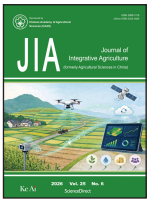
Journal of Integrative Agriculture 2026, 25(6): 2396–2413

Available online at ScienceDirect

Journal of Integrative Agriculture



RESEARCH ARTICLE

Identification of the optimal phenological periods for summer maize yield prediction using UAV-based multispectral data

Qin Dai, Hong Chen, Ziqiang Chen, Chang Liu, Gaoliang Li, Yakun Wang#, Xiaotao Hu#

Key Laboratory of Agricultural Soil and Water Engineering in Arid and Semiarid Areas, Ministry of Education/Northwest A&F University, Yangling 712100, China

Highlights

- A structured framework to identify the optimal phenological periods for summer maize yield prediction using unmanned aerial vehicle (UAV)-based multispectral data was proposed.
- The tasseling stage is the earliest suitable period for maize yield prediction.
- Integrating data from the tasseling, silking, and dough stages is better than either mono-temporal or full-temporal data.

Abstract

Timely and accurate forecasting of crop yields is critical for food management and trade. However, only limited research has explored the impact of integrating crop phenotypic parameters (CPPs) with unmanned aerial vehicle (UAV) data across different phenological stages on maize yield prediction. The extent to which multi-temporal data enhances the accuracy and reliability of yield projections compared to mono-temporal data has yet to be systematically investigated. To attain the proper balance between accuracy and cost in crop yield estimation, this study proposed a structured framework for identifying the optimal phenological periods for summer maize yield prediction using UAV-based multispectral data. Three classical methods of custom mean decrease accuracy (C-MDA), optimal parameters-based geographical detector (OPGD), and grey relational analysis (GRA) were first used to sort and screen both the CPPs and vegetation indices (VIs) derived from UAV-based information over six growth stages. Ridge regression models based on multi-temporal data combinations and mono-temporal data were established separately, and their performance in yield prediction were compared to identify the optimal phenological stages and the corresponding key factors. Our results showed that C-MDA was much better at factor screening and ranking compared to OPGD and GRA. The green normalized difference vegetation index (GNDVI), normalized difference vegetation index (NDVI), and normalized difference red edge index (NDRE) emerged as the top-performing VIs, while the leaf area index (LAI) and above ground biomass (AGB) proved to be the most effective CPPs. When predicting yield using only mono-temporal data, the dough stage delivered the highest predictive accuracy (R^2=0.871, RMSE=0.407 t ha^{-1}), while the tasseling stage was the earliest that achieved yield estimates with acceptable precision (R^2=0.810, RMSE=0.493 t ha^{-1}). In contrast, the integration of UAV data from different crop growth stages markedly enhanced the accuracy of yield estimation. Combinations of data from the tasseling, silking, and dough stages were recommended as the best option (R^2=0.942, RMSE=0.291 t ha^{-1}). These findings indicate that the precise estimation of maize yields in smallholder fields may be attainable, and present both substantial theoretical insights and practical benefits for the advancement of precision agriculture.

Keywords: UAV, yield estimation, multispectral data, geographical detector

1. Introduction

The accurate prediction of grain crop yield is pivotal for the development of informed food policies and the strategic management of agricultural practices (Horie *et al.* 1992; Khaki and Wang 2019). Despite comprising over 80% of farms worldwide, small and medium-sized farms often encounter substantial obstacles in achieving food security. Accurate yield prediction for these plots can assist

Received 25 September 2024; Received in revised form 14 January 2025; Accepted 24 January 2025; Available online 18 February 2025
#Correspondence Yakun Wang, E-mail: wangyakun@nwafu.edu.cn; Xiaotao Hu, E-mail: huxiaotao11@nwsuaf.edu.cn

© 2026 CAAS. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/). Peer review under responsibility of Editorial Board of *Journal of Integrative Agriculture*.
doi: 10.1016/j.jia.2025.02.026

agricultural producers in the early identification of low-yield areas due to issues such as crop insects and plant diseases or improper water and fertilizer management, enabling the timely implementation of intervention measures to support effective agricultural management practices (Samberg *et al.* 2016). Destructive sampling of field crops at multiple scales is a common technique in traditional yield data collection, with the goal of directly obtaining detailed information on crop performance parameters (CPPs). However, the labor-intensive nature, high costs, and susceptibility to variability due to human error when using these methods hinder their widespread application at a regional scale (Khanal *et al.* 2018). Recent advances in remote sensing technology have highlighted its potential as a rapid and efficient method for crop yield prediction. Nevertheless, implementation is challenging due to its susceptibility to weather-related factors and limitations in spatial resolution (Zhou *et al.* 2017; Schut *et al.* 2018; Maimaitijiang *et al.* 2020).

Numerous scholars have harnessed satellite remote sensing techniques to achieve precise crop yield forecasts (Huang *et al.* 2016; Guan *et al.* 2017). Nevertheless, the effectiveness of satellite-based observations is inherently limited by the fixed revisit cycle of satellite platforms, which constrains their ability to obtain flexible and timely data during critical stages of crop development. Furthermore, the intrinsic constraints posed by the resolution of satellite remote sensing data often limit the accuracy of any predictive modeling efforts. While satellite-based observations achieve high accuracy at broader spatial scales, they often cannot provide a detailed representation of dynamic changes at finer spatial resolutions (Bian *et al.* 2022). The high spatial resolution and operational agility of unmanned aerial vehicle (UAV) remote sensing technologies facilitate frequent and high-quality crop imagery acquisition in field trials, so they are highly effective for crop yield prediction applications. Physiological changes in crops during different phenological stages lead to significant differences in the spectral characteristics of the crop canopy, which are captured in the vegetation indices (VIs) used for yield prediction (Dee *et al.* 2018; Yang *et al.* 2024). In the early phenological period, the remote sensing signals are weaker due to the sparser crop canopy, so they are susceptible to interference from external environmental factors (e.g., shading, soil background, etc.) (Liu *et al.* 2023). The mid-phenological stage is a critical growth phase, in which photosynthesis is significantly enhanced, the canopy density increases, and the remote sensing signals become stable and clear. At this stage, remote sensing data can more accurately reflect the growth status of the crop, and the reliability of yield prediction is higher (Fu *et al.* 2024). In the late phenological stage, the accuracy of yield models decreases due to the natural senescence of leaves and the reductions of chlorophyll content and photosynthetic activity. However, as the crop mainly focuses on reproductive growth during this period, it provides valuable data for final yield estimation (Zhang M *et al.* 2022).

Recent studies have proven the utility of leveraging data collected from a limited number of growth stages as sample inputs for yield forecasting models. For instance, Gilliot *et al.* (2021) gathered UAV images taken 126 days post-seeding and found that using UAV data enhanced the precision of predicting maize yield by 15% compared to estimates obtained through manual field sampling. de Villiers *et al.* (2024) used data from four phenological periods to identify UAV-based calculations of VIs and grey-level co-occurrence matrix (GLCM) textural variables as the key indicators for predicting maize yield. Given the limited number of growth stages in these studies, whether the inclusion of data from additional growth stages could yield superior modeling outcomes is not known. Moreover, an extensive body of research has investigated crop yield prediction using data spanning the entire growth cycle of the crop. For example, Wan *et al.* (2020) used UAV image data of the whole growth stage to form a random forest forecasting framework for grain yield, and determined that the combination of multi-temporal normalized difference vegetation index (NDVI), normalized difference yellowness index (NDYI), canopy height, and canopy cover yielded the most accurate predictions for grain yield. Ma *et al.* (2022) showed that the highest accuracy in estimating cotton yield using the RF_ELM Model was achieved by adding textural features to the use of VIs. Although monitoring the entire crop growth cycle can potentially provide a more comprehensive evaluation of crop conditions, this approach requires substantial investments in terms of time, effort, and financial resources. Consequently, the identification of suitable variables and critical growth stages is essential for the development of practical and effective crop yield prediction models.

Different feature selection methods can provide various perspectives and computational strategies when selecting the appropriate variables and key growth stages. These methods help to identify the most influential factors for crop yield prediction in multivariate and complex systems (Yang *et al.* 2022). Optimal parameters-based geographical detector (OPGD) measures the importance of variables by quantifying the explanatory abilities of various factors in geographic spatial data for target variables and is suitable for revealing the spatial heterogeneity between the two (Qiao *et al.* 2022). However, the *q*-value is dependent on spatial unit division, which may lead to information loss (Yang L *et al.* 2024). GRA assesses the importance of variables by measuring the similarity between them, and although it is suitable for complex system analysis, it lacks significance testing and has weak capabilities for processing spatial data (Ikram *et al.* 2020). MDA ranks variables by evaluating their impact on model prediction accuracy. Custom mean decrease accuracy (C-MDA) was developed on this basis, and it allows customized accuracy degradation calculation based on the specific problem and can be adapted to different objective functions. However, it has high computational complexity and may be influenced by model selection (Bénard *et al.* 2022).

To obtain precise crop yield forecasts, a diversity of modeling approaches has been employed. With a better understanding of crop physiological and ecological processes and the continuous development of computer technology, research on crop modelling has gradually emerged. Crop models facilitate the quantitative description and prediction of crop growth and development processes, along with their dynamic interactions with the environment and technology,

enabling the analysis of yield variations under different scenarios. For example, Wanthanaporn *et al.* (2024) used the World Food Studies Crop Model to forecast rice yields in mainland Southeast Asia. Among the variables included in their study, they found that temperature exerted the most significant influence on yield predictions. Richetti *et al.* (2024) demonstrated the efficacy of the APSIM Next Generation Crop Model in predicting spatiotemporal variations in wheat yields at the national, state, and agroecological scales. Crop models can simulate the growth processes of various crops in different regions. However, they often involve complexity and require substantial input data such as meteorological factors, soil properties, crop breeds, and agricultural management techniques (Lecerf *et al.* 2019). Therefore, compared to process-based crop models, statistical models have proven more effective in large-scale yield prediction, so they have achieved broader applications. The simplicity of regression models in statistical modeling also drives their applicability (Gong *et al.* 2018; Paudel *et al.* 2023; Sharma *et al.* 2024). For example, Yuan *et al.* (2021) used a bilinear spectral mixture decomposition model, combined with the product of abundance and VIs, to provide a more accurate prediction of rice yield. In addition, machine learning and deep learning algorithms have garnered significant attention from researchers for their advanced capabilities in modeling complex relationships and conducting efficient computations (Sun *et al.* 2020; Paudel *et al.* 2023). However, the current machine learning yield prediction methods tend to emphasize variations in modeling accuracy across various machine learning algorithms, yet there are few systematic investigations into the selection of dominant control variables across these algorithms. Recently, Sapkota and Paudyal (2023) reported that when estimating maize yield in the hilly regions of Nepal, the green-red vegetation index (GRVI) emerged as the pivotal parameter. Research by Devkota *et al.* (2024) demonstrated the exceptional efficacy of using the ratio of vegetation index (RVI) and NDVI in predicting wheat yield.

Therefore, the novelty of this study lies in simplifying the modeling dataset for the yield prediction model in terms of both the phenological period and its corresponding input control variables, which significantly reduces the detection cost. In addition, the removal of potentially redundant data can both lower the expenses and enhance the prediction accuracy of the resulting model. In this study, VIs were extracted from UAV-based multispectral imagery, and the three CPPs of plant height (PH), leaf area index (LAI), and above-ground biomass (AGB) were measured *in situ* for the purpose of model construction. To improve the model's precision and robustness, these CPP indicators were filtered and ranked using the OPGD, GRA, and C-MDA methods. Ridge regression models, known for their ability to manage multicollinearity and mitigate overfitting, were employed for both mono-temporal and multi-temporal yield estimation tasks (Hoerl and Kennard 1970). Ultimately, the objective of this study was to achieve a significant reduction in the various cost inputs associated with agricultural practices by optimizing the scheduling of UAV observations and developing a streamlined and efficient model for yield estimation. The goal was to answer the following questions: (1) Which specific factors effectively enhance the performance of the maize yield prediction model? (2) Which growth stage of maize is the optimal observation period for yield prediction, considering both lead time and/or accuracy? (3) Compared to mono-stage observations, does maize yield prediction benefit more from combined observations at multiple stages? Hopefully, this study can provide guidance on the accurate prediction of maize yield.

2. Materials and methods

In this study, information derived from six growth stages of maize were systematically collected, encompassing the VIs obtained *via* UAVs and CPPs acquired through ground-based experiments. Subsequently, the three methods of OPGD, GRA, and C-MDA were used to screen the characteristic variables. Based on mono-temporal data, a ridge regression model was constructed to forecast maize production. Then, based on the selected C-MDA method, the multi-temporal data were screened for characteristic variables, and the model was built. The precision of the constructed model was evaluated by the Taylor diagram and Kling-Gupta efficiency (KGE). The main analytical process is illustrated in Fig. 1.

2.1. Study area

The research area is located in the Haiwuhuangjia Modern Agricultural Demonstration Park in Shaanxi Province, China (108°03′E, 34°21′N). This region experiences a temperate semi-humid monsoon climate, with an annual average temperature of 12.9°C and precipitation of 596.8 mm. To demonstrate the applicability of the research method in the Guanzhong Plain of Shaanxi Province, no artificial adjustments were made to water management or fertilization throughout the entire crop-growing period. Instead, farm workers followed the farm's standard cultivation practices to uniformly carry out sowing, fertilization, irrigation, pest control, and harvesting across the entire planting area.

The maize was sown between June 20 and June 27, 2023 and harvested on September 16, 2023. The variety used was Yanke 288, with a row spacing of 0.7 m, plant spacing of 0.3 m, and a planting density of 48,000 plants ha^{-1}. Four typical areas, called A, B, C, and D, were selected for the experiment. Fields A and B were located in the east–west direction of a branch canal, while fields C and D were on one side of a lateral canal. The areas of fields A, B, C, and D were approximately 12,167, 12,920, 7,040, and 10,764 m^2, respectively.

2.2. Data collection

UAV multispectral image collection A DJI Mavic 3 multispectral UAV platform (Xi'an DJI Technology Co., Xi'an, China) was employed to capture images of maize. The multispectral imaging system included four key spectral bands, i.e., the green, red, red edge, and near-infrared bands. The maize images were acquired at the six-leaf, tasseling, silking, blister, milking, and dough stages (i.e., days of year (DOY) 201, 214, 222, 231, 242, and 250, respectively). Image data were collected between 10:00 and 14:00 local

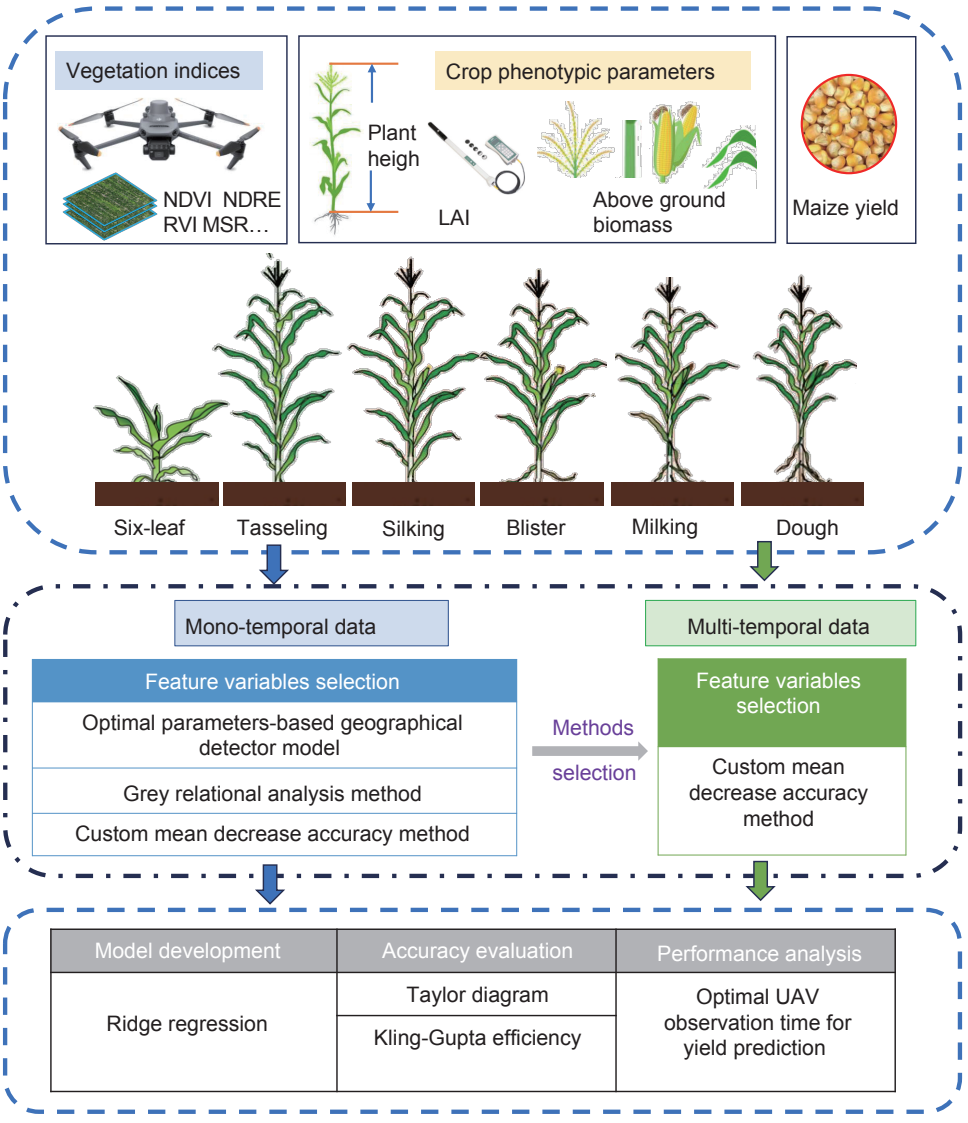


Fig. 1 Main analytical process of this study. NDVI, normalized difference vegetation index; NDRE, normalized difference red edge index; RVI, ratio of vegetation index; MSR, modified simple ratio; LAI, leaf area index; UAV, unmanned aerial vehicle.

time in bright and unclouded weather conditions with minimal wind interference. The UAV was flown at a height of 40 m above the ground, achieving a ground spatial resolution of 2 cm. Flight missions were conducted with a forward overlap of 85% and a side overlap of 75% to ensure comprehensive coverage and data acquisition.

Determination of maize grain yield Maize yield was assessed by harvesting one square meter of mature plants at the time of maturity. The plants were then threshed, cleaned, and dried to a constant weight. Grain yield (t ha^{-1}) was subsequently calculated by dividing the dry weight by the harvested area.

CPPs and VIs Three CPPs and nine VIs, totaling 12 indicators, were used in this study for yield estimation (Table 1). Sampling points were set evenly within the four fields, and three maize plants with uniform growth were randomly selected for plant height measurements in an area of 1 m^2 around each sampling point. To ensure the representativeness of the samples and avoid sample loss, three additional maize plants of similar size to those in the original area were selected outside the area for cutting. The length and maximum leaf width of all leaves were measured and combined with planting density to obtain LAI, and all above-ground material was then dried and weighed to obtain AGB. Spectral indices were calculated using the multispectral images acquired by a UAV in an area of 1 m^2 around the sampling point, excluding the soil background, by applying the formulas in Table 1.

2.3. Analysis methods

OPGD model This study used the "GD" package in R for computation. This package allows for the comparison of *q*-values obtained from different discretization methods, including the equal interval, natural breaks, quantile, geometric, and standard deviation methods, and it automatically selects the method that yields the highest

Table 1 Crop phenotypic parameters (CPPs) and vegetation indices (VIs)

Indicator	Calculation method	Reference
Plant height (PH)	Described in the text	–
Leaf area index (LAI)	Described in the text	Zhai *et al.* (2024)
Above ground biomass (AGB)	Described in the text	–
Normalized difference vegetation index (NDVI)	$NDVI=(R_{NIR}-R_R)/(R_{NIR}+R_R)$	Rouse *et al.* (1974)
Ratio vegetation index (RVI)	$RVI=R_{NIR}/R_R$	Schuerger *et al.* (2003)
Green normalized difference vegetation index (GNDVI)	$GNDVI=(R_{NIR}-R_G)/(R_{NIR}+R_G)$	Gitelson *et al.* (1997)
Modified simple ratio (MSR)	$MSR=(R_{NIR}/R_R-1)/\left[\sqrt{(R_{NIR}/R_R)+1}\right]$	Chen (1996)
Modified soil adjusted vegetation index (MSAVI)	$MSAVI=\left[2R_{NIR}+1-\sqrt{2R_{NIR}+1)^2-8(R_{NIR}-R_R)}\right]/2$	Qi *et al.* (1994)
Modified non-linear vegetation index (MNLI)	$MNLI=[1.5(R_{NIR}^2-R_R)]/(R_{NIR}^2+R_R+0.5)$	Peng *et al.* (2003)
Non-linear vegetation index (NLI)	$NLI=(R_{NIR}^2-R_R)/(R_{NIR}^2+R_R)$	Goel and Qin (1994)
Optimization soil adjusted vegetation index (OSAVI)	$OSAVI=(R_{NIR}-R_R)/(R_{NIR}+R_R+0.16)$	Rondeaux *et al.* (1996)
Normalized difference red edge index (NDRE)	$NDRE=(R_{NIR}-R_{RE})/(R_{NIR}+R_{RE})$	Voitik *et al.* (2023)

q-value and the most optimal partitioning effect for geographic detector analysis. In other words, this package facilitates the identification of the best combination of spatial data discretization methods, spatial hierarchical breakpoints, and spatial scale parameters (Song *et al.* 2020). The OPGD model encompasses factor, interaction, risk, and ecological detection. Our investigation predominantly employed the factor and interaction detectors. The *q*-value was determined using the following equations:

$$q=1-\frac{\sum_{h=1}^{L}N_h\sigma_h^2}{N\sigma^2}=1-\frac{\mathrm{SSW}}{\mathrm{SST}} \quad (1)$$

$$\mathrm{SSW}=\sum_{h=1}^{L}N_h\sigma_h^2,\ \mathrm{SST}=N\sigma^2 \quad (2)$$

where $h=1, ..., L$ is the strata for variable Y or factor X; N_h and N are the numbers of units in stratum h and the whole region, respectively; σ_h^2 and σ^2 are the variances of the variable Y for stratum h and the whole region, respectively; and SSW and SST represent the within-group sum of squares and the overall sum of squares, respectively. The *q*-value ranges from 0 to 1. A higher *q*-value indicates that the independent variable X provides more interpretive ability for feature Y. Interaction detection evaluates if the impacts of variables $X1$ and $X2$ on Y are independent or if they synergize to enhance or diminish the explanatory capability of the dependent variable Y. Potential interactions between the factors can be classified as shown in Table 2.

Grey relational analysis (GRA) method The GRA method evaluates the levels of correlation among system factors by examining the congruity between geometric configurations in data series and curve forms (Li *et al.* 2021). In this study, mean normalization was employed as the dimensionless treatment method. A smaller resolution coefficient corresponds to greater resolution, with optimal resolution achieved when the coefficient is ≤0.5463. Therefore, the resolution coefficient was set to 0.5 in this study. The primary procedural steps of GRA are delineated as follows:

(1) The comparison sequence (evaluation objects) and the reference sequence (evaluation criteria) were identified. In scenarios involving m evaluation objects and n evaluation indicators, the comparison sequence is defined as:

Table 2 Types of interactions

Judge theorem	Interaction
$q(X_1\cap X_2)<\mathrm{Min}(q-X_1,q-X_2)$	Nonlinear weakening
$\mathrm{Min}(q-X_1,q-X_2)<q(X_1\cap X_2)<\mathrm{Max}(q-X_1,q-X_2)$	Single-factor nonlinear attenuation
$q(X_1\cap X_2)>\mathrm{Max}(q-X_1,q-X_2)$	Double-factor enhancement
$q(X_1\cap X_2)=(q-X_1)+(q-X_2)$	Independence
$q(X_1\cap X_2)>(q-X_1)+(q-X_2)$	Nonlinear enhancement

$$[X_1'X_2'\cdots X_m']=\begin{bmatrix}x_1'(1) & x_2'(1)\cdots & x_m'(1)\\ x_1'(2) & x_2'(2)\cdots & x_m'(2)\\ \vdots & \vdots & \vdots\\ x_1'(n) & x_2'(n)\cdots & x_m'(n)\end{bmatrix} \quad (3)$$

The reference sequence is:

$$X_0'=(x_0'(1), x_0'(2), \cdots, x_0'(n))^T \quad (4)$$

(2) Data standardization uses two commonly applied methods: initialization and equalization. In this study, equalization was employed for the standardization process.

(3) The correlation coefficient was calculated as:

$$\xi_i(k)=\frac{\min_i\min_k|x_0(k)-x_i(k)|+\rho\max_i\max_k|x_0(k)-x_i(k)|}{|x_0(k)-x_i(k)|+\rho\max_i\max_k|x_0(k)-x_i(k)|} \quad (5)$$

where $\xi_i(k)$ is the relative difference between the comparison sequence and the reference sequence for the kth indicator; and ρ is the resolution factor, generally taken as 0.5.

(4) The weighted average of the correlation coefficients between each indicator and its corresponding element in the reference sequence was calculated. This aggregate measure is termed the correlation degree, denoted as:

$$ri=\frac{1}{m}\sum_{k=1}^{m}W(k)\xi_i(k) \quad (6)$$

where $W(k)$ is the weight that corresponds to the value of each indicator, which can be determined using the hierarchical analysis method, the expert scoring method, or the composite scoring method.

C-MDA method MDA serves as a feature selection technique within machine learning frameworks like random forests (Breiman 2001). Referring to the MDA method in that study, this study used the C-MDA method. This method denotes the

average change in both the root mean square error (RMSE) and the correlation coefficient (CC) when individual indicators are sequentially excluded from the model, relative to the complete set of indicators used for model construction.

Construction of predictive models: ridge regression The ridge regression method is a biased estimation technique for analyzing multicollinear data that was initially proposed by Hoerl and Kennard (1970). It uses L2 regularized weight decay to address instability in the ordinary least squares estimation caused by multicollinearity in independent variables, with the aim of streamlining the model and enhancing its robustness (Dar *et al.* 2023).

Model accuracy evaluation To comprehensively assess the yield prediction capability of the model, a Taylor diagram was employed to showcase the model results, with KGE and the coefficient of determination (R^2) being used as evaluation metrics. Taylor diagrams are widely used in model evaluation, with critical evaluation metrics including correlation coefficient (CC), standard deviation (SD), and RMSE (Taylor 2001). KGE integrates three factors between the modeled and observed data: CC, RMSE, and variability, so it is more comprehensive and accurate than any single metric for assessing model performance (Knoben *et al.* 2019). The closer KGE is to 1, the more effectively the model performs. KGE was calculated as follows:

$$KGE=1-\sqrt{(CC-1)^2+\left(\frac{\sigma_{sim}}{\sigma_{obs}}-1\right)^2+\left(\frac{\mu_{sim}}{\mu_{obs}}1\right)^2} \quad (7)$$

where σ_{obs} is the standard deviation of the observed data, σ_{sim} is the standard deviation of the model, μ_{sim} is average of the model, and μ_{obs} is average of the observed data.

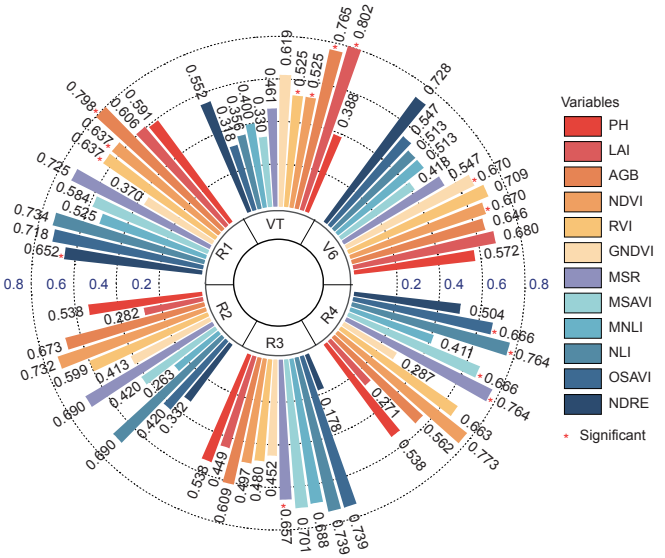


Fig. 2 *q*-values and significance of variables during six growth stages. V6, six-leaf stage; VT, tasseling stage; R1, silking stage; R2, blister stage; R3, milking stage; R4, dough stage. PH, plant height; LAI, leaf area index; AGB, above-ground biomass; NDVI, normalized difference vegetation index; RVI, ratio of vegetation index; GNDVI, green normalized difference vegetation index; MSR, modified simple ratio; MSAVI, modified soil adjusted vegetation index; MNLI, modified non-linear vegetation index; NLI, non-linear vegetation index; OSAVI, optimization soil adjusted vegetation index; NDRE, normalized difference red edge index.

3. Results

3.1. Relationship between mono-temporal data and maize yield

OPGD analysis results The effects of the 12 factors on maize yield varied across the six growth stages, with their corresponding *q*-values presented in Fig. 2. Overall, AGB, NDVI, RVI, and modified simple ratio (MSR) consistently showed high *q*-values (>0.460) across all stages, underscoring their robust interpretive efficacy for maize yield. During the six-leaf stage, all factors displayed *q*-values exceeding 0.500 except for modified soil adjusted vegetation index (MSAVI) (q=0.418). Notably, NDVI and GNDVI significantly influenced the yield in this stage (P<0.05). As the maize progressed through the tasseling, silking, and dough stages, four factors retained statistical significance, while MSR was the sole factor demonstrating significance during the milking stage. At the blister stage, various factors showed notable *q*-values for maize yield, including MSR and NLI at 0.690 each and NDVI at 0.732. However, the *q*-values for the blister stage did not meet the significance threshold, thereby limiting its explanatory power regarding yield.

The contribution proportions of the *q*-values for 12 factors across various growth stages of maize are shown in Fig. 3. At the six-leaf stage, NDRE exhibited the highest proportion, while NDVI dominated at the blister stage. During the milking stage, NLI and OSAVI each represented the highest proportion at 10.984%, while MSR demonstrated the greatest contribution during the dough stage. LAI during the tasseling stage and AGB during the silking stage had the highest *q*-value contributions in their respective growth periods and were found to be statistically significant (Figs. 2 and 3). This implied that these two factors are the most critical for maize yield during the tasseling and silking stages, respectively. At the tasseling stage, there was a notable reduction in the contribution rate of PH, while LAI was exactly the opposite during this stage. The contribution rate of AGB in each period was greater than 8%, with NDVI and RVI exhibiting highly consistent trends. GNDVI generally decreased throughout the whole stage, while MSR, MSAVI, and NLI showed overall increasing trends. OSAVI displayed increasing volatility in general, with changes presenting a complex pattern. The contribution rates of NDVI and NDRE decreased to their lowest levels in the milking stage, while modified non-linear vegetation index (MNLI) was exactly the opposite.

The variation in explanatory power of maize yield due to interactions between pairs of factors (increasing, decreasing, or independent of each other) is shown in Fig. 4. At the six-leaf stage, the double-factor enhancement type predominated, highlighting the effective synergistic enhancement of yield explanatory power by multiple factors. The minimum interaction explanatory power recorded during this period was 0.513, with only AGB∩NDRE exhibiting nonlinear attenuation. During the tasseling stage, the proportion of single-factor nonlinear attenuation types increased. Of the seven sets of factors exhibiting nonlinear weakening, six demonstrated an explanatory power less than 0.050. The percentage of double-

factor enhancement was higher during the silking stage than in the remaining five growth stages, and the interpretive power of the interactions between AGB and multiple factors even exceeded 90%. The blister stage exhibited four co-interaction types, with the explanatory power of LAI∩RVI reaching as high as 0.994. At the milking stage, NDVI∩NDRE, RVI∩NDRE, and GNDVI∩NDRE all exhibited nonlinear enhancement. In these three sets of data, the interactions between the two factors notably boosted the yield explanation, surpassing the combined interpretive power of the individual factors involved. The yield explanatory power of PH for the dough stage was enhanced through its interactions with the other 11 factors. The only instance of nonlinear enhancement observed at this growth stage also appeared in the co-interaction of PH∩LAI.

Results of GRA Grey correlations among the 12 indicators and maize yield across the six growth stages revealed correlation coefficients ranging from 0.510 to 0.865 (Fig. 5). The maximum values of the correlations during the six growth stages corresponded to GNDVI, GNDVI, PH, PH, PH, and PH, respectively, indicating a high degree of synchronization between these two indicators and maize yield. Throughout all growth stages, PH and MSAVI consistently demonstrated correlations with yield exceeding 0.650, while OSAVI consistently surpassed 0.720. Notably, NDVI and GNDVI consistently exhibited correlations exceeding 0.750 across all six growth stages, underscoring their robust association with yield and highlighting their pivotal role as significant variables in yield estimation.

3.2. Maize yield estimation based on mono-temporal data

Screening the mono-temporal feature variables The outcomes of feature variable selection for the six growth periods employing the OPGD, GRA, and C-MDA methods are shown Figs. 6–8. Each method used the top-ranked variable from each stage to standardize the importance of the other variables. The OPGD analysis outcomes indicated that MSR and NLI consistently demonstrated elevated importance values throughout the four phenological periods spanning from the silking to the dough stages. This underscores their significant association with yield, which suggests a pronounced correlation between these variables and crop productivity. The GRA method, focusing on the synchronized variations of indicators with yield, identified PH as the predominant factor throughout the four phenological periods spanning from the silking to the dough stages. Concurrently, the CPPs exhibited higher rankings during these intervals. Notably, NDVI and GNDVI displayed prominent importance values during the six-leaf and tasseling stages, and maintained substantial values across the other growth phases, underscoring their pivotal role as critical variables for maize yield prediction. Negative importance values obtained by using the C-MDA method indicated that model performance improved when the corresponding indicators were removed, compared to modeling with all indicators during these growth stages.

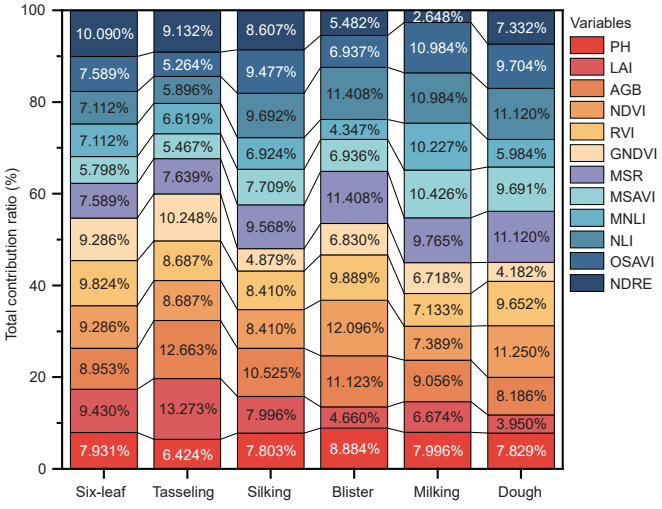


Fig. 3 Contribution proportions of the *q*-values of variables in different stages. PH, plant height; LAI, leaf area index; AGB, above-ground biomass; NDVI, normalized difference vegetation index; RVI, ratio of vegetation index; GNDVI, green normalized difference vegetation index; MSR, modified simple ratio; MSAVI, modified soil adjusted vegetation index; MNLI, modified non-linear vegetation index; NLI, non-linear vegetation index; OSAVI, optimization soil adjusted vegetation index; NDRE, normalized difference red edge index.

The three methods above were used for ranking the importance of feature variables, while the selection of modeling variables was based on RMSE. In the variable selection process, theoretically, the number of input variables corresponding to the minimum RMSE is considered the optimal variable size. However, when determining the appropriate number of feature variables, one must consider the accuracy, robustness, and computational efficiency of the predictive model, with the aim of achieving high modeling precision with a limited number of variables (Kearns and Ron 1999). Significant advances in RMSE convergence efficiency have been attained through prioritization *via* the C-MDA method. A comparative analysis revealed that indicators ranked by the GRA method exhibit notably enhanced stability in RMSE reduction compared to those prioritized using the OPGD method. Within a limited number of variables, the RMSE of the model typically decreased as the number of input variables increased. In this context, both the robustness of the model and the importance of the inputs were considered when determining the number of variables for yield prediction models at various growth stages. The OPGD method set the numbers of variables to 7, 7, 5, 5, 4, and 3 for the six-leaf, tasseling, silking, blister, milking, and dough stages, respectively. The GRA method set the numbers of variables to 7, 5, 5, 6, 5, and 3, while the C-MDA method set the numbers of variables to 7, 5, 6, 4, 3, and 6. In summary, the variables selected and the size of model inputs differed at each growth stage for the various methods, which indicated that the suitability of variables for maize yield prediction is greatly influenced by both the maize growth stage and the feature selection methods employed.

Predictive model performance based on mono-temporal data The precision of crop yield forecasting models after data sorting by the three methods of OPGD, GRA, and C-MDA through the Taylor diagram, as well as the KGE

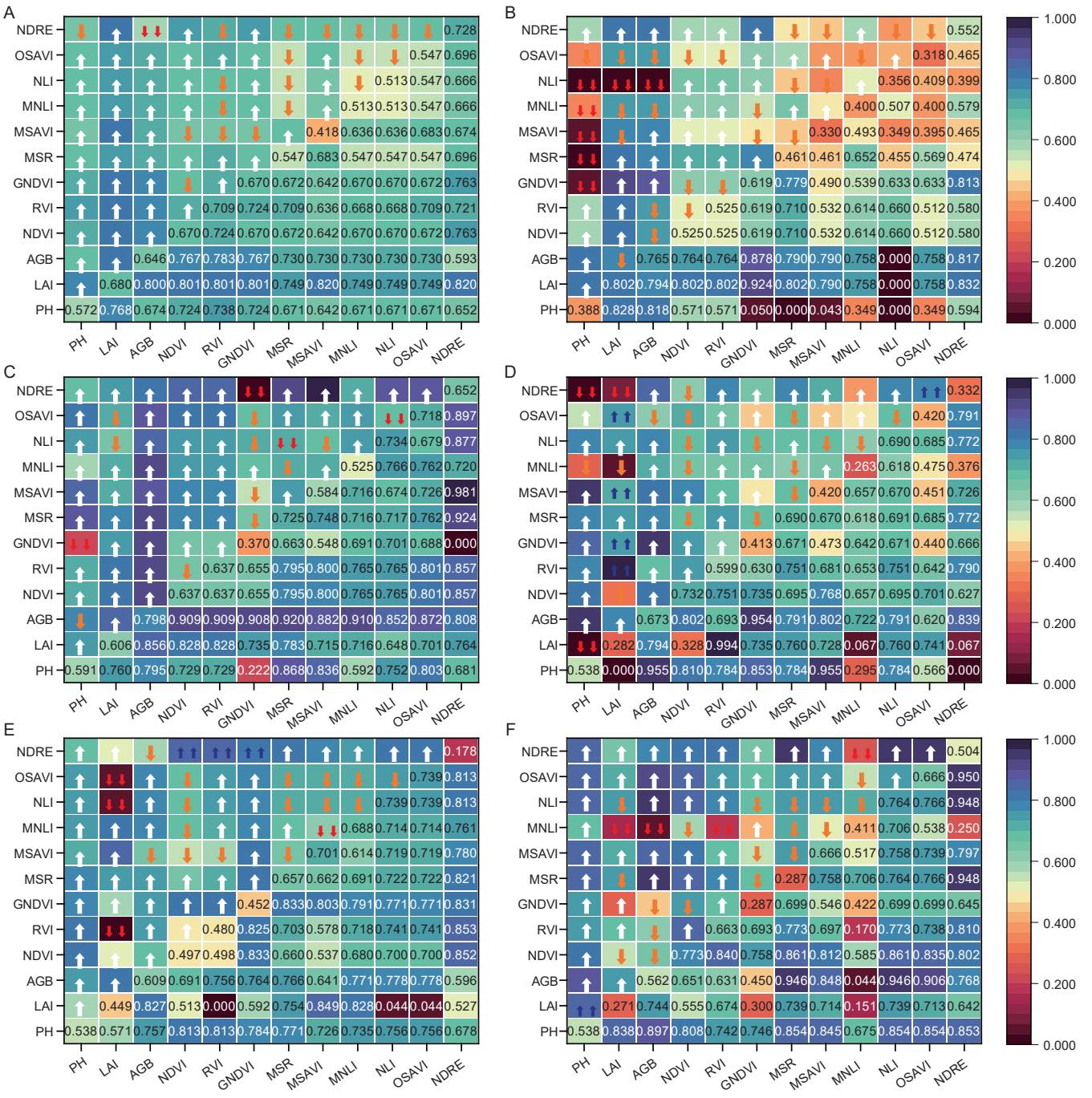


Fig. 4 Interactions and types of variables on yield within six maize growth stages. A, six-leaf stage. B, tasseling stage. C, silking stage. D, blister stage. E, milking stage. F, dough stage. ⇊, nonlinear weakening; ↓, single-factor nonlinear attenuation; ↑, double-factor enhancement; ⇈, nonlinear enhancement. PH, plant height; LAI, leaf area index; AGB, above-ground biomass; NDVI, normalized difference vegetation index; RVI, ratio of vegetation index; GNDVI, green normalized difference vegetation index; MSR, modified simple ratio; MSAVI, modified soil adjusted vegetation index; MNLI, modified non-linear vegetation index; NLI, non-linear vegetation index; OSAVI, optimization soil adjusted vegetation index; NDRE, normalized difference red edge index.

metrics, are shown in Fig. 9. The CC of the models obtained from all three methods exceeded 0.7, with RMSE values staying below 0.8 t ha^{-1}. The Taylor diagram results revealed that among the models obtained using the OPGD method, the model for the tasseling stage demonstrated a notable advantage, with a CC of 0.910 and an RMSE value of 0.493 t ha^{-1}. Conversely, the model for the milking stage performed the worst, with an RMSE value of 0.746 t ha^{-1}. For the models obtained using the GRA method, the accuracies of the models for the silking stage through the dough stage were similar. The model for the six-leaf stage performed the least satisfactorily among the 18 models, with a CC of 0.744. The C-MDA method demonstrated superior overall performance compared to the other two methods, with RMSE values for the models at four stages remaining below 0.6 t ha^{-1} and the CC for the dough stage reaching 0.940. The KGE

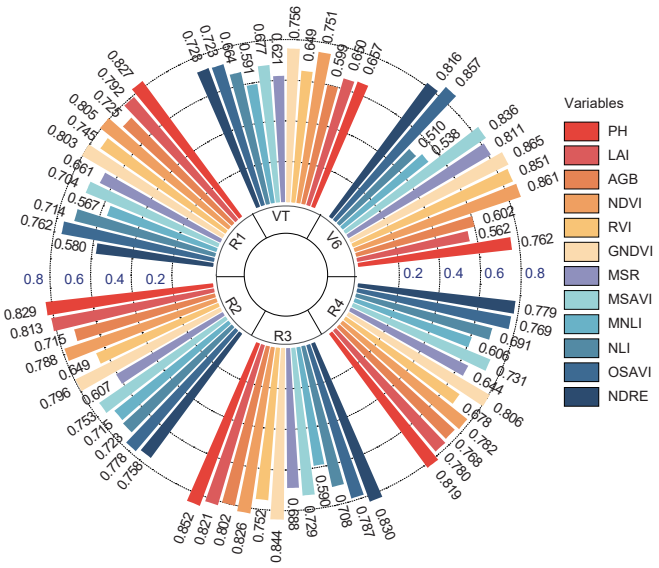


Fig. 5 Grey correlations of variables on yield over the six growth stages of maize. V6, six-leaf stage; VT, tasseling stage; R1, silking stage; R2, blister stage; R3, milking stage; R4, dough stage. PH, plant height; LAI, leaf area index; AGB, above-ground biomass; NDVI, normalized difference vegetation index; RVI, ratio of vegetation index; GNDVI, green normalized difference vegetation index; MSR, modified simple ratio; MSAVI, modified soil adjusted vegetation index; MNLI, modified non-linear vegetation index; NLI, non-linear vegetation index; OSAVI, optimization soil adjusted vegetation index; NDRE, normalized difference red edge index.

values provided a holistic assessment of the performance of each model. When aggregating the KGE values for the models obtained using the three methods, the performance across growth stages was ordered from the highest to lowest as: dough stage>silking stage>tasseling stage>blister stage>milking stage>six-leaf stage. The trend of R^2 was consistent with KGE. The maximum values of R^2 obtained from modeling the data for the six growth stages under the three sorting methods were 0.616, 0.810, 0.804, 0.701, 0.718, and 0.871, respectively. During the dough stage, maize rapidly accumulates starch and nutrients, making observations during this period crucial for highly accurate maize yield predictions. The timeliness of yield estimation is critically important in practical production as it can directly influence subsequent management and decision-making processes. This indicates that the periods around flowering are the most critical for maize yield prediction, highlighting the significant importance of observing various indicators during these stages for achieving high-accuracy yield estimates.

3.3. Maize yield estimation based on multi-temporal data

Generation of multi-temporal data groups Based on the accuracy of yield predictions across the six growth stages detailed in Section 3.2, three multi-temporal datasets were subsequently chosen. These datasets comprised two, three, and four mono-temporal phases, respectively (Fig. 10). These datasets were categorized as bi-temporal, tri-temporal, and quadri-temporal, with 12, 24, and 36 variables, respectively. Variables from the same growth stage are distinguished by naming conventions; for example, R1-X12 refers to the NDRE variable for the dough stage. Building upon the comparison of the three methods discussed in Section 3.2, the C-MDA method was employed for indicator selection in the multi-temporal datasets for yield estimation.

Screening of multi-temporal feature variables The top 20 indicators were selected based on their importance values from the three datasets using the C-MDA method, and the variation of RMSE with the number of model inputs is shown in Fig. 11. Note that the two CPPs significantly influenced both the bi-temporal and tri-temporal datasets. In the tri-temporal dataset, the indicators for dough stage-AGB and tasseling stage-PH had the greatest importance values. In the quadri-temporal dataset, the disparities in importance values among the top 10 indicators were less pronounced compared to the other two datasets. The numbers of model inputs for the three datasets were established as 11, 13, and 11, respectively. Although the RMSE generally decreased with an increase in the amount of input data across all three datasets, the rate of decrease was slower for the quadri-temporal dataset when the number of model inputs was less than 19.

Maize yield estimation based on multi-temporal data As shown in Fig. 12, maize yield prediction models were developed for the three multi-temporal datasets. CC ranged from 0.860 to 0.975, while the RMSE ranged from 0.580 to 0.291 t ha^{-1}. The model constructed using data from the dough and silking stages exhibited a CC value greater than all the mono-temporal data modeling results in Section 3.2, while the corresponding RMSE value was just the opposite. The combination of these two growth stages was shown to enhance the effectiveness of the yield estimation model. Upon incorporating data from the tasseling stage, the model accuracy was further improved, with the RMSE being reduced to below 0.3 t ha^{-1} and the CC value reaching 0.974. Although the addition of data from the blister stage continued to yield a model accuracy higher than those of most mono-temporal modeling results, a notable decline in performance was observed when compared to the accuracy achieved using the bi- and tri-temporal data combinations for model development. In the results of bi- and tri-temporal data modeling, the KGE values were greater than 0.8, and the R^2 values were 0.910 and 0.942, respectively. In contrast, with quadri-temporal data, the R^2 value was only 0.737. These results indicated that the combination of the dough, silking, and tasseling stages is crucial for ensuring the accuracy of yield predictions.

4. Discussion

4.1. Optimizing feature variable selection

The screening and ranking of feature variables facilitate the optimization of the modeling process and enhance model interpretability. The use of OPGD, GRA, and C-MDA techniques in this study effectively reduced dimensionality and bolstered modeling efficiency. Compared to traditional GD models, the OPGD Model employed here enables the ideal integration of spatial data discretization methods,

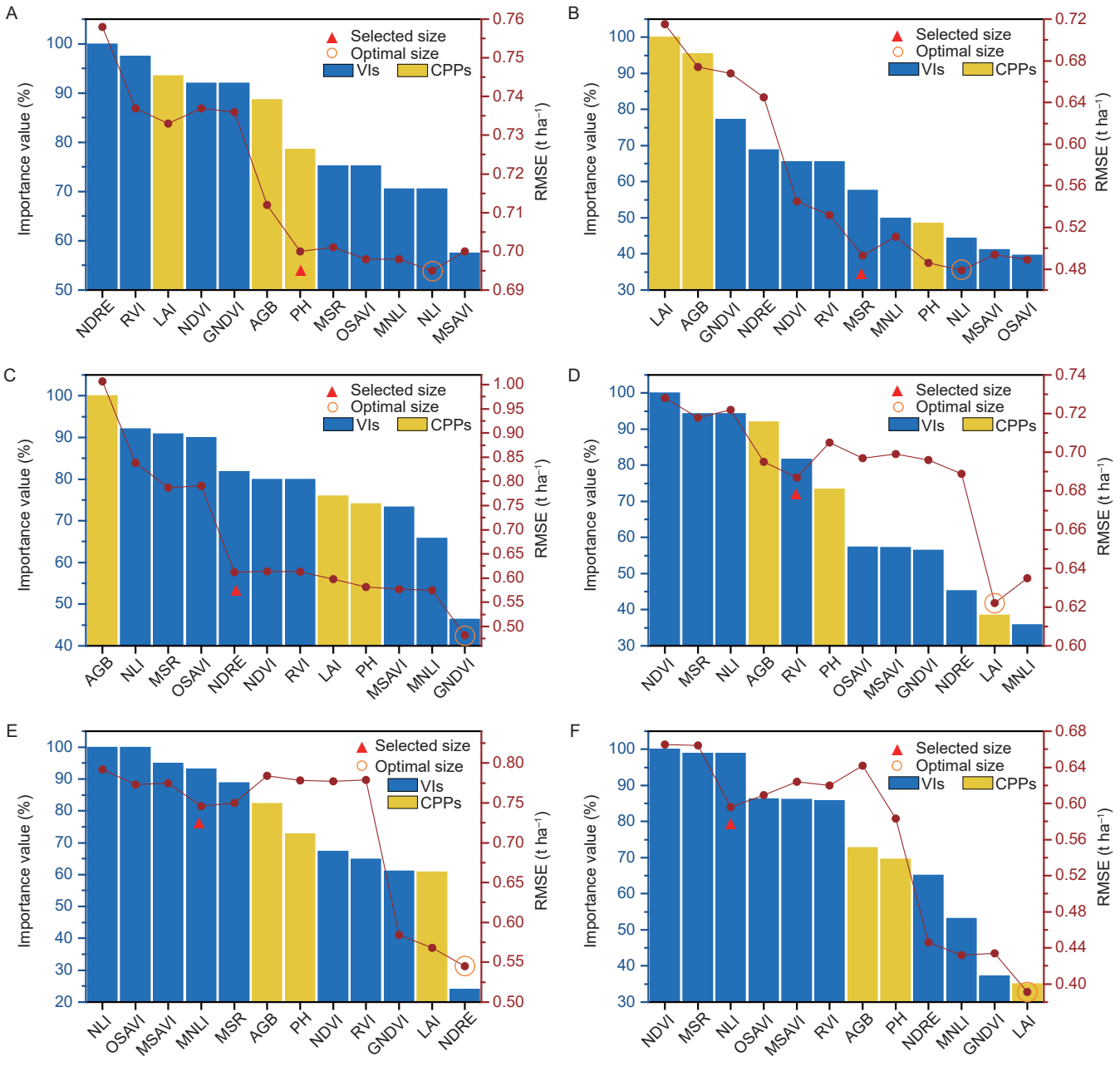


Fig. 6 Ranking of feature variable importance values based on the optimal parameters-based geographical detector (OPGD) method and changes in RMSE with an increasing number of input variables. A, six-leaf stage. B, tasseling stage. C, silking stage. D, blister stage. E, milking stage. F, dough stage. NDRE, normalized difference red edge index; RVI, ratio of vegetation index; LAI, leaf area index; NDVI, normalized difference vegetation index; GNDVI, green normalized difference vegetation index; AGB, above-ground biomass; PH, plant height; MSR, modified simple ratio; MSAVI, modified soil adjusted vegetation index; MNLI, modified non-linear vegetation index; NLI, non-linear vegetation index; OSAVI, optimization soil adjusted vegetation index; NDRE, normalized difference red edge index. VIs, vegetation indices; CPPs, crop phenotypic parameters.

spatial stratification interval counts, and spatial scale variables, thereby providing a more precise, flexible, and efficient solution for geospatial variability assessment (Zhang *et al.* 2022). The OPGD Model has been used extensively in studies on the impacts of topography, climate, and anthropogenic activities on evapotranspiration (Ablikim *et al.* 2023), the kernel normalized difference vegetation index (Gu *et al.* 2024), and other applications. We used the OPGD Model for feature variable screening and ranking, which is a new approach for applying aspects of the OPGD Model. The results obtained from this approach showed that when only 1–2 variables were used for modeling, the RMSE of the models was generally lower than that achieved with the two other feature selection methods. This indicated a clear advantage of the OPGD model in identifying the most influential factors. Furthermore, the interaction detection results revealed that combinations of multiple factors could significantly enhance the explanatory power of the model for yield. Consequently, in future research, we will consider incorporating variable pairs with high *q*-values from the interaction detection phase as

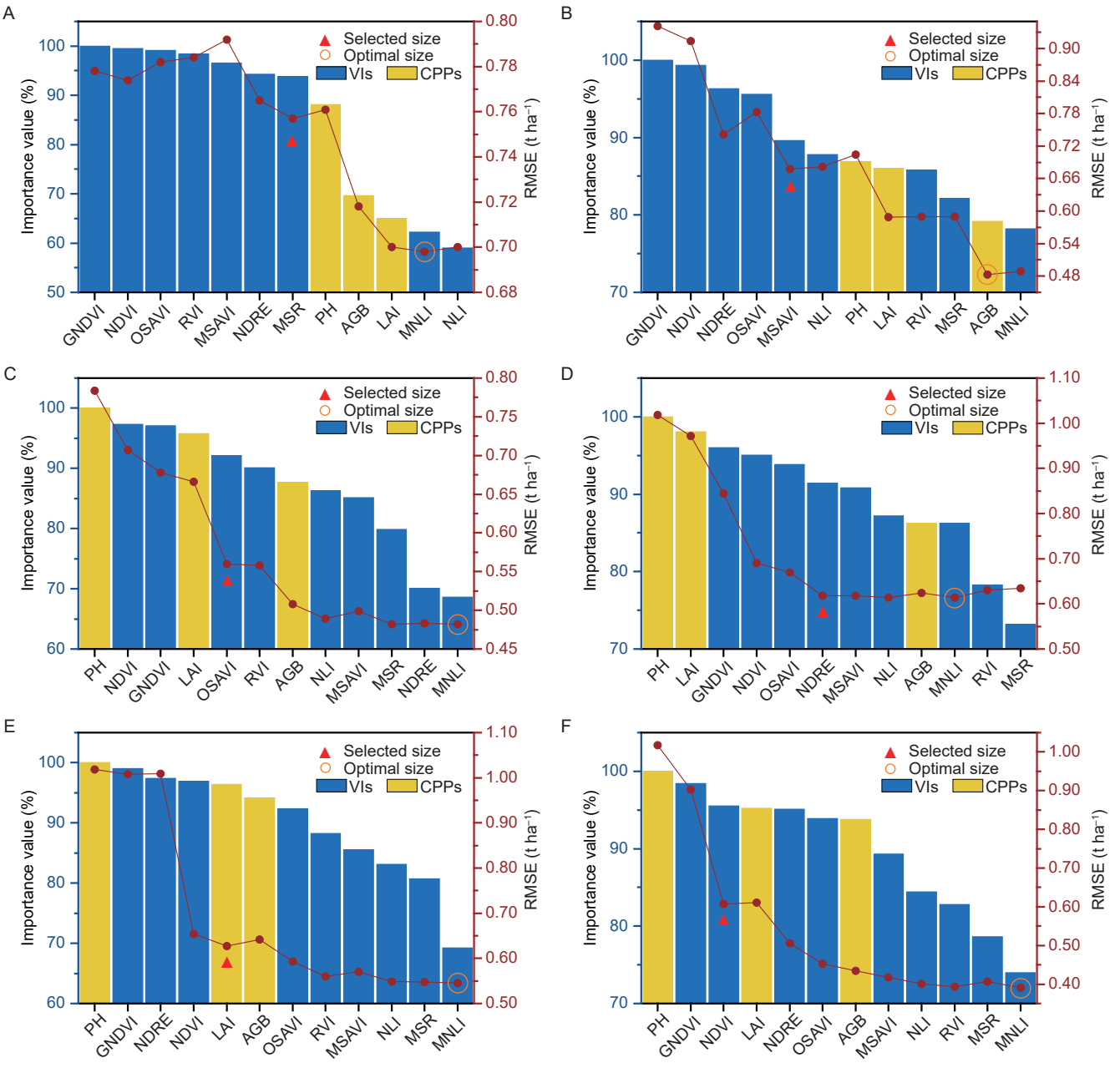


Fig. 7 Ranking of feature variable importance values based on the grey relational analysis (GRA) method and changes in RMSE with an increasing number of input variables. A, six-leaf stage. B, tasseling stage. C, silking stage. D, blister stage. E, milking stage. F, dough stage. GNDVI, green normalized difference vegetation index; NDVI, normalized difference vegetation index; OSAVI, optimization soil adjusted vegetation index; RVI, ratio of vegetation index; MSAVI, modified soil adjusted vegetation index; NDRE, normalized difference red edge index; MSR, modified simple ratio; PH, plant height; AGB, above-ground biomass; LAI, leaf area index; MNLI, modified non-linear vegetation index; NLI, non-linear vegetation index. VIs, vegetation indices; CPPs, crop phenotypic parameters.

feature groups in the model. Compared to the OPGD Model, using the GRA method for feature variable ranking exhibited a more stable trend of decreasing RMSE with an increasing number of input variables. This stability could be attributed to the GRA method's advantage in uncovering underlying structures and tendencies (Ikram *et al.* 2020). In contrast to the approaches of Ramos *et al.* (2020), who relied exclusively on the R^2 metric, and Yang *et al.* (2022), who solely employed the MDA metric for feature variable selection, the C-MDA method integrated both CC and RMSE metrics to improve the efficiency of the feature variable selection process. The performance of the three methods exhibited distinct advantages and disadvantages across the various growth periods. Nevertheless, the C-MDA method had the best sorting effect in general, indicating that the C-MDA method we defined and adopted was a reasonable and effective sorting method.

The feature selection results supported the selection of both VIs and CPPs for yield estimation. In the tri-temporal dataset, two of the CPPs had the highest importance

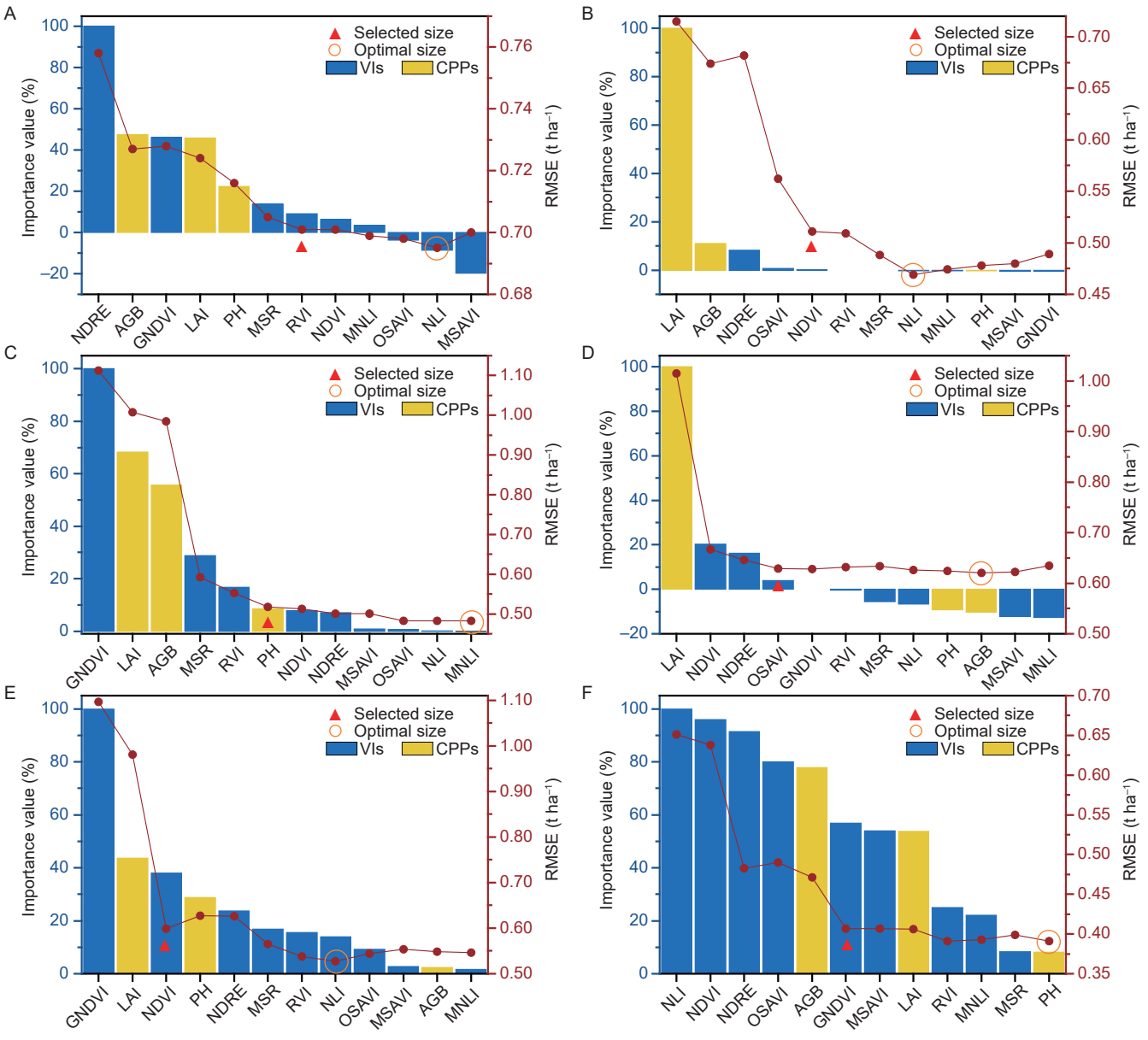


Fig. 8 Ranking of feature variable importance values based on the custom mean decrease accuracy (C-MDA) method and changes in RMSE with an increasing number of input variables. A, six-leaf stage. B, tasseling stage. C, silking stage. D, blister stage. E, milking stage. F, dough stage. NDRE, normalized difference red edge index; AGB, above-ground biomass; GNDVI, green normalized difference vegetation index; LAI, leaf area index; PH, plant height; MSR, modified simple ratio; RVI, ratio of vegetation index; NDVI, normalized difference vegetation index; MSAVI, modified soil adjusted vegetation index; MNLI, modified non-linear vegetation index; NLI, non-linear vegetation index; OSAVI, optimization soil adjusted vegetation index. VIs, vegetation indices; CPPs, crop phenotypic parameters.

among the selected variables. This result demonstrated the importance of CPPs in guaranteeing yield forecasting precision, which was consistent with past research showing that PH, LAI, and AGB are effective in yield prediction (Nada *et al.* 2016; Yu *et al.* 2019; Ji *et al.* 2023). As detailed in a study by Zhang *et al.* (2024), CPPs were reported to exhibit significantly positive correlations with grain filling rates. Consequently, CPPs can be used as a supplement in cases where VIs are saturated and sensitivity is lost during the reproductive stage of maize (Yue *et al.* 2018). This finding revealed the rationale behind the selection of CPPs for yield prediction in this study. On the other hand, our results also demonstrated that VIs were more influential in yield prediction than CPPs. This was reflected in the higher importance values assigned to VIs and the greater number of VIs selected for the prediction model.

The optimal CPPs for yield prediction using mono-temporal data were identified as LAI and AGB, which generally aligned with the conclusions of Li *et al.* (2014) and Yu *et al.* (2019). While the study by John and Reddy (2019) revealed a substantially positive link between yield and PH, our research demonstrated that PH had a lower impact on yield than LAI and AGB. Among the VIs, the three most effective spectral indices for predicting maize yield were identified as GNDVI,

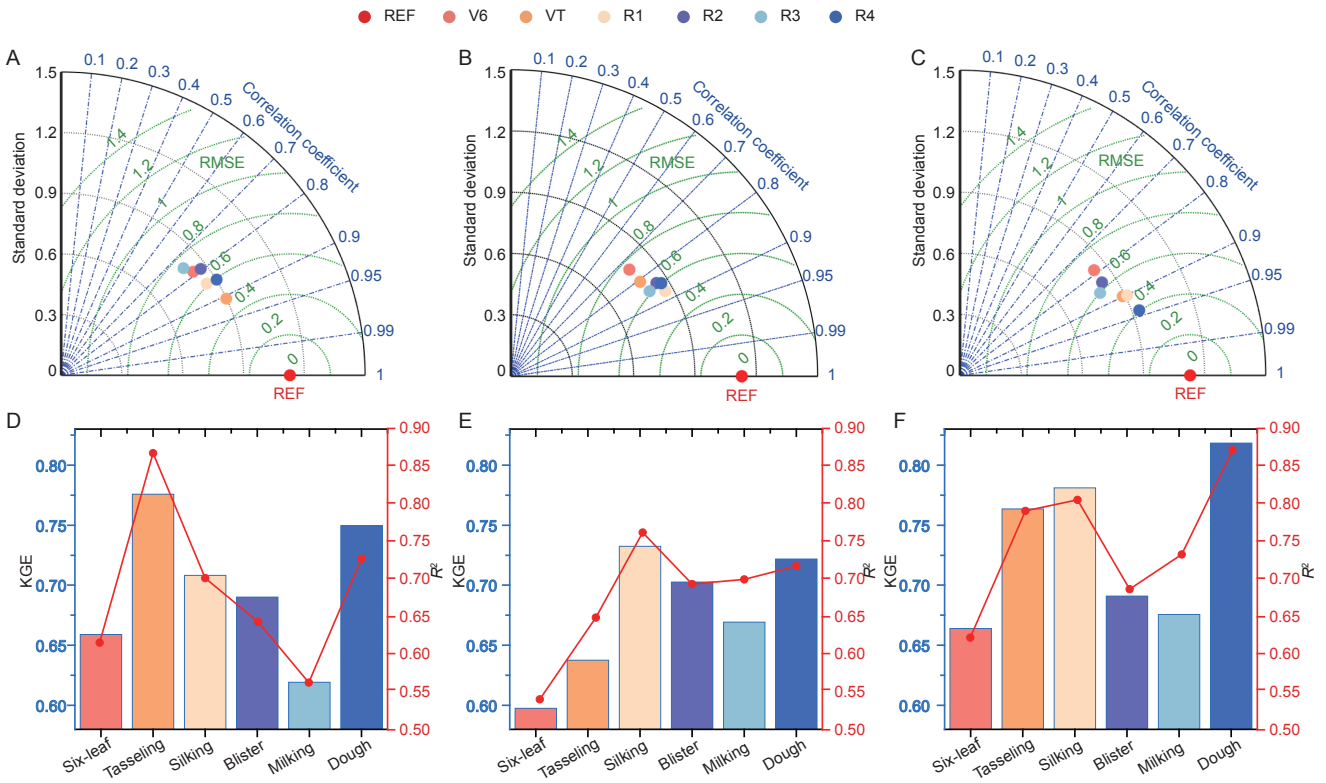


Fig. 9 Accuracy of yield predictions using mono-temporal data. A and D, results of optimal parameters-based geographical detector (OPGD). B and E, results of the grey relational analysis (GRA) method. C and F, results of custom mean decrease accuracy (C-MDA). REF, reference; V6, six-leaf stage; VT, tasseling stage; R1, silking stage; R2, blister stage; R3, milking stage; R4, dough stage. KGE, Kling-Gupta efficiency.



Fig. 10 Compositions of the multi-temporal datasets. V6, six-leaf stage; VT, tasseling stage; R1, silking stage; R2, blister stage; R3, milking stage; R4, dough stage.

NDVI, and NDRE. While the sequencing of these variables varied between the research of Ramos *et al.* (2020) and Yang *et al.* (2022), their top three choices were aligned with our findings. This consistency indicates that these three variables are important and effective for yield estimation. Whereas AGB was the most important CPPs in multi-temporal data in our study, RVI, MSAVI, and OSAVI also served a crucial function in addition to NDRE and GNDVI. These results emphasize the critical importance of using suitable CPPs and VIs for the accurate prediction of crop yield.

4.2. UAV-based study on the optimal temporal period for maize yield prediction

To establish a suitable data-driven yield prediction model, data must be obtained during the growth stages that have a decisive influence on crop yield (Qader *et al.* 2018). Early developmental stages have only minimal impacts on grain number and weight, thus exerting a lesser influence on yield (Széles *et al.* 2012). This study also found that the modeling accuracy during the six-leaf stage was the lowest among the six growth stages considered. The dough period gave the most accurate forecasts in this study when accuracy was considered. In terms of optimal observation periods, the tasseling stage was the earliest to successfully forecast maize yield, while the adjacent silking stage demonstrated the best performance when both accuracy and early forecasting capability were considered. These findings agree with the outcomes of Uno *et al.* (2005), García-Martínez *et al.* (2020), and Su *et al.* (2023), further underscoring the importance of the pre-flowering and post-flowering periods for accurate yield prediction.

The model performance demonstrated a notable improvement with the bi-temporal data compared to the

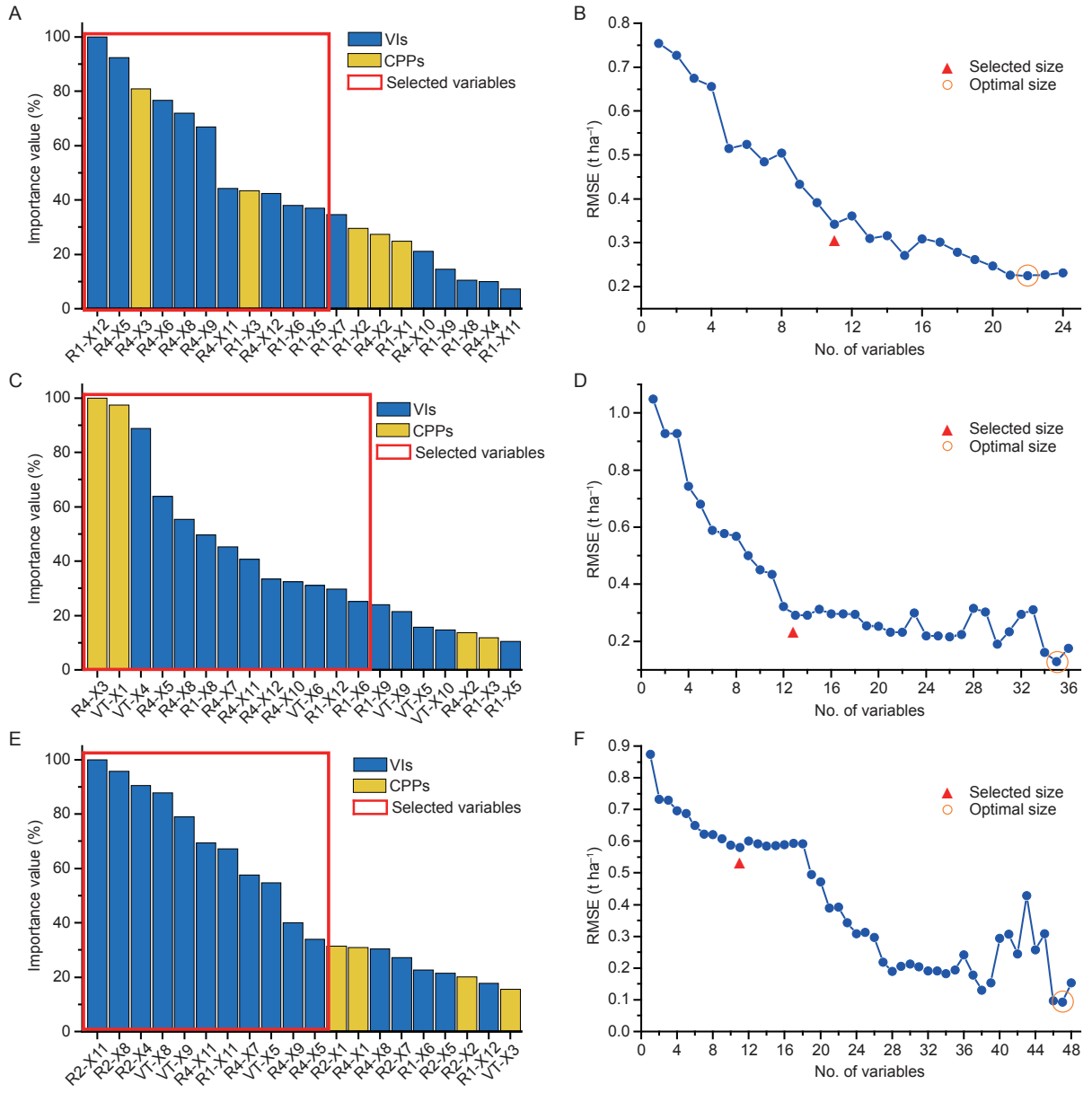


Fig. 11 Significant values and RMSE (t ha^{-1}) of the top 20 ranked variables in the predictive model with an increasing number of input variables when using each of the three sets of multi-temporal data. Corresponding relationships of indicators: VT, tasseling stage; R1, silking stage; R2, blister stage; R4, dough stage; X1, plant height (PH); X2, leaf area index (LAI); X3, above-ground biomass (AGB); X4, normalized difference vegetation index (NDVI); X5, ratio of vegetation index (RVI); X6, green normalized difference vegetation index (GNDVI); X7, modified simple ratio (MSR); X8, modified soil adjusted vegetation index (MSAVI); X9, modified non-linear vegetation index (MNLI); X10, non-linear vegetation index (NLI); X11, optimization soil adjusted vegetation index (OSAVI); X12, normalized difference red edge index (NDRE). VIs, vegetation indices; CPPs, crop phenotypic parameters.

mono-temporal data. Specifically, the R^2 value increased from 0.871 to 0.909, reflecting a 4.363% enhancement. Concurrently, the RMSE decreased from 0.407 t ha^{-1} to 0.342 t ha^{-1}, indicating a reduction of 15.971%. Furthermore, the RMSE of the model built from the tri-temporal data decreased to 0.291 t ha^{-1}, a reduction of 28.501%. This marked decline in RMSE showed that multi-temporal observations provide more comprehensive crop growth information and mitigate the bias of mono-temporal observations. Integrating multi-temporal data could significantly increase the reliability of yield estimates, and this finding agreed with the conclusions of previous studies (Zhou *et al.* 2017). Specifically, the combination of three growth stages — tasseling, silking, and dough — reached the greatest level of prediction accuracy with minimal observations and was thus identified as the most effective UAV observation set for maize yield forecasting. In contrast to our results, García-Martínez *et al.* (2020) concluded that the chlorophyll concentration of foliage gradually decreased during the dough and maturity periods as assimilates were gradually transferred to the kernel. This process led to a gradual weakening of the relationship between canopy spectral

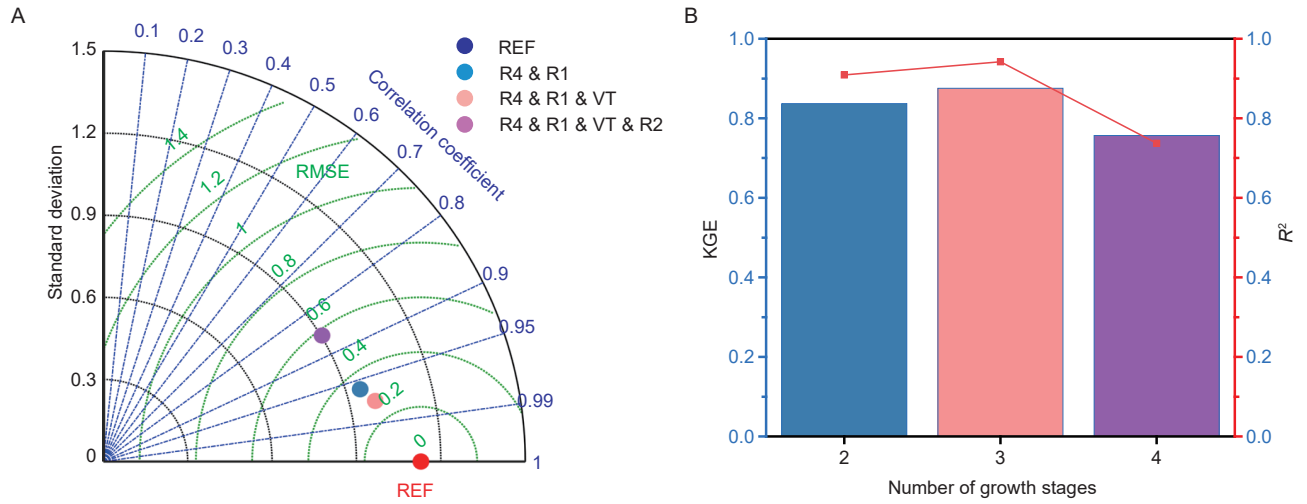


Fig. 12 Accuracy of yield predictions using the bi-, tri-, and quadri-temporal datasets. REF (reference), measured yield; KGE, Kling-Gupta efficiency. VT, tasseling stage; R1, silking stage; R2, blister stage; R4, dough stage.

reflectance and crop production. One potential cause of this discrepancy may be the introduction of the CPP variables in our study. In addition, the RMSE of quadri-temporal data was significantly greater than the modeling results based on bi-temporal and tri-temporal data. This indicated that the tri-temporal data were sufficient to portray the relationship between variables and maize yield.

This study compared three methodologies for factor screening to propose new ideas for elucidating the impacts of CPPs and VIs on maize yield while enabling a structured approach to modeling input factors for yield estimation. The optimal modeling stage was determined by the comparison of six mono-temporal yield estimation models. Meanwhile, the combination of mono-temporal data improved the modeling accuracy and provided scientific guidance for field-derived maize yield estimation. In addition, the effects of canopy structure variability (Cen *et al.* 2019) and abundance (Duan *et al.* 2019) on yield have also been demonstrated. Therefore, future research will focus on a multi-year experimental approach aimed at optimizing the number of factors and integrating diverse variables to achieve high-precision maize yield predictions at the earliest feasible growth stages.

4.3. Application prospects of this study for predicting crop yield in smallholder fields

Currently, many scholars are researching yield estimation for farms and smallholder plots, but most studies often involve the strict control of water and fertilizer management in the experiments (Liu Y *et al.* 2023; Pokhrel *et al.* 2023). In contrast to this study where no special water and fertilizer application control was implemented, Peng *et al.* (2021) conducted deficit irrigation at varying degrees in the study area, while Yang G *et al.* (2024) applied different amounts of nitrogen fertilizer to the fields. These control measures imposed certain limitations in the yield prediction capability of the estimation methods proposed in those studies under uniform water and fertilizer management conditions (Kumar *et al.* 2023; Killeen *et al.* 2024). In contrast, the yield estimation method proposed in this study identifies the key factors and phenological stages and constructs a simplified model, thus reducing monitoring costs while improving the accuracy of yield estimation and enhancing farm management efficiency.

The study area is located at the boundary area of Xianyang City and Yangling District in Shaanxi Province. Based on data provided by the *Shaanxi Statistical Yearbook 2024* (https://tjj.shaanxi.gov.cn/tjsj/ndsj/tjnj/sxtjnj/index.html?2024), we compared the statistical yields, measured yields, and estimated yields for both Xianyang City and Yangling District (Fig. 13). The results showed that the average maize yield in Yangling District was higher than that of Xianyang City. The measured yield mean for the selected plots was slightly higher than the average level in Yangling District, but the overall difference remained within an acceptable range. The model-estimated yields were slightly lower than the measured yields. However, the differences between the statistical yield, measured yield, and estimated yield all fall within the allowable margin of error, indicating that the results of this study have a high degree of accuracy. By using only four plots for sampling, we successfully obtained high-precision yield estimates while effectively reducing observation costs. This outcome provides powerful support for smallholder farmers in developing scientific planting plans and marketing strategies.

Although our current research primarily focuses on yield estimation of maize in Northwest China, we hope to achieve high estimate precision while extending the scope by combining ground-truthing data, UAVs, and satellite remote sensing data in future studies. In particular, more specific and customized yield estimation information will be provided within small farmland areas.

5. Conclusion

This study used multispectral data obtained from a UAV to calculate VIs, while concurrently integrating CPPs gathered from field experiments. Three methodologies were employed to rank and filter these variables. Selected

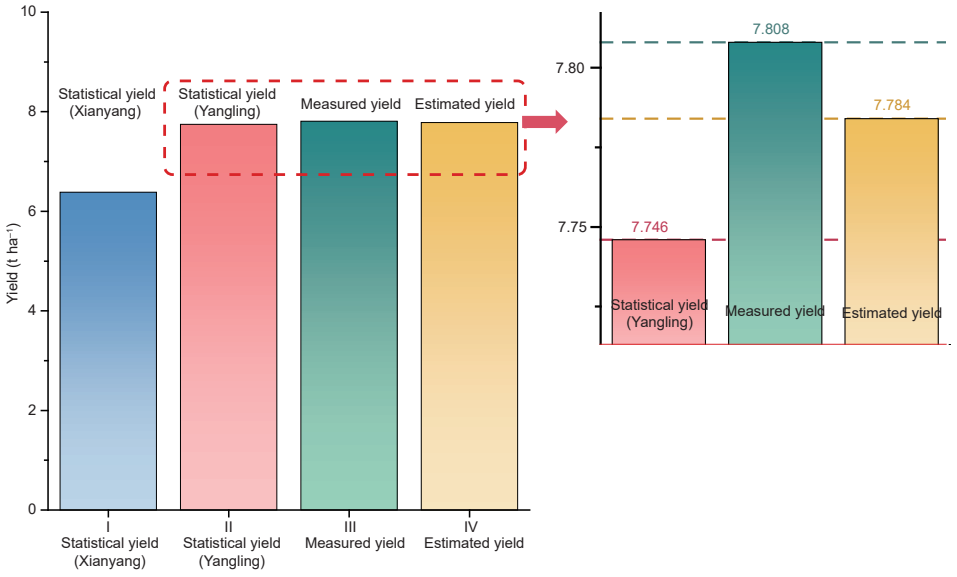


Fig. 13 Statistical yield, measured yield, and estimated yield of maize. Xianyang and Yangling, Shaanxi Province, China.

feature variables were used to construct ridge regression models for the data from six mono-temporal and three multi-temporal data groups. The potential of these data for predicting maize yield across different growth stages was analyzed. Among the three methods for factor selection and ranking, the C-MDA method proved to be the most efficient. The three optimal VIs for predicting maize yield were identified as GNDVI, NDVI, and NDRE, while LAI and AGB proved effective as CPPs. In the mono-temporal maize yield forecast, the tasseling stage was the earliest suitable period. The yield prediction model derived from data collected at the dough stage exhibited the highest accuracy. The integration of UAV data from different crop growth stages enhanced the accuracy of the yield prediction models. Among the three multi-temporal datasets, the model based on data from the tasseling, silking, and dough stages exhibited the best performance. Based on these findings, the precise estimation of maize yields in smallholder fields should be feasible, which holds both theoretical significance and practical value for the advancement of precision agriculture.

Acknowledgements

The study was funded by the National Natural Science Foundation of China (U2243235 and 52309060).

Declaration of competing interest

The authors declare that they have no conflict of interest.

Declaration of generative AI and AI-assisted technologies in the writing process

The authors declare that they did not use AI in the preparation and writing of this manuscript.

References

Ablikim K, Yang H, Mamattursun A. 2023. Spatiotemporal variation of evapotranspiration and its driving factors in the Urumqi River Basin. *Sustainability*, **15**, 13904.

Bénard C, Da Veiga S, Scornet E. 2022. Mean decrease accuracy for random forests: Inconsistency, and a practical solution *via* the Sobol-MDA. *Biometrika*, **109**, 881–900.

Bian C, Shi H, Wu S, Zhang K, Wei M, Zhao Y, Sun Y, Zhuang H, Zhang X, Chen S. 2022. Prediction of field-scale wheat yield using machine learning method and multi-spectral UAV data. *Remote Sensing*, **14**, 1474.

Breiman L. 2001. Random forests. *Machine Learning*, **45**, 5–32.

Cen H, Wan L, Zhu J, Li Y, Li X, Zhu Y, Weng H, Wu W, Yin W, Xu C, Bao Y, Feng L, Shou J, He Y. 2019. Dynamic monitoring of biomass of rice under different nitrogen treatments using a lightweight UAV with dual image-frame snapshot cameras. *Plant Methods*, **15**, 32.

Chen J M. 1996. Evaluation of vegetation indices and a modified simple ratio for boreal applications. *Canadian Journal of Remote Sensing*, **22**, 229–242.

Dar I S, Chand S, Shabbir M, Kibria B M G. 2023. Condition-index based new ridge regression estimator for linear regression model with multicollinearity. *Kuwait Journal of Science*, **50**, 91–96.

Dee J R, Adams H D, Palmer M W. 2018. Belowground annual ring growth coordinates with aboveground phenology and timing of carbon storage in two tallgrass prairie forb species. *American Journal of Botany*, **105**, 1975–1985.

Devkota K P, Bouasria A, Devkota M, Nangia V. 2024. Predicting wheat yield gap and its determinants combining remote sensing, machine learning, and survey approaches in rainfed Mediterranean regions of Morocco. *European Journal of Agronomy*, **158**, 127195.

Duan B, Fang S, Zhu R, Wu X, Wang S, Gong Y, Peng Y. 2019. Remote estimation of rice yield with unmanned aerial vehicle (UAV) data and spectral mixture analysis. *Frontiers in Plant Science*, **10**, 00204.

Fu T, Tian S, Zhan Q. 2024. Phenological analysis and yield estimation of rice based on multi-spectral and SAR data in Maha Sarakham, Thailand. *Journal of Spatial Science*, **69**, 149–165.

García-Martínez H, Flores-Magdaleno H, Ascencio-Hernández R, Khalil-Gardezi A, Tijerina-Chávez L, Mancilla-Villa O R, Vázquez-Peña M A. 2020. Corn grain yield estimation from vegetation indices, canopy cover, plant density, and a neural network using multispectral and RGB images acquired with unmanned aerial vehicles. *Agriculture*, **10**, 277.

Gilliot J M, Michelin J, Hadjard D, Houot S. 2021. An accurate method for predicting spatial variability of maize yield from UAV-based plant height estimation: A tool for monitoring agronomic field experiments. *Precision Agriculture*, **22**, 897–921.

Gitelson A A, Merzlyak M N, Grits Y. 1997. Remote estimation of chlorophyll content in higher plant leaves. *International Journal of Remote Sensing*, **18**, 2691–2697.

Goel N S, Qin W. 1994. Influences of canopy architecture on relationships between various vegetation indices and LAI and Fpar: A computer simulation. *Remote Sensing Reviews*, **10**, 309–347.
Gong Y, Duan B, Fang S, Zhu R, Wu X, Ma Y, Peng Y. 2018. Remote estimation of rapeseed yield with unmanned aerial vehicle (UAV) imaging and spectral mixture analysis. *Plant Methods*, **14**, 70.
Gu Z, Chen X, Ruan W, Zheng M, Gen K, Li X, Deng H, Chen Y, Liu M. 2024. Quantifying the direct and indirect effects of terrain, climate and human activity on the spatial pattern of kNDVI-based vegetation growth: A case study from the Minjiang River Basin, Southeast China. *Ecological Informatics*, **80**, 102493.
Guan K, Wu J, Kimball J S, Anderson M C, Frolking S, Li B, Hain C R, Lobell D B. 2017. The shared and unique values of optical, fluorescence, thermal and microwave satellite data for estimating large-scale crop yields. *Remote Sensing of Environment*, **199**, 333–349.
Hoerl A E, Kennard R. 1970. Ridge regression: Applications to nonorthogonal problems. *Technometrics*, **12**, 69–82.
Horie T, Yajima M, Nakagawa H. 1992. Yield forecasting. *Agricultural Systems*, **40**, 211–236.
Huang J, Sedano F, Huang Y, Ma H, Li X, Liang S, Tian L, Zhang X, Fan J, Wu W. 2016. Assimilating a synthetic Kalman filter leaf area index series into the WOFOST model to improve regional winter wheat yield estimation. *Agricultural and Forest Meteorology*, **216**, 188–202.
Ikram M, Zhang Q, Sroufe R, Shah S Z A. 2020. Towards a sustainable environment: The nexus between ISO 14001, renewable energy consumption, access to electricity, agriculture and CO_2 emissions in SAARC countries. *Sustainable Production and Consumption*, **22**, 218–230.
Ji Y, Liu R, Xiao Y, Cui Y, Chen Z, Zong X, Yang T. 2023. Faba bean above-ground biomass and bean yield estimation based on consumer-grade unmanned aerial vehicle RGB images and ensemble learning. *Precision Agriculture*, **24**, 1439–1460.
John K, Reddy P R. 2019. Correlation and path analysis for yield and yield attributes in groundnut (*Arachis hypogaea* L.). *Legume Research*, **42**, 518–522.
Kearns M, Ron D. 1999. Algorithmic stability and sanity-check bounds for leave-one-out cross-validation. *Neural Computation*, **11**, 1427–1453.
Khaki S, Wang L. 2019. Crop yield prediction using deep neural networks. *Frontiers in Plant Science*, **10**, 621.
Khanal S, Fulton J, Klopfenstein A, Douridas N, Shearer S. 2018. Integration of high resolution remotely sensed data and machine learning techniques for spatial prediction of soil properties and corn yield. *Computers and Electronics in Agriculture*, **153**, 213–225.
Killeen P, Kiringa I, Yeap T, Branco P. 2024. Corn grain yield prediction using UAV-based high spatiotemporal resolution imagery, machine learning, and spatial cross-validation. *Remote Sensing*, **16**, 683.
Knoben W J M, Freer J E, Woods R A. 2019. Technical note: Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores. *Hydrology and Earth System Sciences*, **23**, 4323–4331.
Kumar C, Mubvumba P, Huang Y, Dhillon J, Reddy K. 2023. Multi-stage corn yield prediction using high-resolution UAV multispectral data and machine learning models. *Agronomy*, **13**, 1277.
Lecerf R, Ceglar A, López-Lozano R, Van Der Velde M, Baruth B. 2019. Assessing the information in crop model and meteorological indicators to forecast crop yield over Europe. *Agricultural Systems*, **168**, 191–202.
Li G, Zhang J, Yang C, Song Y, Zheng C, Liu Z, Wang S, Tang S, Ding Y. 2014. Yield and yield components of hybrid rice as influenced by nitrogen fertilization at different eco-sites. *Journal of Plant Nutrition*, **37**, 244–258.
Li L, Liu Z, Du X J. 2021. Improvement of analytic hierarchy process based on grey correlation model and its engineering application. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems* (Part A: Civil Engineering), **7**, 04021007.
Liu S B, Jin X L, Bai Y, Wu W B, Cui N B, Cheng M H, Liu Y D, Meng L, Jia X, Nie C W, Yin D M. 2023. UAV multispectral images for accurate estimation of the maize LAI considering the effect of soil background. *International Journal of Applied Earth Observation and Geoinformation*, **121**, 103383.
Liu Y, Sun L, Liu B, Wu Y, Ma J, Zhang W, Wang B, Chen Z. 2023. Estimation of winter wheat yield using multiple temporal vegetation indices derived from UAV-based multispectral and hyperspectral imagery. *Remote Sensing*, **15**, 4800.
Ma Y, Ma L, Zhang Q, Huang C, Yi X, Chen X, Hou T, Lv X, Zhang Z. 2022. Cotton yield estimation based on vegetation indices and texture features derived from RGB image. *Frontiers in Plant Science*, **13**, 925986.
Maimaitijiang M, Sagan V, Sidike P, Hartling S, Esposito F, Fritschi F B. 2020. Soybean yield prediction from UAV using multimodal data fusion and deep learning. *Remote Sensing of Environment*, **237**, 111599.
Nada H, Siniša J, Anto M, Vladimir M, Dragana M. 2016. Correlation and path analysis of yield and yield components of confectionary sunflower. *Genetika*, **48**, 827–835.
Paudel D, de Wit A, Boogaard H, Marcos D, Osinga S, Athanasiadis I N. 2023. Interpretability of deep learning models for crop yield forecasting. *Computers and Electronics in Agriculture*, **206**, 107663.
Peng G, Ruiliang P, Biging G S, Larrieu M R. 2003. Estimation of forest leaf area index using vegetation indices derived from Hyperion hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, **41**, 1355–1362.
Peng X, Han W, Ao J, Wang Y. 2021. Assimilation of LAI derived from UAV multispectral data into the SAFY model to estimate maize yield. *Remote Sensing*, **13**, 1094.
Pokhrel A, Virk S, Snider J L, Vellidis G, Hand L C, Sintim H Y, Parkash V, Chalise D P, Lee J M, Byers C. 2023. Estimating yield-contributing physiological parameters of cotton using UAV-based imagery. *Frontiers in Plant Science*, **14**, 1248152.
Qader S H, Dash J, Atkinson P M. 2018. Forecasting wheat and barley crop production in arid and semi-arid regions using remotely sensed primary productivity and crop phenology: A case study in Iraq. *Science of the Total Environment*, **613**, 250–262.
Qi J, Chehbouni A, Huete A R, Kerr Y H, Sorooshian S. 1994. A modified soil adjusted vegetation index. *Remote Sensing of Environment*, **48**, 119–126.
Qiao Y, Wang X, Han Z, Tian M, Wang Q, Wu H, Liu F. 2022. Geodetector based identification of influencing factors on spatial distribution patterns of heavy metals in soil: A case in the upper reaches of the Yangtze River, China. *Applied Geochemistry*, **146**, 105459.
Richetti J, Lawes R A, Whan A, Gaydon D S, Thorburn P J. 2024. How well does APSIM NextGen simulate wheat yields across Australia using gridded input data? Validating continental-scale crop model simulations. *European Journal of Agronomy*, **158**, 127212.
Ramos A P M, Osco L P, Furuya D E G, Gonçalves W N, Santana D C, Teodoro L P R, Silva Junior C A D, Capristo-Silva G F, Li J, Baio F H R, Marcato Junior J, Teodoro P E, Pistori H. 2020. A random forest ranking approach to predict yield in maize with UAV-based vegetation spectral indices. *Computers and Electronics in Agriculture*, **178**, 105791.
Rondeaux G, Steven M, Baret F. 1996. Optimization of soil-adjusted vegetation indices. *Remote Sensing of Environment*, **55**, 95–107.
Rouse J W, Haas R H, Schell J A, Deering D W. 1974. Monitoring vegetation systems in the Great Plains with ERTS. *NASA Special Publication*, **351**, 309–317.
Samberg L H, Gerber J S, Ramankutty N, Herrero M, West P C. 2016. Subnational distribution of average farm size and smallholder contributions to global food production. *Environmental Research Letters*, **11**, 124010.
Sapkota S, Paudyal D R. 2023. Growth monitoring and yield estimation of maize plant using unmanned aerial vehicle (UAV) in a hilly region. *Sensors*, **23**, 5432.
Schuerger A C, Capelle G A, Di Benedetto J A, Mao C, Thai C N, Evans M D, Richards J T, Blank T A, Stryjewski E C. 2003. Comparison of two hyperspectral imaging and two laser-induced fluorescence instruments for the detection of zinc stress and chlorophyll concentration in bahia grass (*Paspalum notatum* Flugge.). *Remote Sensing of Environment*, **84**, 572–588.
Schut A G T, Traore P C S, Blaes X, de By R A. 2018. Assessing yield and fertilizer response in heterogeneous smallholder fields with UAVs and satellites. *Field Crops Research*, **221**, 98–107.
Sharma V, Honkavaara E, Hayden M, Kant S. 2024. UAV remote sensing phenotyping of wheat collection for response to water stress and yield prediction using machine learning. *Plant Stress*, **12**, 100464.

Song Y, Wang J, Ge Y, Xu C. 2020. An optimal parameters-based geographical detector model enhances geographic characteristics of explanatory variables for spatial heterogeneity analysis: Cases with different types of spatial data. *GIScience & Remote Sensing*, **57**, 593–610.

Su X, Wang J, Ding L, Lu J, Zhang J, Yao X, Cheng T, Zhu Y, Cao W, Tian Y. 2023. Grain yield prediction using multi-temporal UAV-based multispectral vegetation indices and endmember abundance in rice. *Field Crops Research*, **299**, 108992.

Sun J, Lai Z, Di L, Sun Z, Tao J, Shen Y. 2020. Multilevel deep learning network for county-level corn yield estimation in the U.S. corn belt. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **13**, 5048–5060.

Széles A V, Megyes A, Nagy J. 2012. Irrigation and nitrogen effects on the leaf chlorophyll content and grain yield of maize in different crop years. *Agricultural Water Management*, **107**, 133–144.

Taylor K E. 2001. Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research* (Atmospheres), **106**, 7183–7192.

Uno Y, Prasher S O, Lacroix R, Goel P K, Karimi Y, Viau A, Patel R M. 2005. Artificial neural networks to predict corn yield from Compact Airborne Spectrographic Imager data. *Computers and Electronics in Agriculture*, **47**, 149–161.

de Villiers C, Mashaba-Munghemezulu Z, Munghemezulu C, Chirima G J, Tesfamichael S G. 2024. Assessing maize yield spatiotemporal variability using unmanned aerial vehicles and machine learning. *Geomatics*, **4**, 213–236.

Voitik A, Kravchenko V, Pushka O, Kutkovetska T, Shchur T, Kocira S. 2023. Comparison of NDVI, NDRE, MSAVI and NDSI indices for early diagnosis of crop problems. *Agricultural Engineering*, **27**, 47–57.

Wan L, Cen H, Zhu J, Zhang J, Zhu Y, Sun D, Du X, Zhai L, Weng H, Li Y, Li X, Bao Y, Shou J, He Y. 2020. Grain yield prediction of rice using multi-temporal UAV-based RGB and multispectral images and model transfer — A case study of small farmlands in the South of China. *Agricultural and Forest Meteorology*, **291**, 108096.

Wanthanaporn U, Supit I, Chaowiwat W, Hutjes R W A. 2024. Skill of rice yields forecasting over Mainland Southeast Asia using the ECMWF SEAS5 ensemble prediction system and the WOFOST crop model. *Agricultural and Forest Meteorology*, **351**, 110001.

Yang B, Zhu W, Rezaei E E, Li J, Sun Z, Zhang J J R S. 2022. The optimal phenological phase of maize for yield prediction with high-frequency UAV remote sensing. *Remote Sensing*, **14**, 1559.

Yang G, Li Y, Yuan S, Zhou C, Xiang H, Zhao Z, Wei Q, Chen Q, Peng S, Xu L. 2024. Enhancing direct-seeded rice yield prediction using UAV-derived features acquired during the reproductive phase. *Precision Agriculture*, **25**, 1014–1037.

Yang L, Ji X, Li M, Yang P, Jiang W, Chen L, Yang C, Sun C, Li Y. 2024. A comprehensive framework for assessing the spatial drivers of flood disasters using an Optimal Parameter-based Geographical Detector–machine learning coupled model. *Geoscience Frontiers*, **15**, 101889.

Yu X, Zhang Q, Gao J, Wang Z, Borjigin Q, Hu S, Zhang B, Ma D. 2019. Planting density tolerance of high-yielding maize and the mechanisms underlying yield improvement with subsoiling and increased planting density. *Agronomy*, **9**, 370.

Yuan N, Gong Y, Fang S, Liu Y, Duan B, Yang K, Wu X, Zhu R. 2021. UAV remote sensing estimation of rice yield based on adaptive spectral endmembers and bilinear mixing model. *Remote Sensing*, **13**, 2190.

Yue J, Feng H, Yang G, Li Z. 2018. A comparison of regression techniques for estimation of above-ground winter wheat biomass using near-surface spectroscopy. *Remote Sensing*, **10**, 66.

Zhai L, Zhang L, Cui Y, Zhai L, Zheng M, Yao Y, Zhang J, Hou W, Wu L, Jia X. 2024. Combined application of organic fertilizer and chemical fertilizer alleviates the kernel position effect in summer maize by promoting post-silking nitrogen uptake and dry matter accumulation. *Journal of Integrative Agriculture*, **23**, 1179–1194.

Zhang B, Gu L, Dai M, Bao X, Sun Q, Qu X, Zhang M, Liu X, Fan C, Gu X H, Zhen W H. 2024. Estimation of grain filling rate and thousand-grain weight of winter wheat (*Triticum aestivum* L.) using UAV-based multispectral images. *European Journal of Agronomy*, **159**, 127258.

Zhang M, Kafy A A, Ren B, Zhang Y, Tan S, Li J. 2022. Application of the optimal parameter geographic detector model in the identification of influencing factors of ecological quality in Guangzhou, China. *Land*, **11**, 1303.

Zhang P, Lu B, Shang J, Wang X, Hou Z, Jin S, Yang Y, Zang H, Ge J, Zeng Z. 2024. Ensemble learning for oat yield prediction using multi-growth stage UAV images. *Remote Sensing*, **16**, 4575.

Zhou X, Zheng H B, Xu X Q, He J Y, Ge X K, Yao X, Cheng T, Zhu Y, Cao W X, Tian Y C. 2017. Predicting grain yield in rice using multi-temporal vegetation indices from UAV-based multispectral and digital imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, **130**, 246–255.

Executive Editor-in-Chief Shaokun Li
Managing Editor Ning Wang