temperature as the most important predictors of yield, surpassing purely climatic models and improving understanding of nutrient-climate interactions (Asamoah et al., 2024). Related studies using RF and other ML algorithms have shown that including both soil and weather variables substantially improves prediction of maize yield under zero N fertilization and in drought-stressed environments. These advances highlight the potential of combining soil nutrient information with climate variables in robust predictive frameworks, but also reveal gaps: many models rely on limited nutrient descriptors, treat climate and soil separately, or focus on short time periods and narrow environments.

Building on this progress, the present study focuses on prediction of maize yield based explicitly on soil nutrient status and climate variables, aiming to better capture their joint effects. The main research contents are: (1) construction of a comprehensive feature set describing soil nutrients (e.g., N, P, K, organic matter, pH and related properties) and key climate factors (temperature, precipitation, radiation, humidity) relevant to maize growth; (2) development and comparison of data-driven yield prediction models, with emphasis on ensemble methods such as Random Forest and other ML/DL techniques that have shown strong performance in crop yield prediction; and (3) quantitative analysis of variable importance and interaction patterns between soil nutrients and climate variables, to identify critical drivers of yield variation and potential leverage points for management. The technical route begins with data collection and preprocessing, including quality control and normalization of soil and climate data. Next, the dataset is split into training and testing subsets, and multiple candidate models are trained, tuned, and evaluated using metrics such as coefficient of determination (R^2) and root mean square error (RMSE), following best practices from recent ML yield-prediction studies. Finally, model interpretation techniques (e.g., variable importance analysis and partial response analysis) are applied to quantify how specific combinations of soil nutrients and climate variables influence predicted maize yield, providing both a practical prediction tool and theoretical insight for nutrient management and climate adaptation strategies.

Across diverse environments, maize yield is jointly controlled by soil nutrient status and climate conditions, and their interaction largely determines both productivity and stability. While process-based models and ML/DL approaches have advanced yield prediction, there remains a need for models that explicitly integrate detailed soil nutrient descriptors with key climate variables and provide interpretable guidance for management. This study addresses that gap by constructing and evaluating data-driven maize yield prediction models grounded in soil-climate interactions, aiming to support more precise fertilization, risk management, and climate-smart maize production.

2 Analysis of Factors Influencing Maize Yield

2.1 Mechanism of soil nutrients on maize growth

Maize yield is jointly controlled by soil nutrient supply and climate conditions throughout the growing season. Understanding how these drivers act individually and in combination is essential for reliable yield prediction and targeted management. Adequate N, P, and K fertilization strongly enhances maize growth traits such as plant height, leaf area, cob number, and grain weight, which together raise biomass accumulation and grain yield by large margins compared with unfertilized controls (Kaleri et al., 2026). Long-term NPK application improves key soil properties-including soil organic carbon and available N, P, and K-which in turn explain a larger share of yield variation than phenological factors in the North China Plain (Wang et al., 2024).

Nutrient deficiency, especially of nitrogen and phosphorus, markedly reduces yield and dry matter accumulation in maize-based systems (Sun et al., 2024). Under N, P, or K deficiency, maize root growth and activity are inhibited, and hundreds of genes related to nutrient transport, hormones, and transcription factors are differentially expressed, indicating complex molecular regulation of root adaptation to low nutrient supply (Nana et al., 2020).

2.2 Effects of climate factors on maize yield

Temperature, precipitation, drought, and vapor pressure deficit (VPD) strongly shape maize yield anomalies at regional to global scales. Temperature-related extremes generally show stronger associations with yield deviations than precipitation alone, although irrigation can partially buffer high-temperature damage (Figure 1) (Vogel et al., 2019). In Northeast China, compound drought and heat cause greater yield loss than either stress alone, with