Computational Molecular Biology 2026, Vol.16, No.2, 98-113
http://bioscipublisher.com/index.php/cmb

Research Insight Open Access

Prediction of Maize Yield Based on Soil Nutrients and Climate Variables

Jinhua Cheng, Wei Wang
Institute of Life Sciences, Jiyang College of Zhejiang A&F University, Zhuji, 311800, Zhejiang, China
Corresponding author: wei.wang@jicat.org
Computational Molecular Biology, 2026, Vol.16, No.2 doi: 10.5376/cmb.2026.16.0008
Received: 02 Feb., 2026
Accepted: 08 Mar., 2026
Published: 21 Mar., 2026

Preferred citation for this article:
Cheng J.H., and Wang W., 2026, Prediction of maize yield based on soil nutrients and climate variables, Computational Molecular Biology, 16(2): 98-113 (doi: 10.5376/cmb.2026.16.0008)

Abstract Maize yield prediction plays an essential role in ensuring food security and promoting sustainable agricultural management. This study explores a prediction framework based on soil nutrient characteristics and climate variables to improve the accuracy and reliability of maize yield estimation. Key soil indicators, including nitrogen, phosphorus, potassium, organic matter, and pH value, were combined with climate factors such as temperature, precipitation, and accumulated growing degree days. Multiple prediction models, including traditional statistical approaches, machine learning algorithms, and deep learning methods, were constructed and compared. The study further analyzed the interaction effects between soil and climate variables and evaluated model performance using indicators such as RMSE, MAE, and R^2. A regional case study was conducted to verify the applicability and robustness of the proposed framework. The results demonstrate that integrating soil nutrient and climate data can significantly enhance maize yield prediction accuracy and provide valuable support for precision agriculture, crop management, and agricultural decision-making.
Keywords Maize yield prediction; Soil nutrients; Climate variables; Machine learning; Precision agriculture

1 Introduction

Global demand for maize is rising steadily as it underpins food, feed, and industrial supply chains, yet production is increasingly constrained by climate variability and degraded soils. Temperature extremes, altered rainfall, and declining soil fertility jointly threaten yield stability, especially in regions already facing food insecurity. Improving the accuracy of maize yield prediction by explicitly linking soil nutrients with key climate variables is therefore essential for optimizing fertilization, managing risk, and designing climate-smart production systems. Maize yields respond strongly to interactions between climate conditions and soil nutrient status. Studies in sub-Saharan Africa and China show that nitrogen (N), phosphorus (P), and potassium (K) inputs can buffer or amplify the impacts of changing CO_2, temperature, and rainfall on yield, and that soil indigenous nutrients strongly modulate yield losses under warming (Falconnier et al., 2020. Long-term experiments further indicate that soil fertility improvements (e.g., higher total and available N and P) enhance yield stability and sustainability, while climate warming tends to reduce yields where soil fertility is low. At the same time, nutrient management alone is insufficient; integrating soil, climate, and management information is needed to maintain productivity under ongoing climate change (Ocwa et al., 2023). In this context, a predictive framework that couples soil nutrient properties with climate variables can support more precise fertilizer recommendations, reduce environmental risks, and improve resilience of maize-based systems.

Internationally, two main directions have emerged. First, process-based crop models are used to simulate maize yield responses to climate scenarios and N management, revealing strong interactions between N inputs, soil N dynamics, and climate drivers in both low-input and intensive systems (Falconnier et al., 2020). Second, data-driven approaches, especially machine learning (ML) and deep learning (DL), increasingly predict crop yields from large datasets combining soil, climate, and management information. Systematic reviews show that temperature, rainfall, soil type, soil nutrients, and vegetation indices are among the most frequently used predictors, and that algorithms such as Random Forest (RF), Support Vector Machines, Artificial Neural Networks, CNNs, and LSTMs dominate recent work. For maize specifically, RF models trained on multi-year field trials in Ghana identified soil properties (e.g., organic carbon, total N, exchangeable bases) and maximum