

Article

Grain Yield Estimation of Rice Germplasm Resources Using Time-Series UAV Imagery and Dynamic Clustering Process

Qi Ke ^{1,†} , Di Wang ^{2,†}, Yan Zhao ¹, Caili Guo ¹, Xiaoxu Han ¹, Ankang Zhang ², Chongya Jiang ¹, Xia Yao ^{1,3} , Tao Cheng ¹ , Weixing Cao ¹, Yan Zhu ¹  and Hengbiao Zheng ^{1,3,*}

¹ National Engineering and Technology Center for Information Agriculture (NETCIA), MARA Key Laboratory of Crop System Analysis and Decision Making, MOE Engineering Research Center of Smart Agriculture, Jiangsu Key Laboratory for Information Agriculture, Institute of Smart Agriculture, Nanjing Agricultural University, Nanjing 211800, China; 2025101170@stu.njau.edu.cn (Q.K.); 2024101179@stu.njau.edu.cn (Y.Z.); clguo@njau.edu.cn (C.G.); 2022201010@stu.njau.edu.cn (X.H.); 2023003@njau.edu.cn (C.J.); yaoxia@njau.edu.cn (X.Y.); tcheng@njau.edu.cn (T.C.); caow@njau.edu.cn (W.C.); yanzhu@njau.edu.cn (Y.Z.)

² Huaiyin Institute of Agricultural Sciences of Xuhuai Region in Jiangsu, Huai'an 223001, China; wangdirice@163.com (D.W.); 13905239366@139.com (A.Z.)

³ Zhongshan Biological Breeding Laboratory, Nanjing 210014, China

* Correspondence: zhenghb@njau.edu.cn

† These authors contributed equally to this work.

Abstract

Traditional methods for measuring rice yield are often labor-intensive, time-consuming, and difficult to implement at scale. Conversely, remote sensing-based yield prediction models typically exhibit limited applicability across diverse genetic materials. In this study, we propose a high-precision yield prediction approach that integrates UAV-based time-series imagery with dynamic process clustering. Field experiments were conducted over two years involving 630 rice germplasm accessions in Rugao and Huaian, Jiangsu Province. UAV-mounted RGB and multispectral cameras were employed to acquire canopy imagery throughout the rice growth period. A range of features, including spectral reflectance, vegetation indices, canopy height (CH), and canopy volume (CV), were extracted from the UAV data. The K-Shape clustering algorithm was applied to dynamically group the temporal growth curves, enabling the construction of a cluster-based yield prediction model. Among the vegetation indices, the Enhanced Vegetation Index (EVI2) demonstrated the best performance ($R^2 = 0.73$, RMSE = 599.53 kg/hm²). Models based on temporal features of CH and CV showed satisfactory accuracy ($R^2 = 0.70$, RMSE = 640.96 kg/hm²). Notably, a dual-modal model combining vegetation indices with structural parameters significantly improved predictive performance ($R^2 = 0.80$, RMSE = 511.42 kg/hm²). This study demonstrates that multi-feature cluster analysis enhances the accuracy and robustness of yield prediction models across diverse genotypes. The proposed methodology provides valuable technical support for high-yield rice breeding initiatives.



Academic Editor: Jiehao Li

Received: 10 April 2026

Revised: 8 May 2026

Accepted: 10 May 2026

Published: 12 May 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: K-shape clustering; rice breeding; UAV remote sensing; time series; yield prediction

1. Introduction

Rice (*Oryza sativa* L.) is a staple food for more than half of the world's population, and accurate yield prediction is essential for ensuring food security and optimizing agricultural resource allocation. Traditional yield estimation methods, which rely on manual sampling of yield components (e.g., panicle number, grains per panicle, thousand-grain weight), are

labor-intensive, time-consuming, and prone to human error. Moreover, these measurements can only be conducted at maturity, precluding early-stage yield prediction [1]. These limitations highlight the urgent need for high-throughput, non-destructive, and timely yield estimation approaches.

In recent years, unmanned aerial vehicle (UAV) remote sensing has emerged as a transformative technology for agricultural monitoring, offering centimeter-level spatial resolution and flexible revisit frequency [2,3]. Equipped with multispectral, hyperspectral, or RGB sensors, UAVs can capture canopy spectral signatures, three-dimensional structural attributes, and thermal information throughout the crop growing season [4,5]. Compared with satellite remote sensing, which suffers from coarse spatial resolution (10 m to 1 km) and long revisit cycles (2–7 days), UAV-based platforms enable plot-scale monitoring with unprecedented detail [6].

Vegetation indices (VIs) derived from UAV multispectral imagery have been widely used for crop yield prediction. Commonly employed VIs include the Normalized Difference Vegetation Index (NDVI) [7], Green Normalized Difference Vegetation Index (GNDVI) [8], Red Edge Normalized Difference Vegetation Index (NDRE) [9], Chlorophyll Indices (CI_{green} , $CI_{\text{red edge}}$) [10], and Enhanced Vegetation Index 2 (EVI2) [11]. These indices are sensitive to canopy chlorophyll content, green biomass, and photosynthetic activity, making them effective predictors of final yield. For example, Yang et al. [1] achieved yield prediction with $R^2 = 0.78$ across 230 rice accessions using RGB imagery. However, studies have shown that relying solely on spectral information may be insufficient for accurate yield estimation [12]. Canopy structural parameters, such as plant height and canopy volume, provide complementary information about biomass accumulation and plant architecture. With the advent of Structure from Motion (SfM) and Multi-View Stereo (MVS) algorithms, UAV-based RGB imagery can generate dense point clouds and digital surface models (DSMs), enabling the extraction of canopy height (CH) with high accuracy.

The temporal dimension of crop growth carries critical information about yield formation. Multiple studies have confirmed that yield prediction accuracy improves when multi-temporal remote sensing data are used instead of single-stage observations [13,14]. Time-series analysis allows the capture of growth dynamics, including the rate of biomass accumulation, peak greenness timing, and senescence patterns, all of which are indicative of final yield. Machine learning approaches, including random forest (RF), support vector machine (SVM), extreme gradient boosting (XGBoost), and deep learning architectures such as convolutional neural networks (CNNs) and long short-term memory (LSTM) networks, have been widely applied to UAV-based yield prediction. Studies have reported that combining multi-temporal features with appropriate machine learning models can achieve R^2 values exceeding 0.80 for rice yield prediction [15,16].

For time-series data, clustering methods play a crucial role in grouping samples with similar temporal patterns before model construction. Traditional clustering methods, such as k-means, rely on Euclidean distance to measure similarity between sequences. However, Euclidean distance is sensitive to amplitude differences and phase shifts, meaning that two sequences with similar shapes but different magnitudes or temporal offsets may be incorrectly classified as dissimilar. This limitation makes k-means less suitable for time-series data where shape patterns are more informative than absolute values.

Dynamic time warping (DTW) is another widely used approach for time-series similarity measurement. DTW aligns sequences by allowing non-linear matching of time indices, which effectively handles temporal shifts. However, DTW has two major drawbacks: first, its computational complexity is $O(N^2)$, making it prohibitive for large-scale datasets; second, DTW does not provide a straightforward method for computing cluster centroids, which limits its use in centroid-based clustering algorithms.

In contrast, the K-Shape algorithm, proposed by Paparrizos and Gravano [17], is specifically designed for time-series clustering. K-Shape uses normalized cross-correlation (NCC) as the similarity measure, which is scale-invariant and translation-invariant. This means that K-Shape focuses on the shape of the sequence—such as rising, falling, peaking, or plateauing—rather than absolute values. Compared with k-means, K-Shape is more robust to amplitude differences and phase shifts. Compared with DTW, K-Shape has a lower computational complexity ($O(N \log N)$) and provides an efficient method for centroid computation based on maximizing the average NCC within each cluster. These properties make K-Shape particularly attractive for clustering crop growth time series, where different accessions may have similar shape patterns despite differences in absolute height or growth timing.

Despite the growing body of research on UAV-based yield prediction, several critical gaps remain. First, most studies have been conducted on uniform canopies (single or few varieties), leaving a gap in understanding how to handle diverse genetic materials. Second, many studies rely exclusively on spectral vegetation indices, with limited integration of structural parameters across multiple growth stages. Third, while K-Shape has been successfully applied in other domains [18], its application to crop yield prediction remains underexplored. Fourth, most models are validated within the same year and location [19,20], raising questions about their generalizability to new environments and germplasm pools.

To address these gaps, this study proposes a high-precision yield prediction framework that integrates UAV-based time-series imagery with K-Shape clustering. The specific objectives are: (1) to evaluate the predictive power of single-modal versus dual-modal (spectral + structural) time-series features; (2) to apply K-Shape clustering to group accessions by temporal growth curve similarity; (3) to quantify the contributions of temporal information, clustering strategy, and multi-modal fusion through ablation experiments; and (4) to validate the proposed model using an independent cross-year and cross-location dataset. Through these innovations, this study aims to provide a robust and interpretable yield prediction tool for rice breeding programs.

2. Materials and Methods

2.1. Experimental Design

Experiment 1: This experiment was conducted from June 2022 to November 2022 at the Rugao Base of the National Engineering and Technology Center for Information Agriculture, located in Rugao City, Jiangsu Province ($32^{\circ}30' \text{ N}$, $120^{\circ}20' \text{ E}$) (Figure 1). The average annual temperature in the experimental area is 14.6°C , with an average of 121.3 rainy days, an average annual precipitation of 1055.5 mm, and an average annual frost-free period of 215.6 days. The soil type is loam. The experimental materials consisted of 230 breeding accessions from the Taihu Lake Basin and 10 conventional high-yield varieties, totaling 240 plots. Sowing was carried out on 20 May, and each plot had an area of $1.8 \text{ m} \times 2 \text{ m} = 3.6 \text{ m}^2$.

Experiment 2: This experiment was conducted from June 2024 to November 2024 at the Rice and Wheat Scientific Research and Breeding Base in Qingjiangpu District, Huai'an City, Jiangsu Province ($119^{\circ}01' \text{ E}$, $33^{\circ}35' \text{ N}$) (Figure 1). Qingjiangpu District is situated in the north-central part of Jiangsu Province, with an average annual temperature of 14°C , an average annual frost-free period of 240 days, and an average annual rainfall of 940 mm. The experimental materials were new elite lines from the Academy of Agricultural Sciences, with a total of 390 plots. Sowing was performed on 21 May, and each plot measured $2.05 \text{ m} \times 5 \text{ m} = 10.25 \text{ m}^2$.

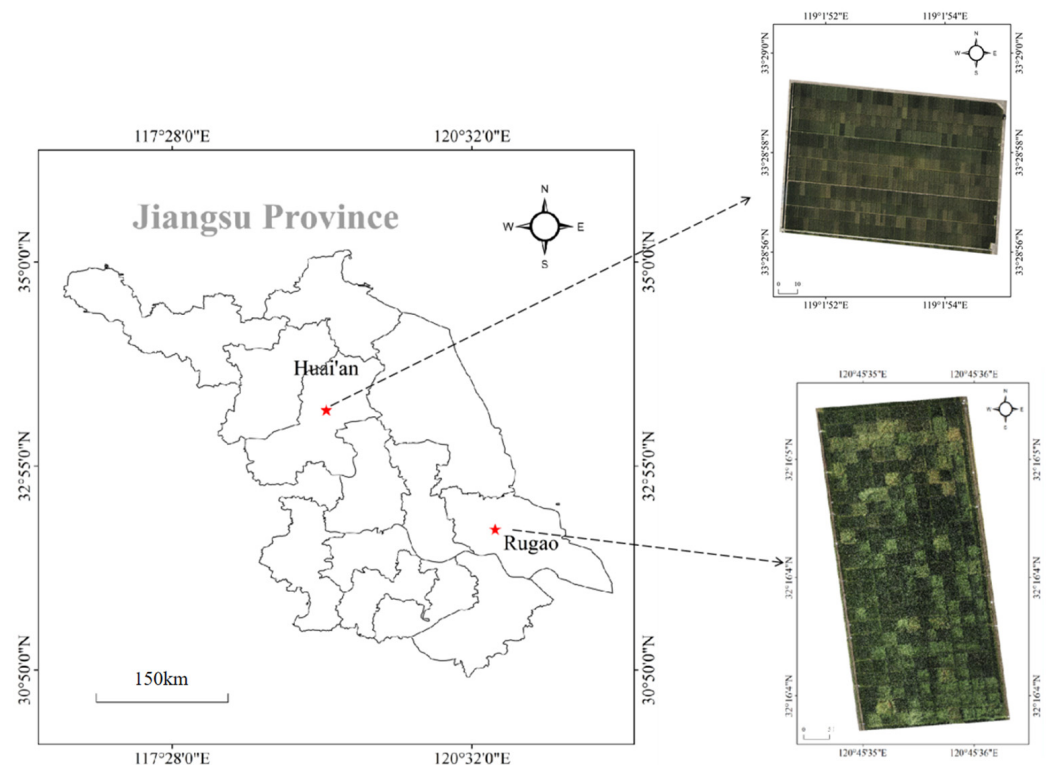


Figure 1. Experimental site of this study.

The genetic materials differ substantially: the Rugao experiment mainly included traditional Taihu Lake Basin germplasm accessions (230 accessions plus 10 conventional varieties), while the Huai'an experiment comprised newly bred elite lines (390 accessions). Plot sizes also varied (3.6 m^2 in Rugao vs. 10.25 m^2 in Huai'an), and climatic conditions (temperature, precipitation) as well as agronomic management schedules (e.g., fertilization timing) were not strictly identical between the two locations. Despite these differences, both experiments adopted consistent UAV flight parameters (altitude 30 m, 80% overlap) and comparable sensor configurations, ensuring data acquisition uniformity.

2.2. Data Collection

2.2.1. UAV Data

Multi-source remote sensing images of the rice canopy were acquired using a Phantom 4 Pro quadrotor UAV and an M300RTK UAV manufactured by DJI Technology Co., Ltd., Shenzhen, China, each equipped with RGB and multispectral cameras. The specific sensor parameters are summarized in Table 1. The UAVs followed pre-designed flight paths, with imaging conducted between 11:00 and 13:00 under clear, cloudless sky conditions. Flight altitude was set to 30 m, with a forward overlap of 80% and a side overlap of 80%. The UAV gimbal maintained the camera in a vertically downward orientation during image capture, and camera parameters were kept constant throughout the aerial survey. UAV images were collected at intervals of 5 to 10 days, resulting in a total of nine data acquisition phases (S1 to S8) spanning the entire rice growing season. These phases corresponded to key developmental stages, including tillering, jointing, booting, heading, and maturity.

Table 1. UAV sensor parameters.

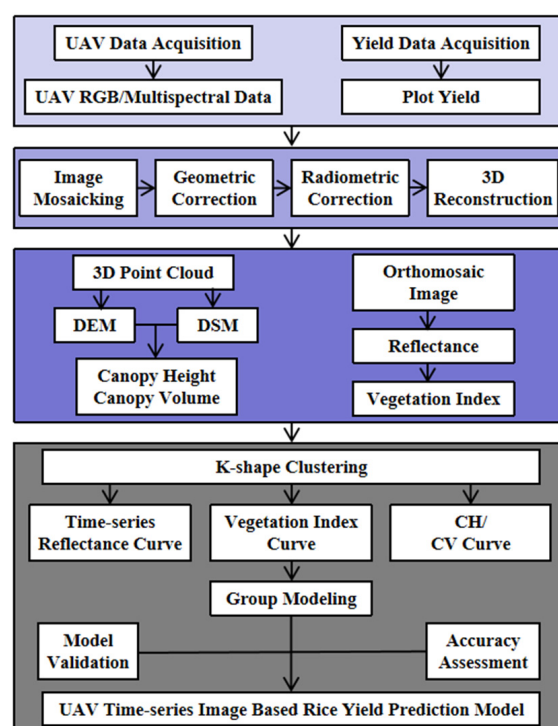
Sensor	DJI Zenmuse P1	DJI Phantom 4 Pro RGB	RedEdge	Changguang YuChen MS600
Number of Bands	3 R\G\B	3 R\G\B	5 R\G\B\RE\NIR	6 R\G\B\RE1\RE2\NIR
Field of View	63.5°	84°	49.6° × 38.3°	49.5° × 38.1°
Image Size	8192 × 5460	5472 × 3078	1456 × 1088	3648 × 2736
Ground Resolution (cm)	0.38	0.82	2.25	0.9
Exposure Mode	Auto Exposure	Auto Exposure	Fixed Exposure	Fixed Exposure

2.2.2. Yield Data

Rice yield was estimated through a manual measurement method. All rice plants within the designated quadrat were harvested, threshed, and air-dried prior to yield determination. The yield was expressed in units of kilograms per hectare (kg/hm²).

2.3. Image Processing

The technical workflow of this study comprised four key steps (Figure 2). First, UAV-based RGB and multispectral images, along with field yield data, were collected. Second, the raw images underwent preprocessing procedures, including mosaicking, geometric correction, radiometric correction, and 3D reconstruction. Third, canopy height (CH) and canopy volume (CV) were derived from the point cloud data, while reflectance and vegetation indices were extracted from multispectral orthomosaics. Finally, K-Shape clustering was applied to the time-series features, and yield prediction models were developed and validated for each resulting cluster.

**Figure 2.** Research roadmap.

2.3.1. RGB Image Processing

RGB images were processed using Agisoft Metashape software (version 1.8, Agisoft LLC., St. Petersburg, Russia). The processing workflow consisted of the following steps:

(1) Image alignment and feature matching: Based on the Structure from Motion (SfM) algorithm, feature points between images were automatically detected and matched using high-accuracy mode.

(2) Ground control point optimization: Bundle block adjustment was performed using field-collected ground control points (GCPs). GCP coordinates were measured using real-time kinematic differential GPS (RTK-GPS) to ensure high positioning accuracy.

(3) Dense point cloud generation and 3D reconstruction: A high-density three-dimensional point cloud was generated using the Multi-View Stereo (MVS) algorithm, from which digital surface models (DSMs) and digital elevation models (DEMs) were constructed.

(4) Orthophoto generation: Orthorectification of the original images was performed based on the DSM to produce georeferenced RGB orthophotos.

(5) Canopy height model generation: The canopy height model (CHM) was obtained by subtracting the DEM from the DSM on a per-pixel basis, which was then used for subsequent canopy height and canopy volume extraction.

2.3.2. Multispectral Data Processing

Two multispectral cameras (RedEdge (MicaSense, Inc., Seattle, WA, USA) and MS600 (Yusense, Qingdao, China)) were used in this study, with data from different sites processed using their respective dedicated software.

RedEdge Multispectral Camera Processing (Rugao Site)

Images acquired with the RedEdge camera were processed using Pix4Dmapper software (version 4.5.6, Pix4D SA, Lausanne, Switzerland) following the procedures below:

(1) Image mosaicking and aerial triangulation: After importing all raw images, feature matching points between images were automatically extracted using the SIFT algorithm. Aerial triangulation was performed using bundle block adjustment, with keypoint extraction set to high-accuracy mode and matching set to optimal mode. The root mean square reprojection error was controlled within 0.5 pixels.

(2) Radiometric correction and reflectance calculation: Radiometric correction was performed using a standard reflectance panel (50% reflectance) captured synchronously with the flight, according to the following equation:

$$R_{\text{target}} = \frac{DN_{\text{target}} - DN_{\text{dark}}}{DN_{\text{panel}} - DN_{\text{dark}}} \times DN_{\text{panel}} \quad (1)$$

where R_{target} is the target reflectance, DN_{target} is the target digital number, DN_{panel} is the panel DN value, R_{panel} is the panel reflectance, and DN_{dark} is the dark current noise. Each band was corrected independently, and the corrected reflectance images were exported as GeoTIFF files.

(3) Geometric registration: Multispectral images were geometrically registered to the RGB orthophotos as references. A second-order polynomial transformation model and nearest-neighbor resampling method were applied to resample the multispectral images to the same spatial resolution as the RGB images (1 cm/pixel), with the root mean square error of registration controlled within 1 pixel.

(4) Band fusion: The registered multispectral bands were fused with the RGB images to generate eight-band composite images containing RGB and five multispectral bands (blue, green, red, red edge, near-infrared).

MS600 Multispectral Camera Processing (Huai'an Site)

Images acquired with the MS600 camera were processed using Ysense Map software (version 2.0, Changguang Yuchen Technology Co., Ltd., Qingdao, China). The main processing steps included band registration, image mosaicking, and reflectance calibration, with processing parameters kept at software default settings. Reflectance calibration was performed using the same standard reflectance panel and the same calibration equation as described above.

Region of Interest (ROI) Extraction

Plot boundaries for each experimental plot were manually delineated as regions of interest (ROIs) on each orthophoto using ArcGIS Pro software (version 2.9, ESRI, Redlands, CA, USA). To minimize edge effects between adjacent plots, ROI boundaries were contracted inward by 2 pixels (approximately 2–4 cm). Spectral reflectance and canopy height values for all pixels within each ROI were extracted for subsequent vegetation index calculation and canopy parameter extraction.

2.4. Vegetation Index Construction

To comprehensively monitor rice growth status at different phenological stages and identify the spectral features most relevant to yield, six representative vegetation indices were selected in this study. The selection criteria were as follows: (1) the indices have been widely reported in the literature to be strongly correlated with crop yield or biomass; (2) they cover different spectral response mechanisms, including sensitivity to chlorophyll, canopy structure, and the red-edge region; and (3) they are suitable for characterizing rice canopy structural features. The calculation formula, physiological significance, and selection rationale for each index are detailed in Table 2.

Table 2. Vegetation indices used in this study.

Vegetation Index	Formula	Physiological/Canopy Significance	Applicable Characteristics	Reference
Normalized difference vegetation index	$NDVI = (NIR - R) / (NIR + R)$	Reflects green vegetation coverage and photosynthetic activity	Early to mid-growth stages	[7]
Green normalized difference vegetation index	$GNDVI = (NIR - G) / (NIR + G)$	More sensitive to chlorophyll content changes than NDVI	Mid to late growth stages	[8]
Red edge normalized difference vegetation index	$NDRE = (NIR - RE) / (NIR + RE)$	Indicates leaf chlorophyll content and nitrogen nutritional status	Effective under high coverage	[9]
Red edge chlorophyll index	$CI_{RE} = (NIR/RE) - 1$	Linearly correlated with leaf chlorophyll content	Sensitive to canopy structure	[10]
Green chlorophyll index	$CI_G = (NIR/G) - 1$	Sensitive to chlorophyll content and nitrogen stress	Suitable for nutrient diagnosis	[21]
Enhanced vegetation index 2	$EVI2 = 2.5 \times (NIR - R) / (NIR + 2.4 \times R + 1)$	Reflects vegetation vitality under high biomass conditions	Saturation-resistant for mid-late stages	[11]

2.5. Extraction of Canopy Height and Canopy Volume

Prior to rice transplanting, a DEM of the terrain was acquired to serve as a baseline for subsequent canopy height assessments. For each aerial survey, a DSM was generated, and the DEM was subtracted from the DSM to produce a Rice Canopy Height Model (CHM). The rice canopy region was accurately delineated by defining ROIs on the CHM, and the average canopy height within each ROI was calculated as the canopy height for the corresponding plot.

Following the extraction of the CHM, the Canopy Volume Model (CVM) was constructed utilizing ground sampling distance (GSD) information corresponding to each pixel. Similar to the calculation of canopy height, the canopy volume was obtained using the ROI tool. The formula for calculating canopy volume (CV) is provided in Equation (1):

$$CV = \sum CH_i \times GSD2 \quad (2)$$

where CH_i denotes the canopy height of the i -th pixel.

2.6. K-Shape Clustering

The primary clustering method employed in this study is K-Shape clustering, which is characterized by high accuracy and strong extensibility. As a clustering technique specifically tailored for time-series data, K-Shape groups sequences based on shape similarity rather than relying on absolute values or phase alignment [17]. However, K-Shape is relatively sensitive to noise and imposes strict requirements on input data quality, necessitating data preprocessing and uniform sequence lengths [22]. This algorithm is built upon Normalized Cross-Correlation (NCC) and utilizes Shape-Based Distance (SBD) to evaluate similarity, assessing shape correspondence by maximizing the cross-correlation value of sequences after sliding window alignment. To mitigate the effects of dimensional interference, time-series data must be standardized to have a mean of zero and a standard deviation of one prior to clustering. The algorithm comprises four main steps: (1) random initialization of cluster centers; (2) assignment of data points based on SBD; (3) updating of cluster centers through maximization of cross-correlation; and (4) iterative optimization until convergence is achieved. Its distinctive advantage lies in its robustness to phase shifts, enabling effective identification of global shape patterns such as trends and peaks while maintaining computational efficiency for medium-scale datasets.

2.6.1. Construction of Clustering Input Features

Three types of temporal features were independently used as inputs for K-Shape clustering: (a) EVI2 time series—eight EVI2 values extracted from each plot across eight growth stages, forming an 8-dimensional temporal vector; (b) canopy height (CH) time series—eight CH values extracted from each plot across eight growth stages, forming an 8-dimensional temporal vector; (c) canopy volume (CV) time series—eight CV values calculated from each plot across eight growth stages using Equation (1), forming an 8-dimensional temporal vector. All time-series vectors were standardized prior to clustering to eliminate amplitude differences while preserving shape characteristics.

2.6.2. Determination of the Optimal Number of Clusters

The optimal number of clusters (k) was determined using a combination of the Elbow method (based on within-cluster sum of squared distances) and the Silhouette coefficient. Values of k ranging from 2 to 8 were evaluated. The optimal k was selected as the value that maximized the Silhouette coefficient while ensuring that each cluster contained at least 25 samples for subsequent model calibration.

2.6.3. Assignment of Cluster Labels for the Validation Set

For the independent validation set, re-clustering and centroid re-computation were performed on the validation data. The assignment procedure was as follows: (1) extract the same temporal features for each validation plot; (2) standardize the time-series sequences using the mean and standard deviation calculated from the training set (to avoid data leakage); (3) compute the Shape-Based Distance (SBD) between each validation sequence

and each training-derived cluster centroid; and (4) assign the plot to the cluster with the minimum SBD.

2.6.4. Post-Clustering Modeling Workflow

After clustering, a separate yield prediction model was trained independently for each cluster. The specific workflow was as follows: (1) within-cluster data partitioning—for each cluster, samples assigned to that cluster were extracted, with no sample sharing across different clusters; (2) model selection—Random Forest was adopted as the base model for each cluster; (3) model calibration—for each cluster, a Random Forest model was calibrated using the samples within that cluster, with five-fold cross-validation applied to optimize hyperparameters and prevent overfitting; and (4) prediction for new samples—for a new sample in the validation set, its cluster membership was first determined according to the rule described in 2.6.3, and then the corresponding cluster-specific Random Forest model was called to generate the yield prediction.

2.7. Model Accuracy Evaluation

Two accuracy evaluation indicators were selected in this study, namely the Coefficient of Determination (R^2) and Root Mean Square Error (RMSE).

2.8. AI Tool Usage in Materials and Methods

Generative artificial intelligence tools (ChatGPT 4.0, OpenAI, San Francisco, CA, USA; Doubao-Seed-2.0 Pro model (ByteDance, Beijing, China)) were used for language polishing and grammatical correction of the Materials and Methods section to improve the clarity and readability of the manuscript. All experimental designs, data collection, and analytical methods described in this section were designed, conducted, and verified by the authors, and the authors take full responsibility for the accuracy and integrity of all content.

3. Results

3.1. Rice Yield Variability

Figure 3 illustrates the distribution of rice yield data for the breeding accessions at two ecological sites, Huai'an and Rugao. The overall yield distribution approximates a normal distribution, exhibiting considerable variability. At the Huai'an site, a total of 390 samples were collected, with yields ranging from 2400.5 to 11,118.9 kg/hm², an average yield of 6849.14 kg/hm², and a coefficient of variation (CV) of 17.77%. At the Rugao site, 240 samples were measured, with yields ranging from 2450 to 15,000 kg/hm², an average yield of 8680 kg/hm², and a CV of 25%.

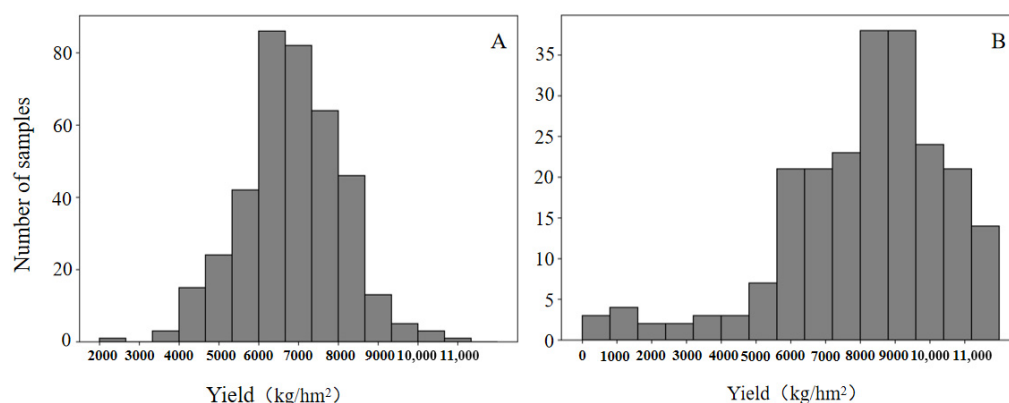


Figure 3. Histogram of the actual measured yield of Huai'an (A) and Rugao (B).

3.2. Canopy Height Extraction

Using the rice breeding experiment at Huai'an as an example, Figure 4 presents a comparison between the measured plant height and the UAV-derived plant height. The canopy height obtained through UAV remote sensing exhibited a strong linear correlation with the manually measured plant height ($R^2 = 0.96$), with all 390 sampling points densely clustered near the 1:1 reference line. In the subsequent modeling process, the canopy height parameters derived from UAV data were used exclusively as input variables, ensuring high-precision data for the yield prediction model.

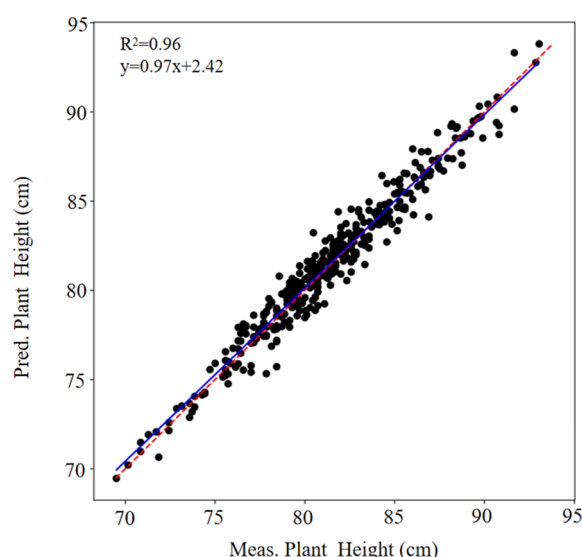


Figure 4. Comparison of rice canopy height derived from UAV RGB images with measured values.

3.3. Temporal Dynamic Variation of Image Features

Figure 5 displays the temporal variation curves of EVI2, canopy height (CH), and canopy volume (CV) within the breeding plots. The results indicate that EVI2 initially increases and subsequently decreases over the time series. In contrast, CH and CV exhibit more stable variation patterns, characterized by a monotonic increase followed by a plateau, reaching saturation.

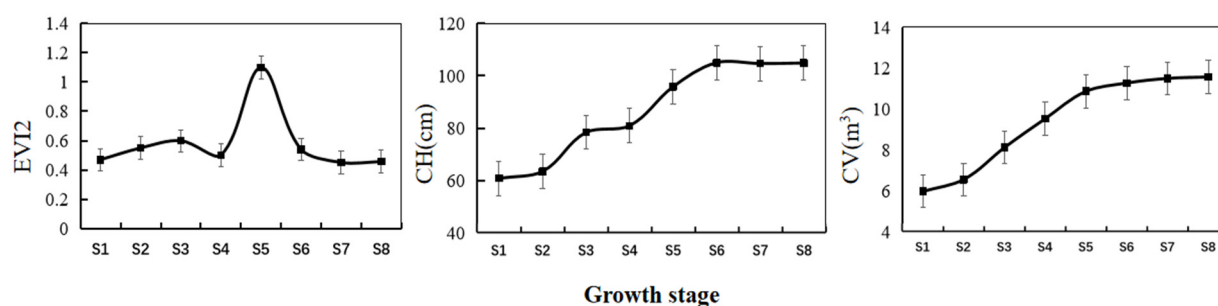


Figure 5. Time-series variation curves of reflectance, vegetation indices and canopy height in breeding plots.

3.4. Rice Yield Prediction Based on Single Image Features

Table 3 presents the p -values from correlation tests between vegetation indices, canopy height, canopy volume, and rice yield at various growth stages. The analysis reveals that, at the single growth stage level, most remote sensing parameters (e.g., NDVI, NDRE, and Cirededge) exhibited statistically significant correlations with yield from S5 to S9 ($p < 0.01$), with EVI2 at the heading stage (S5) showing the most significant correlation ($p = 0.00004$). In

the analysis combining all growth stages, EVI2 showed a statistically significant correlation with yield ($p = 0.03$), indicating that integrating multi-temporal information can enhance yield prediction performance.

Table 3. Significance levels (p -values) of correlation tests between yield and vegetation indices, canopy height (CH), and canopy volume (CV) at each growth stage and across all stages combined.

Growth Stage	NDVI	GNDVI	NDRE	CI _{green}	CI _{red edge}	EVI2	CH	CV
S1	0.003	0.001	0.0008	0.001	0.0007	0.0008	0.002	0.007
S2	0.003	0.001	0.0008	0.0009	0.0009	0.009	0.003	0.002
S3	0.001	0.0002	0.0006	0.0003	0.0009	0.006	0.009	0.008
S4	0.002	0.002	0.002	0.003	0.003	0.0001	0.01	0.01
S5	0.000	0.002	0.003	0.003	0.003	0.00004	0.002	0.001
S6	0.000	0.0009	0.001	0.0015	0.000	0.000	0.001	0.0009
S7	0.000	0.0001	0.000	0.000	0.000	0.0004	0.003	0.003
S8	0.001	0.000	0.000	0.000	0.0003	0.0005	0.003	0.0006
S9	0.000	0.000	0.0003	0.0004	0.0003	0.0005	0.004	0.0007
All	0.008	0.008	0.006	0.001	0.01	0.03	0.008	0.006

Figure 6 presents the results of rice yield prediction using the random forest algorithm based on remote sensing features derived from multiple growth stages. The model's prediction performance was notably limited, with R^2 values generally below 0.4 and root mean square error (RMSE) exceeding 1000 kg/hm².

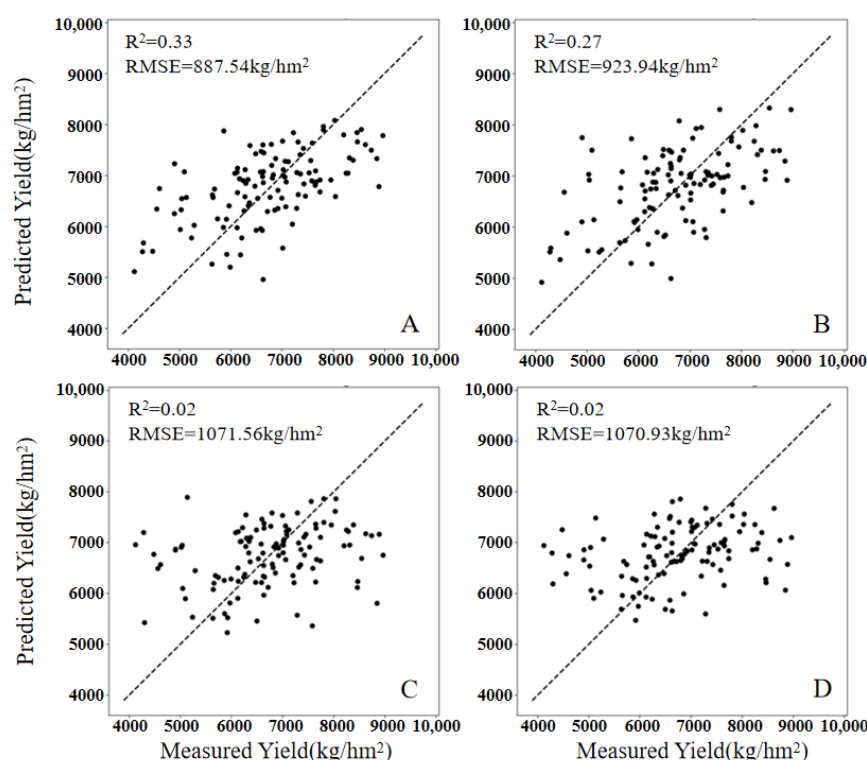


Figure 6. Rice yield estimation based on multi-temporal remote sensing features ((A) R840, (B) EVI2, (C) CH, (D) CV).

3.5. Rice Yield Prediction Based on Dynamic Process Clustering

3.5.1. Clustering Analysis Based on Temporal Vegetation Indices

Figure 7 illustrates the performance of EVI2 and NDRE in yield prediction with varying numbers of clusters. EVI2 achieves optimal results when the number of clusters is set to 8, with an R^2 approaching 0.72 and RMSE reduced to approximately 600 kg/hm².

Conversely, NDRE demonstrates relatively high R^2 values and lower RMSE when the number of clusters is set to 5.

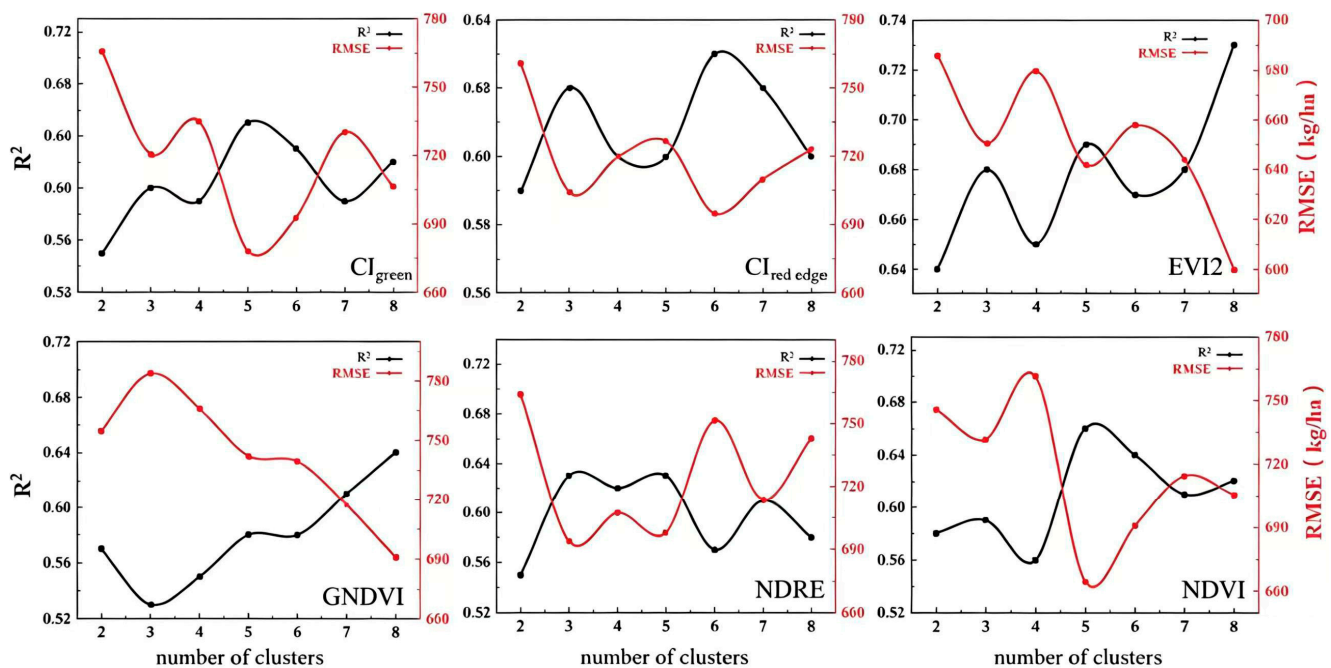


Figure 7. Trends in multi-period rice yield estimation results based on time-series vegetation index curves clustering.

As illustrated by the clustering optimization trends in Figure 7, the predictive performance of each vegetation index exhibited notable fluctuations with varying numbers of clusters. Notably, EVI2 demonstrated the most stable and superior predictive capability when dynamically partitioned into eight specific clusters.

3.5.2. Clustering Analysis Based on Temporal CH and CV

Figure 8 presents the clustering analysis based on time-series CH and CV. As the number of clusters varies, the R^2 and RMSE values for yield prediction based on CH and CV exhibit evident fluctuation trends. For the CH curve clustering, the optimal performance is achieved when the number of clusters is set to eight, with R^2 reaching approximately 0.72 and RMSE decreasing to around 600 kg/hm². In the case of CV curve clustering, the optimal results are attained at seven clusters, where R^2 and RMSE similarly demonstrate the best combination.

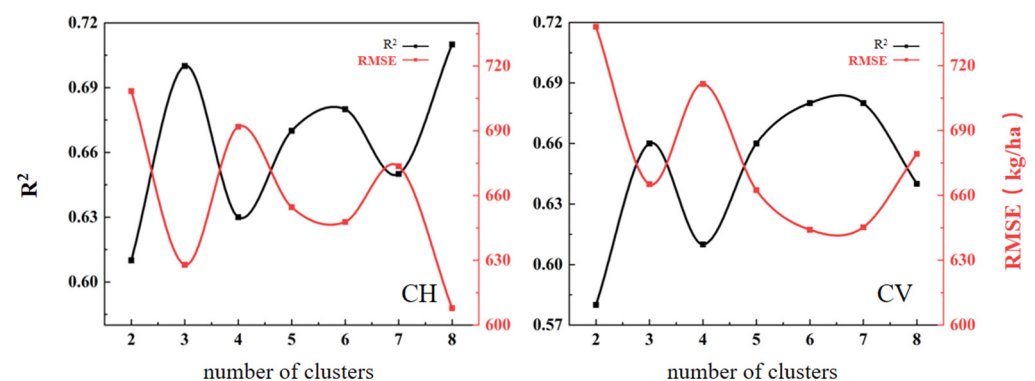


Figure 8. Trends in multi-period rice yield estimation results based on time-series CH and CV curves clustering.

3.6. Rice Yield Prediction Based on Dual-Modal Temporal Data

Table 4 demonstrates the significant advantages of multi-modal data fusion for rice yield prediction. The combination of EVI2 and CH yields the optimal performance, with a modeling R^2 of 0.80 and an RMSE of only 511.42 kg/hm²; in the independent validation, the R^2 is 0.79 and the RMSE is 950.56 kg/hm². This represents a substantial enhancement in prediction accuracy compared to using single modalities, with the EVI2 modality achieving $R^2 = 0.73$ and RMSE = 599.53 kg/hm², and the CH modality achieving $R^2 = 0.70$ and RMSE = 640.96 kg/hm².

Table 4. Comparison of rice yield prediction accuracy by fusing different vegetation indices and CH.

Fused Data Type	Coefficient of Determination (R^2)	Root Mean Square Error RMSE (kg/hm ²)	Independent Validation (R^2)	Independent Validation RMSE (kg/hm ²)
EVI2 + CH	0.80	511.42	0.79	950.56
NDVI + CH	0.73	598.71	0.70	1056.83
NDRE + CH	0.76	564.82	0.74	1018.72
GNDVI + CH	0.72	601.53	0.70	1062.45
CI _{green} + CH	0.73	590.17	0.71	1045.62
CI _{red edge} + CH	0.76	568.35	0.69	1078.31
EVI2	0.73	599.53	0.76	1067.67
CH	0.70	640.96	0.73	1122.39

For other vegetation indices (NDVI, NDRE, GNDVI, CI_{green}, CI_{red edge}) combined with CH, both the modeling and validation results show higher R^2 values and significantly lower RMSE compared to individual modalities. Notably, the NDRE + CH combination achieves an R^2 of 0.76 and an RMSE of 564.82 kg/hm², while the CI_{red edge} + CH combination also performs prominently with an R^2 of 0.76 and an RMSE of 568.35 kg/hm².

3.7. Validation of the Rice Yield Prediction Model

Figure 9 presents the results of independent validation using data collected from the Rugao experimental site in 2022. The clustering models based on EVI2, CH, and CV all demonstrate strong predictive performance, with R^2 values ranging from 0.68 to 0.76, significantly surpassing the accuracy of traditional multi-period regression models. Among these, the EVI2-based model achieved the highest validation accuracy, with an R^2 of 0.76 and an RMSE of 1067.67 kg/hm². These findings indicate that the approach utilizing clustering of time-series curve features effectively predicts rice yield, particularly for rice breeding accessions.

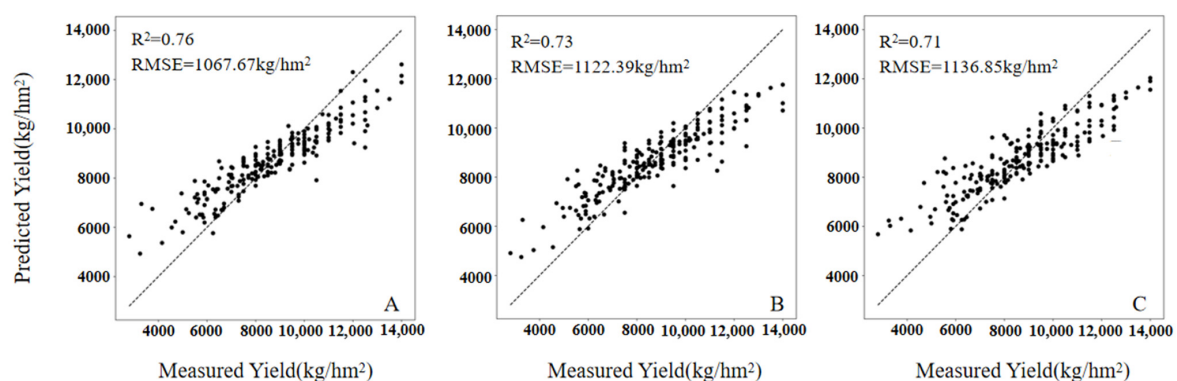


Figure 9. Validation results of rice yield prediction model based on time-series clustering. ((A) EVI2, (B) CH, (C) CV).

Similarly, the independent validation set, constructed from multi-source remote sensing data collected at the Rugao experimental site in 2022, was employed to evaluate the performance of the dual-modal rice yield prediction model (Figure 10). The results demonstrate a significant enhancement in predictive accuracy when combining EVI2 with structural parameters, with R^2 reaching 0.79 and RMSE decreasing to 950.56 kg/hm². These findings corroborate the results discussed above and further confirm the effectiveness of the proposed method in predicting rice yield for breeding accessions.

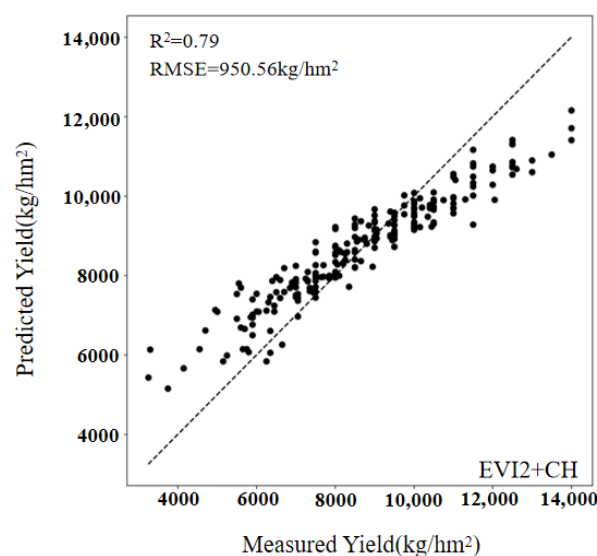


Figure 10. The validation results of the yield prediction model based on multi-modal data.

3.8. Comparison with Baseline Models and Ablation Study

The superiority of the proposed clustering-based method is evident when compared against four mainstream baseline models (PLSR, SVR, RF, and XGBoost). As shown in Table 5, XGBoost achieved the highest R^2 of 0.54 among the baseline models, followed by RF ($R^2 = 0.27$). PLSR and SVR performed poorly, with R^2 values of 0.03 and 0.00, respectively. In contrast, the proposed method achieved an R^2 of 0.73 and an RMSE of 599.53 kg/hm², substantially outperforming all baseline models.

Table 5. Performance comparison of the proposed method against mainstream baseline models.

Model	(R^2)	RMSE (kg/hm ²)
PLSR	0.03	1903.89
SVR	0.00	3654.44
RF	0.27	1485.04
XGBoost	0.54	923.94
Proposed method	0.73	599.53

Note: The proposed method refers to K-Shape clustering with within-cluster Random Forest modeling.

The ablation study (Table 6) further elucidates the sources of performance improvement. Among the three components examined, the clustering strategy contributed the largest gain ($\Delta R^2 = 0.560$), followed by temporal information ($\Delta R^2 = 0.481$) and multi-modal fusion ($\Delta R^2 = 0.07$ – 0.10). The full model achieved an R^2 of 0.80 and an RMSE of 511.42 kg/hm² (Table 6), outperforming all ablation variants.

Table 6. Ablation study results quantifying the contribution of each component.

Experiment	(R ²)	RMSE (kg/hm ²)	Performance Drop (ΔR^2)
Ablation: single-stage	0.32	1000.3	−0.481
Ablation: no clustering	0.24	1637.1	−0.560
Ablation: EVI2 only	0.73	599.53	−0.07
Ablation: CH only	0.70	640.96	−0.10
Full model	0.80	511.42	—

Note: The performance drop (ΔR^2) is calculated relative to the full model (0.80).

4. Discussion

4.1. Limitations of Single-Modal Features in Yield Prediction of Breeding Materials

Single-modal features are inherently limited in their capacity to capture the complex, dynamic processes underlying rice yield formation. Single-band reflectance is susceptible to spectral similarity, whereby accessions with differing yield potentials exhibit comparable reflectance values. For instance, using only the green band for yield prediction yielded an R^2 of only 0.58 and an RMSE of 742.85 kg/hm², indicating poor discriminative ability [13]. Although vegetation indices that combine multiple spectral bands improve sensitivity to physiological parameters, they remain vulnerable to interference from soil background, atmospheric conditions, and other external factors. Moreover, different indices respond differently to crop status; NDVI tends to saturate under dense canopies, whereas EVI2 may lose informational content during late growth stages. Consequently, models based on a single vegetation index typically achieve R^2 values between 0.68 and 0.73, which are insufficient for the high-precision screening required in breeding programs [14].

Canopy structural parameters, such as CH and CV, reflect the three-dimensional architecture of rice, but their relationship with yield is nonlinear and influenced by variety and planting density. Additionally, point cloud extraction is prone to external disturbances, leading to unstable data quality. Prediction models based solely on structural parameters typically achieve $R^2 \approx 0.70$ and RMSE ≈ 640.96 kg/hm² [23]. Critically, single-modal features cannot adequately address the phenological heterogeneity common in breeding populations. Different rice genotypes exhibit substantial variation in growth progression, and static features from single time points are highly susceptible to phenological asynchrony, which can reduce prediction accuracy by 20–30% under multi-genotype conditions [24].

Nevertheless, single-modal features retain value in specific contexts. Spectral information from the flowering stage shows the strongest correlation with yield, with sensitive bands centered around 590 nm [14]. Furthermore, the predictive utility of individual features varies by variety: the Ronaldo variety relies heavily on multi-temporal NDVI ($R^2 = 0.76$), whereas the Gladio variety is better predicted by single-temporal NDVI ($R^2 = 0.88$) [24]. These observations underscore the importance of selective feature application, as effectiveness depends strongly on genotype, growth stage, and environmental conditions.

To contextualize our findings, we compared our results with several recent UAV-based yield prediction studies. Zhou et al. [16] developed a CNN-M2D model for multi-variety rice yield prediction, achieving an R^2 of 0.73 and an RRMSE of 8.13% on their validation set. In contrast, our method achieved an R^2 of 0.80 under a more demanding independent validation scheme (cross-year and cross-location). Feng et al. [19] reported an R^2 of 0.88 using an MLT-Bi-LSTM architecture, but their study was conducted on a relatively uniform set of varieties under similar management conditions. Our study involved 630 accessions with diverse genetic backgrounds across two locations. The lower R^2 observed here (0.80) likely reflects the greater difficulty of predicting yield in highly heterogeneous breeding

populations and the rigorous cross-year validation protocol. Liu et al. [13] achieved an R^2 of 0.94 at the grain-filling stage using deep neural networks, but their study was restricted to a single location with limited genotypic variation. Collectively, these comparisons indicate that although the absolute R^2 of our method (0.80) is lower than some reported values, the validation conditions—cross-year, cross-location, and diverse germplasm—are substantially more challenging, and our method demonstrates competitive or superior generalizability.

4.2. Potential of K-Shape Clustering Algorithm in Yield Prediction

The K-Shape algorithm quantifies the shape similarity of time-series curves using normalized cross-correlation (NCC), thereby overcoming limitations associated with single-modal features. Unlike traditional k-means clustering, which relies on Euclidean distance and is sensitive to amplitude variations, K-Shape is scale-invariant and translation-invariant. This property allows it to capture similarities in crop growth patterns based on shape rather than amplitude, reducing systematic errors caused by phenological asynchrony [17,22]. Further analysis of clustering outcomes revealed insights into genotype-by-environment interactions, such as differential responses to nitrogen fertilization and barren tolerance among materials assigned to different clusters [20], providing valuable theoretical and practical guidance for precision breeding. Moreover, K-Shape enables multi-stage prediction; clustering based on data from the heading and filling stages achieved accuracy comparable to that obtained using full-growth-stage data, thereby improving data collection efficiency [25].

Recent advances in deep temporal sequence clustering (DTSC), including improved deep clustering autoencoders (DCAE) with attention mechanisms and Transformer architectures, have further enhanced the ability to model complex nonlinear temporal patterns [26]. For example, a Transformer decoder based on K-Shape clustering has shown superior performance in complex time-series prediction tasks, offering a novel approach for agricultural time-series analysis [27] and further improving the robustness of rice yield prediction models.

4.3. Advantages of Multimodal Data Fusion in Yield Prediction

Multimodal data fusion offers significant advantages for predicting yield in rice breeding materials, owing to the complementary nature of different data modalities. Our bimodal model, integrating vegetation indices (EVI2) with structural parameters (CH, CV), achieved an R^2 of 0.80 and an RMSE of 511.42 kg/hm², representing a 10–15% improvement over single-modal models. The comprehensive integration of physiological information (vegetation indices) and morphological information (structural parameters) enables a more complete characterization of rice growth dynamics, while K-Shape clustering mitigates phenological asynchrony. This integration enhances model robustness and adaptability; the redundancy inherent in heterogeneous data allows mutual validation, thereby increasing model reliability [28]. Furthermore, under suboptimal weather conditions, the degradation in prediction accuracy of multimodal models is less pronounced than that of single-modal models. Temporal fusion of multimodal data captures dynamic growth characteristics—such as the dominance of structural parameters during tillering and the importance of EVI2 during filling—and dynamic weighted fusion strategies further improve performance [29,30].

Multimodal data fusion also supports intelligent breeding decisions by providing multi-dimensional insights into rice growth processes. It has the potential to reduce breeding costs and improve screening efficiency by narrowing the number of accessions that require manual field measurement [31,32]. Machine learning techniques, including Random Forest and convolutional neural networks (e.g., ResNet50), have demonstrated

superior performance in multimodal data integration, especially when combined with attention mechanisms [20]. However, the effectiveness of multimodal fusion depends strongly on data quality and fusion strategy; precise preprocessing, comprehensive feature extraction [31], and appropriate fusion methods are essential prerequisites for optimal performance [33].

One notable observation in our results is the higher RMSE on the independent validation set (950.6 kg/hm²) compared with the cross-validation error within the training set (511.4 kg/hm²). This increase likely reflects the combined effects of cross-year (2024 vs. 2022) and cross-location (Huai'an vs. Rugao) differences. Several factors may contribute to this performance gap. First, genetic materials differed between sites: traditional Taihu Lake Basin germplasm in Rugao versus newly bred elite lines in Huai'an. These genotypic differences may interact with environmental conditions, leading to distinct phenotypic responses. Second, climatic conditions varied, with Rugao having a higher annual temperature (14.6 °C vs. 14.0 °C) and greater precipitation (1055.5 mm vs. 940 mm), which may affect phenological development rates and canopy structure dynamics. Third, plot sizes differed (3.6 m² in Rugao vs. 10.25 m² in Huai'an), potentially influencing the representativeness of canopy measurements. Despite these spatiotemporal differences, our model maintained a validation R² of 0.79, indicating reasonable robustness. This level of cross-location generalizability is comparable to or better than previous studies that evaluated model transferability across diverse environments [16,25]. Nevertheless, the elevated RMSE suggests that further improvements are needed for practical deployment across diverse conditions. Future work should explore domain adaptation techniques, such as transfer learning or adversarial domain alignment, to further enhance cross-location and cross-year generalizability.

5. Conclusions

This study proposes an integrated method combining UAV time-series imagery and growth process clustering, which effectively addresses the key issue of low prediction accuracy for rice breeding material yield. The study confirms that it is challenging to attain high-precision prediction by relying solely on single-stage spectral or structural parameters, whereas characterizing growth dynamics through time-series curves and conducting K-Shape clustering can significantly optimize prediction performance. The central findings indicate that the bimodal fusion model integrating vegetation indices and canopy structural parameters demonstrates the optimal performance, with a coefficient of determination (R²) of 0.80 and a root mean square error (RMSE) of 511.42 kg/hm², reducing the prediction error by 14.7% compared with the single-modal system. This study integrates multi-source time-series data with clustering analysis, thereby providing reliable technical support for high-yield crop breeding.

Author Contributions: Conceptualization, H.Z., W.C. and Y.Z. (Yan Zhu); Methodology, H.Z., W.C. and Y.Z. (Yan Zhu); Software, Q.K., X.H. and Y.Z. (Yan Zhao); Validation, Q.K., D.W., C.G. and X.H.; Formal Analysis, Q.K., D.W., C.G. and X.H.; Investigation, Q.K., D.W., and X.H.; Resources, D.W., A.Z., W.C. and H.Z.; Data Curation, Q.K., D.W., Y.Z. (Yan Zhao) and X.H.; Writing—Original Draft Preparation, Q.K., D.W., and X.H.; Writing—Review and Editing, H.Z., W.C., Y.Z. (Yan Zhu), A.Z., C.J., X.Y. and T.C.; Visualization, Q.K., Y.Z. (Yan Zhao) and D.W.; Supervision, H.Z., W.C., Y.Z. (Yan Zhu), A.Z., C.J., X.Y. and T.C.; Project Administration, H.Z., W.C., Y.Z. (Yan Zhu), A.Z., C.J., X.Y. and T.C.; Funding Acquisition, H.Z., W.C., Y.Z. (Yan Zhu), A.Z., C.J., X.Y. and T.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Key Research and Development Program of China (2024YFD2301100), the Zhongshan Biological Breeding Laboratory (ZSBBL-KY2025-1), the Key Independent Research Project of Jiangsu Key Laboratory of Information Agriculture (KLIAZZ2301),

and the Jiangsu Collaborative Innovation Center for Modern Crop Production (JCICMCP), the Jiangsu Provincial Young Elite Scientists Sponsorship Program (JSTJ-2024-429).

Data Availability Statement: The data presented in this study are available on request from the corresponding author due to privacy or ethical restrictions.

Acknowledgments: The raw UAV imagery and intermediate data processing files used in this study are very large in volume. At the current stage, these data are available from the corresponding author upon reasonable request. During the preparation of this manuscript, the authors used ChatGPT 4.0 (OpenAI) and Doubao-Seed-2.0 Pro model for the purposes of language polishing, manuscript revision, and formatting assistance. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Yang, Q.; Shi, L.S.; Han, J.Y.; Zha, Y.Y.; Zhu, P.H. Deep convolutional neural networks for rice grain yield estimation at the ripening stage using UAV-based remotely sensed images. *Field Crops Res.* **2019**, *235*, 142–153. [\[CrossRef\]](#)
2. Feng, A.J.; Zhou, J.F.; Vories, E.D.; Sudduth, K.A.; Zhang, M.N. Yield estimation in cotton using UAV-based multi-sensor imagery. *Biosyst. Eng.* **2020**, *193*, 101–114. [\[CrossRef\]](#)
3. Du, M.; Noguchi, N. Monitoring of wheat growth status and mapping of wheat yield's within-field spatial variations using color images acquired from UAV-camera system. *Remote Sens.* **2017**, *9*, 289. [\[CrossRef\]](#)
4. Yang, G.; Liu, J.; Zhao, C. Comprehensive estimation of yield-related traits using UAV-derived multispectral images to improve rice grain yield prediction. *Comput. Electron. Agric.* **2022**, *192*, 106580.
5. Yang, Q.; Shi, L.S.; Han, J.Y.; Chen, Z.W.; Yu, J. A VI-based phenology adaptation approach for rice crop monitoring using UAV multispectral images. *Field Crops Res.* **2022**, *277*, 13. [\[CrossRef\]](#)
6. Claverie, M.; Ju, J.; Masek, J.G.; Dungan, J.L.; Vermote, E.F.; Roger, J.C.; Skakun, S.V.; Justice, C. The Harmonized Landsat-Sentinel-2 surface reflectance data set. *Remote Sens. Environ.* **2018**, *219*, 145–161. [\[CrossRef\]](#)
7. Rouse, J.W., Jr.; Haas, R.H.; Deering, D.W.; Schell, J.A.; Harlan, J.C. *Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation*; NASA/GSFC Final Report; NASA: Greenbelt, MD, USA, 1974.
8. Gitelson, A.; Merzlyak, M.N. Spectral reflectance changes associated with autumn senescence of *Aesculus hippocastanum* L. and *Acer platanoides* L. leaves. Spectral features and relation to chlorophyll estimation. *J. Plant Physiol.* **1994**, *143*, 286–292. [\[CrossRef\]](#)
9. Fitzgerald, G.J.; Rodriguez, D.; Christensen, L.K.; Belford, R.; Sadras, V.O.; Clarke, T.R. Spectral and thermal sensing for nitrogen and water status in rainfed and irrigated wheat environments. *Precis. Agric.* **2006**, *7*, 233–248. [\[CrossRef\]](#)
10. Gitelson, A.A.; Gritz, Y.; Merzlyak, M.N. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *J. Plant Physiol.* **2003**, *160*, 271–282. [\[CrossRef\]](#)
11. Shen, Y.; Zhang, X.; Gao, S.; Zhang, H.K.; Schaaf, C.; Wang, W.; Ye, Y.; Liu, Y.; Tran, K.H. Analyzing GOES-R ABI BRDF-adjusted EVI2 time series by comparing with VIIRS observations over the CONUS. *Remote Sens. Environ.* **2024**, *302*, 17. [\[CrossRef\]](#)
12. Shi, G.W.; Du, X.; Du, M.W.; Li, Q.Z.; Tian, X.L.; Ren, Y.T.; Zhang, Y.; Wang, H.Y. Cotton yield estimation using the remotely sensed cotton boll index from UAV images. *Drones* **2022**, *6*, 254. [\[CrossRef\]](#)
13. Yang, X.; Xu, L.; Zhang, J.; Xu, T.; Ma, W. Rice yield estimation based on vegetation index and fluorescence spectral information from UAV hyperspectral remote sensing. *Remote Sens.* **2022**, *14*, 4892.
14. Dong, T.; Liu, J.; Qian, B.; Shang, J. Improving regional crop yield forecasts by assimilating multi-source remote sensing data into the WOFOST model with a particle filter. *Remote Sens. Environ.* **2021**, *258*, 112379.
15. Liu, J.K.; Wang, W.Q.; Su, X.X.; Li, J.; Nian, Y.; Zhu, X.Q.; Ma, Q.; Li, X.W. Prediction of rice yield and nitrogen use efficiency based on UAV multispectral imagery and machine learning. *Trans. Chin. Soc. Agric. Eng.* **2025**, *41*, 127–138.
16. Zhou, H.; Huang, F.; Lou, W.; Gu, Q.; Ye, Z.; Hu, H.; Zhang, X. Yield prediction through UAV-based multispectral imaging and deep learning in rice breeding trials. *Agric. Syst.* **2025**, *223*, 104214. [\[CrossRef\]](#)
17. Paparrizos, J.; Gravano, L. k-shape: Efficient and accurate clustering of time series. In Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data, Melbourne, Australia, 31 May–4 June 2015; pp. 1855–1870.
18. Li, Q.; Zhao, S.; Du, L.; Luo, S. Multi-genotype rice yield prediction based on time-series remote sensing images and dynamic process clustering. *Agriculture* **2025**, *15*, 64. [\[CrossRef\]](#)

19. Feng, X.; Li, Z.; Yang, P.; Hong, W.; Wang, A.; Qin, J.; Zhang, H.; Senou, P.D.K.; Zhang, Y.; Wang, D.; et al. Enhance the accuracy of rice yield prediction through an advanced preprocessing architecture for time series data obtained from a UAV multispectral remote sensing platform. *Eur. J. Agron.* **2025**, *165*, 127542. [[CrossRef](#)]
20. Ji, W.H.; Zheng, H.B.; Wang, D.; Tang, W.J.; Zhang, X.H.; Guo, C.L.; Yao, X.; Jiang, C.Y.; Zhu, Y.; Cao, W.X.; et al. Yield prediction of rice breeding materials based on UAV imagery and convolutional neural networks. *J. Nanjing Agric. Univ.* **2025**, *48*, 1001–1012.
21. Muramatsu, K.; Yoneda, E.; Soyama, N.; López-Ballesteros, A.; Thanyapraneeekul, J. Use of light response curve parameters to estimate gross primary production capacity from chlorophyll indices of global observation satellite and flux data. *Sci. Remote Sens.* **2024**, *302*, 17. [[CrossRef](#)]
22. Bariya, M.; Meier, A.; Paparrizos, J.; Franklin, M.J. k-shape stream: Probabilistic streaming clustering for electric grid events. In Proceedings of the 2021 IEEE Madrid PowerTech, Madrid, Spain, 28 June–2 July 2021; pp. 1–6.
23. Kaushal, S.; Gill, H.S.; Billah, M.M.; Khan, S.N.; Halder, J.; Bernardo, A.; St Amand, P.; Bai, G.; Glover, K.; Maimaitjiang, M.; et al. Enhancing the potential of phenomic and genomic prediction in winter wheat breeding using high-throughput phenotyping and deep learning. *Front. Plant Sci.* **2024**, *15*, 1410249. [[CrossRef](#)] [[PubMed](#)]
24. Perros, N.; Kalivas, D.; Giovos, R. Spatial analysis of agronomic data and UAV imagery for rice yield estimation. *Agriculture* **2021**, *11*, 809. [[CrossRef](#)]
25. Li, Z.; Chen, Z.; Cheng, Q.; Fei, S.; Zhou, X. Deep learning models outperform generalized machine learning models in predicting winter wheat yield based on multispectral data from drones. *Drones* **2023**, *7*, 505. [[CrossRef](#)]
26. Alqahtani, A.; Ali, M.; Xie, X.; Jones, M.W. Deep time-series clustering: A review. *Electronics* **2021**, *10*, 3001. [[CrossRef](#)]
27. Yang, H.; Yan, C.; Wang, P. A K-Shape clustering based transformer-decoder model for predicting multi-step potentials of urban mobility field. *IEEE Trans. Intell. Transp. Syst.* **2024**, *25*, 10891–10902. [[CrossRef](#)]
28. Ji, Y.; Liu, Z.; Cui, Y.; Liu, R.; Chen, Z.; Zong, X.; Yang, T. Faba bean and pea harvest index estimations using aerial-based multimodal data and machine learning algorithms. *Plant Physiol.* **2024**, *194*, 1512–1526. [[CrossRef](#)]
29. Che, Y.; Wang, Q.; Li, S.; Li, B.; Ma, Y. Monitoring of maize phenotypic traits using super-resolution reconstruction and multimodal data fusion. *Trans. Chin. Soc. Agric. Eng.* **2021**, *37*, 169–178.
30. Zhao, X.; Wang, B.; Du, X.; Wang, W.; Ding, Z.; Zhou, C.; Zhang, K. Evapotranspiration prediction model of tea garden based on temporal convolutional network and Transformer. *Trans. Chin. Soc. Agric. Mach.* **2024**, *55*, 337–346.
31. Wang, Y.; Ma, X.; Tan, S.; Jia, X.; Chen, J.; Qin, Y.; Hu, X.; Zheng, H. Inverting rice nitrogen content with multimodal data fusion of unmanned aerial vehicle remote sensing and ground observations. *Trans. Chin. Soc. Agric. Eng.* **2024**, *40*, 100–109.
32. Wang, C.; Liu, L. Winter wheat yield prediction at county-level based on multi-modal data fusion. *Trans. Chin. Soc. Agric. Eng.* **2025**, *41*, 162–172.
33. Saki, M.; Keshavarz, R.; Franklin, D.; Abolhasan, M.; Lipman, J.; Shariati, N. A data-driven review of remote sensing-based data fusion in precision agriculture from foundational to transformer-based techniques. *IEEE Access* **2025**, *13*, 166188–166209. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.