversity and representativeness required by the model, but in the field of agriculture, data quality often faces multiple challenges (Ambarwari et al., 2020).

Data quality problems can result from errors in the data collection process, including but not limited to missing data, outliers, labeling errors, or inaccurate labeling. To solve these problems, data cleaning, standardization and correction are needed to ensure the integrity and accuracy of data. Data availability is also a challenge. Agricultural data often comes from multiple sources, and the format and standards are inconsistent, so it needs to be integrated and unified. In addition, some data may not be publicly available or shared, making data acquisition difficult.

To solve the problem of data quality and availability, data preprocessing technology, feature engineering and data integration methods should be integrated. At the same time, it is necessary to strengthen the standardization of data collection and data sharing in order to make more extensive use of high-quality data for the training and optimization of machine learning models. Effective handling of data quality and availability issues will help improve the accuracy and practicality of machine learning algorithms in predicting wheat drought response.

3.2 Model generalization ability

The model generalization ability of a machine learning algorithm is a key factor in evaluating its performance on new data, which refers to the model's performance on previously unseen data. For predicting wheat response to drought, the generalization ability of the model determines its applicability and reliability in real scenarios.

If a model performs well only on training data, but poorly on new data, it indicates that the model is overfitting. Overfitting means that the model overadapts to the characteristics of training data, resulting in poor generalization ability on new data. On the contrary, if the model performs well on both training data and new data, it indicates that the model has strong generalization ability (Cao et al., 2021).

For the prediction of wheat response to drought, the model generalization ability is affected by data quality, model complexity and training methods. In order to improve the model generalization ability, appropriate model evaluation methods, such as cross-validation and data set partitioning, should be adopted. In addition, techniques such as feature selection and model regularization also help reduce overfitting and improve the generalization ability of the model.

When machine learning algorithm is applied to wheat drought response prediction, evaluating the generalization ability of the model is an important step to ensure the reliability and practicability of the model. A model with good generalization ability can predict wheat response to drought more accurately and provide more accurate guidance and decision support for agricultural production.

3.3 Comparison with the effect of traditional research methods

The advantages and limitations of the machine learning algorithm and the traditional research methods were compared for wheat drought response. Traditional research methods focus on laboratory observation, physiological testing and statistical analysis when exploring wheat drought response. These methods are conducive to in-depth understanding of physiological processes, but limited by the size and complexity of data, it is difficult to fully capture the comprehensive impact of drought on wheat growth and yield. Relatively speaking, machine learning algorithms rely on large data sets and algorithm learning to perform well in processing large-scale data and pattern recognition, and their efficient data processing capabilities enable more accurate prediction and analysis of wheat responses under different drought conditions (Feng et al., 2019). However, machine learning models also have certain limitations, such as high dependence on data quality and label accuracy, and relatively weak interpretability of their results.

In practical research, the combination of traditional methods and machine learning methods may be a more ideal path. The complementary approach of traditional methods, which drill down into physiological processes, and machine learning methods, which can process large-scale data more quickly and accurately, can help to better understand wheat's response to drought. This comprehensive application can make up for the limitations of a single method, and provide more accurate and efficient strategies and guidance for agricultural production. In the future, with the continuous development of technology and the continuous optimization of methods, the combination of machine learning and traditional research methods may become an important direction of wheat drought response research.

3.4 Model improvement suggestion

In the machine learning model for predicting wheat response to drought, continuous improvement and optimization of the model is an important step to improve the accuracy and practicability of the prediction. The feature selection of the model is one of the keys. Through in-depth understanding of the physiological and molecular response mechanism of wheat to drought, more representative features can be extracted to enhance the accurate prediction ability of the model to drought response. Suitable feature selection can reduce the complexity of the model and improve the generalization ability of the model.

The optimization of the model needs to consider the algorithm parameters and model architecture. For wheat drought response prediction, different machine learning algorithms can be explored and their parameters adjusted, such as support vector machines, decision trees, neural networks, etc., to find a more suitable model for the problem. At the same time, adjusting the hyperparameters and network structure of the model, such as increasing the number of layers and adjusting the learning rate, can help improve the model performance (Sundararajan et al., 2021).

Data quality and quantity are also critical to model improvement. Ensuring the accuracy and completeness of the data, while collecting more and more comprehensive sample data, can help the model to better capture the complex response relationship of wheat to drought and improve the prediction accuracy. Model improvement also requires continuous verification and evaluation. The stability and generalization ability of the model are verified by cross-validation, maintaining validation set and other methods, so as to determine whether the model improvement is effective.

4 Conclusion and Prospect

By reviewing a large number of previous studies, we found several important conclusions. Machine learning models show remarkable potential in analyzing wheat's response to drought, and can accurately predict wheat growth and yield under drought conditions. Secondly, the study shows that machine learning algorithms can use multi-source data, such as soil properties, meteorological data, and remote sensing information, to provide a more comprehensive perspective for predicting wheat drought response. Most importantly, these models significantly improve the forecasting accuracy and efficiency compared with traditional methods, providing more forward-looking and accurate decision support for wheat agricultural production.

In the future, researchers can focus their research on several aspects. First, we need to further optimize and improve the machine learning algorithm to improve the accuracy and stability of the model in predicting wheat drought response. And explore the method of multi-model fusion, combining the advantages of different algorithms to build more powerful prediction models. In addition, the improvement of data quality and usability is also a focus of future attention, including in-depth analysis of data quality and the use of more laboratory and field validation data to ensure the robustness and adaptability of the model.

At the same time, strengthening the research on the physiological and molecular links between drought and wheat growth and development will contribute to a deeper understanding of drought response mechanisms, so as to better optimize agricultural production strategies, and then apply machine learning technology to actual agricultural production to develop data-driven agricultural management measures to promote wheat drought resistance and improve yield and quality. These future research directions will push machine learning algorithms to play a more significant role in predicting wheat's response to drought, providing more reliable solutions to the challenges climate change poses to agriculture.

References

Ahmed M.U., and Hussain I., 2022, Prediction of wheat production using machine learning algorithms in northern areas of Pakistan, Telecommunications Policy, 46(6): 102370.
https://doi.org/10.1016/j.telpol.2022.102370

Ambarwari A., Adrian Q.J., and Herdiyeni Y., 2020, Analysis of the effect of data scaling on the performance of the machine learning algorithm for plant identification, Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi), 4(1): 117-122.
https://doi.org/10.29207/resti.v4i1.1517

Cai H.H., Peng J., Liu W.Y., Luo D.F., Wang Y.Z., Bai J.D., and Bai Z.J., 2021, Inversion and mapping of soil pH valve based on in-situ hyperspectral data in cotton field, Shuitu Baochi Tongbao (Bulletin of Soil and Water Conservation), 41(4): 189-195.

Cao J., Zhang Z., Luo Y., Zhang L., Zhang J., Li Z., and Tao F., 2019, Wheat yield predictions at a county and field scale with deep learning, machine learning, and google earth engine, European Journal of Agronomy, 123: 126204.
https://doi.org/10.1016/j.eja.2020.126204

Feng P., Wang B., Liu D.L., and Yu Q., 2019, Machine learning-based integration of remotely-sensed drought factors can improve the estimation of agricultural drought in South-Eastern Australia, Agricultural Systems, 173: 303-316.
https://doi.org/10.1016/j.agsy.2019.03.015

Mwadzingeni L., Shimelis H., Dube E., Laing M.D., and Tsilo T.J., 2016, Breeding wheat for drought tolerance: Progress and technologies, Journal of Integrative Agriculture, 15(5): 935-943.
https://doi.org/10.1016/S2095-3119(15)61102-9

Rijal B., Baduwal P., Chaudhary M., Chapagain S., Khanal S., Khanal S., and Poudel P.B., 2021, Drought stress impacts on wheat and its resistance mechanisms, Malaysian Journal of Sustainable Agriculture, 5(2): 67-76.
https://doi.org/10.26480/mjsa.02.2021.67.76

Sallam A., Alqudah A.M., Dawood M.F.A., Baenziger P.S., and Börner A., 2019, Drought stress tolerance in wheat and barley: Advances in physiology, breeding and genetics research, Int. J. Mol. Sci., 20(13): 3137.
https://doi.org/10.3390/ijms20133137
PMid:31252573 PMCid:PMC6651786

Sundararajan K., Garg L., Srinivasan K., Bashir A.K., Kaliappan J., Ganapathy G.P., Selvaraj S.K., and Meena T., 2021, A contemporary review on drought modeling using machine learning approaches, Computer Modeling in Engineering and Sciences, 128 (2): 447-487.
https://doi.org/10.32604/cmes.2021.015528

Zhang B.Y., Li X., and Zhang X.L., 2023, Influences of drought events on ecological resilience of Larix principis-rupprechtii and Pinus tabulaeformis, Hebei Nongye Daxue Xuebao (Journal of Agricultural University of Hebei), 46(4): 65-73.

Zhang J.B., Xue X.P., Li N., Li H.Y., Zhang L., and Song J.P., 2019, Effects of drought stress on physiological characteristics and dry matter production of winter wheat during water critical period, Shamo yu Lüzhou Qixiang (Desert and Oasis Meteorology), 13(3): 124-130.