UAV-based multi-sensor data fusion and machine learning algorithm for yield prediction in wheat

Shuaipeng Fei1 · Muhammad Adeel Hassan2,5 · Yonggui Xiao2 · Xin Su3 · Zhen Chen1 · Qian Cheng1 · Fuyi Duan1 · Riqiang Chen6 · Yuntao Ma4

Accepted: 30 June 2022 / Published online: 3 August 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract
Early prediction of grain yield helps scientists to make better breeding decisions for wheat. Use of machine learning (ML) methods for fusion of unmanned aerial vehicle (UAV)-based multi-sensor data can improve the prediction accuracy of crop yield. For this, five ML algorithms including Cubist, support vector machine (SVM), deep neural network (DNN), ridge regression (RR) and random forest (RF) were used for multi-sensor data fusion and ensemble learning for grain yield prediction in wheat. A set of thirty wheat cultivars and breeding lines were grown under three irrigation treatments i.e., light, moderate and high irrigation treatments to evaluate the yield prediction capabilities of a low-cost multi-sensor (RGB, multi-spectral and thermal infrared) UAV platform. Multi-sensor data fusion-based yield prediction showed higher accuracy compared to individual-sensor data in each ML model. The coefficient of determination ($R^2$) values for Cubist, SVM, DNN and RR models regarding grain yield prediction were observed from 0.527 to 0.670. Moreover, the results of ensemble learning through integrating the above models illustrated further increase in accuracy. The predictions of ensemble learning showed high $R^2$ values up to 0.692, which was higher as compared to individual ML models across the multi-sensor data. Root mean square error (RMSE), residual prediction deviation (RPD) and ratio of prediction performance to inter-quartile range (RPIQ) were calculated to be 0.916 t ha$^{-1}$, 1.771 and 2.602, respectively. The results proved that low altitude UAV-based multi-sensor data can be used for early grain yield prediction using data fusion and an ensemble learning framework with high accuracy. This high-throughput phenotyping approach is valuable for improving the efficiency of selection in large breeding activities.

Keywords Data fusion · Machine learning · Phenotyping · Wheat · Unmanned aerial vehicle
Introduction

Wheat provides approximately 20% of the food to fulfill calories and protein requirements of 4.5 billion people (Nigro et al., 2014). The crop breeder’s major goal is to produce diverse germplasm which could have resistance to climate change and have high-yielding attributes (Shafiee et al., 2021). Present methods that enable breeders to estimate grain yield through secondary traits at early plant development stages, are destructive, low throughput, costly and time consuming (Lee et al., 2018; Montesinos-López et al., 2017; Shafiee et al., 2021). Recently, use of unmanned aerial vehicles (UAVs)-based optical sensors have been introduced as a low-cost, efficient, high throughput phenotyping platform (HTPP) for estimation of important attributes of crops (Hassan et al., 2018; Yang et al., 2020). This would increase the selection intensity and accuracy to obtain high genetic gain for grain yield (Rutkoski et al., 2016). Despite successful phenotyping of several crop attributes through UAV-based information, there are significant limitations in using a single sensor to estimate particular traits. Recently, multi-sensor data fusion of RGB, multi-spectral (MS) and thermal information has been used to improve the prediction accuracy of important crop traits such as chlorophyll content, above ground biomass, leaf area index (LAI), nitrogen level and yield (Feng et al., 2020a; Liu et al., 2021; Maimaitijiang et al., 2017, 2020). Multi-sensor data fusion by integrating the benefits of spectral, spatial, structural and temperature information from multiple sensors can improve estimates of plant traits (Maimaitijiang et al., 2020).

UAV-based multi-sensor platforms are able to collect large amounts of high-dimensional data in a short period of time, which poses a challenge to modeling methods (Montesinos-López et al., 2017). The rapid development of computer science has promoted the advancement of machine learning (ML) algorithms, which has become a hot topic in quantitative remote sensing. ML methods such as ridge regression (RR), Gaussian process (GP), deep neural networks (DNN), random forest (RF) and support vector machine (SVM) are being used to build predictive models for plant traits using multiple types of remote sensing data as input (Fu et al., 2019; Jin et al., 2020; Matese & Di Gennaro, 2021). ML methods have increased the prediction accuracy and robustness of UAV-based sensor data sets. However, the optimal ML model for estimating crop traits varies both spatially and temporally. For example, although RF was shown to be the best model for estimating wheat biomass (Han et al., 2019; Wang et al., 2016), it was found to be less accurate for evaluating alfalfa yield using hyperspectral data (Feng et al., 2020b). A similar limitation has also been reported in the use of extreme learning machine (ELM) for estimating crop traits. Maimaitijiang et al. (2017) reported that ELM was the best model for estimating soybean traits, with significantly improved accuracy over partial least squares regression (PLSR) and SVM. However, in a recent study, ELM accuracy was lower than PLSR and RF when estimating sorghum chlorophyll (Bhadra et al., 2020). This suggests that the models that have been well-trained could perform poorly for estimation of plant parameters in other species and under different growth environments. This is most likely because all of these regression techniques are based on an individual prediction model and are susceptible to overfitting when working with less training data (Pal, 2007). Ensemble learning models provide better predictive performance than individual models. Stacking regression is a powerful ensemble learning technique that employs diverse learners and increases their differences to enhance ultimate accuracy (Feng et al., 2020b). The condition of diversity ensures the ability to provide information complementarity between different
base learners, and is key to ensure that final models achieve correct results. The stacking method has been used increasingly in the field of precision agriculture. It has achieved higher accuracy than the best-performing individual model to estimate winter wheat yield (Fei et al., 2021), tobacco photosynthetic capacity (Fu et al., 2019) and alfalfa yield (Feng et al., 2020b). To date, no studies have been reported to predict winter wheat yield with multiple sensors combined with the stacking ensemble learning method.

Taking all this into consideration, the main objectives of the study were (1) to evaluate UAV-based multi-sensor data fusion for wheat yield prediction at grain filling stage and (2) to develop an ensemble learning framework to improve yield prediction accuracy of ML models.

**Materials and methods**

**Experiment location and design**

The experiment was conducted at a research site of Chinese Academy of Agricultural Sciences (113° 45′ 40″ E, 35° 8′ 11″ N) in Xinxiang, Henan province, China. Thirty cultivars which have been released in the last few decades in China’s Yellow and Huai Valleys Winter Wheat Zone (YHVWWZ) were used in this study. Cultivars were cultivated in plots with dimensions of 11.2 m², 8 m in length and 1.4 m in width (Fig. 1). Each plot represented one cultivar with six rows spaced at 0.20 m. A randomized block design was implemented with three irrigation treatments (light irrigation, moderate irrigation and high irrigation) and replications were adopted for each treatment with a total number of 180 plots. The irrigation was done at different growth stages through a large movable sprinkler irrigation device. Detailed information on irrigation and environmental conditions (sunshine duration, precipitation and maximum temperature) for the winter wheat growing season in 2019–2020 at the experimental location is given in Fig. 2. Field fertilization, pest and disease management were all maintained at optimal levels according to local agricultural standards.

![Fig. 1 Experimental design](image-url)
Grain yield measurement

All the cultivated plots were harvested using a plot combine harvester in early June. The wheat grains from each plot were collected in plastic mesh bags for drying and weighed at a moisture content of approximately 12.5%. The t-test results revealed that irrigation had a significant effect on grain yield ($P < 0.001$) (Fig. 3).

Fig. 2 The profile of meteorological variables during the wheat growing season in 2019–2020 and the volume of sprinkler irrigation. a Sunshine duration, b max temperature, c precipitation, and d irrigation volume and time point for the three irrigation treatments. Meteorological data was gathered from local weather stations.

Fig. 3 Violin diagram and t-test of measured grain yield under three irrigation treatments. ***Indicates significant at the 0.001 level.
During the trials, an M210 (SZ DJI Technology Co., Shenzhen, China) equipped with a Red-Edge MX MS camera (MicaSense Inc., Seattle, USA), a Zenmuse XT2 camera (SZ DJI Technology Co., Shenzhen, China) with RGB and thermal infrared (TIR) lenses was used for multi-sensor data acquisition (Fig. 4). The detailed information of the sensors is given in Table 1. The aerial surveys were carried out at the grain filling stage due to the proven high accuracy of yield predictions during this period (Guan et al., 2017; Hassan et al., 2019; Qader et al., 2018).
The Red-Edge MX has five spectral lenses: blue, green, red, red-edge and near-infrared. The sunlight sensor of the Red-Edge MX camera automatically adjusts the ambient light effects to reduce the inaccuracy of MS images (Hassan et al., 2019). Each monochrome sensor captured the same resolution (1280×960) image with a 10 nm bandwidth for the red and red-edge band (half maximum bandwidth), 20 nm bandwidth for blue and green and 40 nm bandwidth for the infrared. The radiometric calibration was done before and after each flight using a calibration board in order to translate the DN value of the MS data into reflectance. The Zenmuse XT2 has two lenses i.e., TIR lens (640×512 pixels) and a RGB lens (4000×3000 pixels). Temperature observations in 7.5–13.5 μm spectral region are recorded with a 5 °C thermal sensitivity using the TIR lens. During the flights, a portable thermometer was used to measure the surface temperature of boards for calibration of thermal images.

The DJI ground station software was utilized as an automated flight control system, letting users create custom missions and establish their own air routes. Flight missions were carried out from 11.00 to 13.00 under clear sky. All flights were flown at 30 m height to get high quality photos, with 85% front and 80% side image overlap for all cameras. Each sensor was featured with a built-in GNSS module that was used to get the location information of the images. Co-ordinates of ground control points (GCPs) were recorded with millimeter precision using a differential global navigation satellite system.

Image processing

The post-imaging process for data extraction is shown in Fig. 5. The images from the three sensors were pre-processed in the Pix4D mapper (Pix4D, Lausanne, Switzerland) to generate ortho-mosaics, digital surface model (DSM) and to implement radiometric calibration. The important steps of this procedure included geolocating the images, importing GCPs, aligning the images, building dense point cloud, DSM and ortho-mosaic and calibrating radiometric information (Han et al., 2019). Dense point clouds were generated using the structure from motion (SfM) method in Pix4Dmapper along with photogrammetric workflow. The ortho-mosaic images of each sensor were segmented into 180 polygon shapes with given IDs identifying the various plots to retrieve relevant information for each plot in the field. ArcMap 10.5 (Environmental Systems Research Institute, Inc., Redlands, USA) was used to create polygon forms. To minimize the impact of marginal effects, the edge of each plot was omitted while creating the shp file. Corresponding images and a polygon shapefile layer of plot borders with individual IDs were all submitted to ENVI 5.3 (Exelis Visual Information Solutions, Inc., Boulder, USA) for feature extraction. The average of all the extracted pixel values of each plot was used as the corresponding feature.

For the MS images, radiometric calibration was utilized to enhance the radiometric quality of the data and convert the DN values to reflectance by utilizing radiometric calibration images with known reflectance values. For thermal image calibration, the DN values of the blackboard were extracted using ArcMap 10.5. A linear relationship ($T = 0.2813\times\text{DN} - 295.98$) was established between the actual temperature and the DN of the blackboard. The temperature distribution for the whole experiment was then calculated using this equation with the raster calculator in ArcMap 10.5. Figure 6a shows the scatter plot of the measured temperature of the blackboard vs. the temperature derived from the thermal image of the blackboard. The results of a t-test showed that there was a significant difference in canopy temperature between different irrigation treatments (Fig. 6b).
Crop height, texture and vegetation index extraction

The features extracted from each sensor image are shown in Table 2. Crop height model (CHM), which has been extensively used to extract crop height information from various crops (Bendig et al., 2015), was used to calculate crop height. The UAV-based RGB images were acquired prior to plant emergence on October 24, 2019, in order to generate a bare-soil digital elevation model (DEM) including photogrammetric 3D point clouds. A distinct 3D point cloud model depicting a digital surface model (DSM) of all objects on the ground was created using RGB images acquired on April 20, 2020. A CHM was obtained by subtracting DSM and DEM pixel-by-pixel. To assess the accuracy of CHM, ten plants from each plot were selected for crop height measurement and averaged to represent the heights of the plot. High correlation ($R^2 = 0.84$) was observed between UAV and ground-based crop height (Fig. 6c). Significant differences ($P < 0.001$) were found in UAV-based crop height under different irrigation treatments (Fig. 6d).

Texture features were extracted from the MS-based blue, green, red, red-edge, NIR bands and RGB-based crop height and thermal-based canopy temperature. The widely utilized gray level co-occurrence matrix (GLCM) was chosen to investigate the yield prediction with texture information. The ENVI 5.3 software was used to calculate GLCM-based texture features, including mean (ME), variance (VA), homogeneity (HO), contrast (CO), dissimilarity (DI), entropy (EN), second moment (SE) and correlation (COR). Vegetation
indices (VIs) are commonly employed to measure crop attributes in agricultural fields. In this study, 24 VIs that have been found to be highly linked with crop traits were used (Table 2). These VIs were calculated from DN values of each band in the RGB images or reflectance of each band in the MS images.

**Ensemble learning framework**

Five widely used machine learning (ML) methods including Cubist (Quinlan, 1992), support vector machine (SVM) (Cortes & Vapnik, 1995), deep neural network (DNN) (Qiu et al., 2014), ridge regression (RR) (Hoerl & Kennard, 1970) and random forest (RF) (Breiman, 2001), were utilized as base learners in this study to explore yield prediction accuracy of ensemble learning and multi-sensor data fusion. Appropriate hyper-parameters play a crucial role in the successful application of ML algorithms; the candidate hyper-parameters of these five algorithms are shown in Table 3.

Stacking regression is an ensemble learning model proposed by Wolpert (1992) to blend the predictors and improve prediction accuracy (Fei et al., 2021). To validate the stacking
## Table 2 Definitions of the features derived from various sensors

| Sensor          | Feature                                         | Formulation                                      | References                           |
|-----------------|-------------------------------------------------|--------------------------------------------------|--------------------------------------|
| RGB             | Color intensity                                 | INT = (R + G + B) / 3                            | Ahmad and Reid (1996)                |
|                 | Kawashima index                                 | IKAW = (R − B) / (R + B)                         | Kawashima and Nakatani (1998)        |
|                 | Principal component analysis index              | IPCA = 0.994 | R − B | 0.961 | G − B | 0.914 | G − R | 0.994 | uni007C.var R−B | uni007C.var G−B | uni007C.var G−R | uni007C.var | | Saberioon et al. (2014) |
|                 | Excess red index                                | ExR = 1.4R − G                                   | Meyer and Neto (2008)                |
|                 | Excess green index                              | ExG = 2G − R − B                                | Meyer and Neto (2008)                |
|                 | Excess green minus excess red index             | ExGR = 2G − R − B − (1.4R − G)                  | Meyer and Neto (2008)                |
|                 | Modified Green Red Vegetation Index             | MGRVI = (G^2 − R^2) / (G^2 + R^2)               | Bendig et al. (2015)                 |
|                 | Red Green Blue Vegetation Index                 | RGBVI = (G^2 − B * R) / (G^2 + B * R)           | Bendig et al. (2015)                 |
|                 | Crop height                                     | DSM-DEM                                          | /                                   |
| MS              | Gray-level co-occurrence matrix                 | ME, VA, HO, CO, DI, EN, SE, COR                 | Haralick and Shanmugam (1973)        |
|                 | Normalized Difference Vegetation Index          | NDVI = (NIR − R) / (NIR + R)                     | Rouse et al. (1974)                  |
|                 | Green-NDVI                                      | GNDVI = (NIR − G) / (NIR + G)                    | Gitelson et al. (1996)               |
|                 | Ratio Vegetation Index                          | RVI = (NIR/R)                                    | Tucker (1979)                        |
|                 | Normalized difference red-edge index            | NDREI = (NIR − RE) / (NIR + RE)                  | Barnes et al. (2000)                 |
|                 | Enhanced Vegetation Index                       | EVI = 2.5 * (NIR − R) / (1 + NIR − 2.4 * R)     | Huete et al. (2002)                  |
|                 | Optimized Soil-Adjusted Vegetation Index        | OSAVI = (NIR − R) / (NIR − R + 0.16)             | Rondeaux et al. (1996)               |
|                 | Modified chlorophyll absorption in reflectance index | MCARI = [(RE − R) − 0.2 * (RE − G)] * (RE/R) | Daugorthy et al. (2000)              |
|                 | Transformed chlorophyll absorption in reflectance index | TCARI = 3 * [(RE − R) − 0.2 * (RE − G) * (RE/R)] | Haboudane et al. (2002)             |
|                 | Nitrogen Reflectance Index                      | NRI = (G − R) / (G + R)                          | Schleicher et al. (2001)             |
|                 | Transformational Vegetation Index               | TVI = √NDVI + 0.5                                | Broge and Leblanc (2001)             |
|                 | Modified Simple Ratio Index                     | MSR = [(NIR/R) − 1] / √NIR/R + 1                 | Chen (1996)                          |
|                 | Structure Insensitive Pigment Index             | SIP = (NIR − B) / (NIR + B)                      | Penuelas et al. (1995)               |
|                 | Plant Senescence Reflectance Index              | PSRI = (R − B) / NIR                            | Merzlyak et al. (1999)               |
|                 | Chlorophyll Index Red-Edge                     | CIRE = NIR/RE − 1                                | Gitelson et al. (2003)               |
|                 | MCARI/OSAVI                                    | MCARI/OSAVI                                      | Daugorthy et al. (2000)              |
|                 | TCARI/OSAVI                                    | TCARI/OSAVI                                      | Haboudane et al. (2002)              |
Multi-spectral vegetation indices were calculated from the reflectance of each band, and the RGB vegetation indices were calculated from the DN value of each band. 

*MS* multi-spectral, *TIR* thermal infrared, *DSM* digital surface model, *DEM* digital elevation model, *ME* mean, *VA* variance, *HO* homogeneity, *CO* contrast, *DI* dissimilarity, *EN* entropy, *SE* second moment, *CO* correlation

Table 2 (continued)

| Sensor | Feature                             | Formulation         | References                           |
|--------|-------------------------------------|---------------------|--------------------------------------|
|        | Gray-level co-occurrence matrix     | ME, VA, HO, CO, DI, EN, SE, COR | Haralick and Shanmugam (1973)       |
| TIR    | Canopy temperature                  | /                   |                                      |
|        | Gray-level co-occurrence matrix     | ME, VA, HO, CO, DI, EN, SE, COR | Haralick and Shanmugam (1973)       |
ensemble learning (Fig. 7a), all of the data were initially separated into training and test sets in a 4:1 ratio. Each of the five ML methods was used to do a five-fold outer cross-validation (CV) on the training set. The training data were randomly and uniformly divided into five equal sets. Four fifths of which were used for outer CV training and one for outer

| Machine learning method | Hyperparameters |
|-------------------------|-----------------|
| Cubist                  | Committees: 1 to 25; neighbors: 1 to 9 |
| SVM                     | Kernel function: Radial basis function; \( \text{Sigma} \): from 0.01 to 0.02 with increments of 0.0001; \( \text{C} \): from 0.5 to 0.7 with increments of 0.01 |
| DNN                     | Units: from 10 to 100 with increments of 10; epochs: from 2 to 120 with increments of 2; hidden layers: 1, 2, 3 and 4; regularization method: dropout; activation function: rectified linear activation unit function |
| RR                      | \( K \): from 0 to 0.02 with increments of 0.001 |
| RF                      | \( ntree \): from 1000 to 10,000 with increments of 500; \( mtry \): from 1 to 100 with increments of 2 |

SVM support vector machine, DNN deep neural network, RR ridge regression, RF random forest

Fig. 7 A workflow of a multi-sensor data fusion and ensemble learning, and b outer and inner cross-validation. MS multi-spectral features, TIR thermal infrared features, SVM support vector machine, DNN deep neural network, RR ridge regression, RF random forest, CV cross-validation, \( P \) and \( p \) model predictions at different modeling stages
CV validation, and the process was cycled five times, all of the data were utilized as either training samples or validation samples. At the same time, the hyper-parameters of each ML algorithm were being screened. The outer CV training data were randomly divided into 10 equal halves. Nine of the ten aliquots were used for inner CV training and one for inner CV validation, and the process was repeated ten times (Fig. 7b). All hyper-parameter combinations of each ML were submitted to a full ten-fold inner CV, with the hyper-parameter combination with the best average accuracy being sent to the outer CV training data. The five ML models trained by outer CV were tested for model accuracy on the initial test set, and an out-of-sample prediction matrix was constructed on the outer CV validation set. After completing the five-fold CV, five sets of prediction results for the initial test set were generated for each model and averaged to generate the test set prediction matrix for further utilization as the secondary learner’s test set. The RR model was employed as secondary learner to combine the prediction capability of multiple base learners. In level 2, the five-fold CV was also used with the out-of-sample prediction matrix to train the RR (StRR) model. After running the model on the test set prediction matrix, five sets of predicted values were generated, and the final predicted values were obtained by averaging them. The division of the original data into training and test sets in this study was carried out 80 times, and the same division method was used for different data sources or data fusions. In addition, the same split of the five-fold CV was used for different ML models in the same partitioning of the training and test sets, so that the prediction accuracy of different methods can be fairly compared.

### Accuracy evaluation parameters

For accuracy evaluation, coefficient of determination ($R^2$), root mean square error (RMSE), residual prediction deviation (RPD) and ratio of prediction performance to interquartile range (RPIQ), were used to quantitatively estimate the prediction accuracy of ML models to predict grain yield. The lower the RMSE and the higher the $R^2$, RPD and RPIQ, the better the indication of the model’s predictive power. Equations 1–4 were used for estimation of $R^2$, RMSE, RPD and RPIQ, respectively. The formulae for calculating the above parameters are as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}$$  \hspace{1cm} (1)

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{n}}$$  \hspace{1cm} (2)

$$\text{RPD} = \frac{\text{SD}}{\text{RMSE}}$$  \hspace{1cm} (3)

$$\text{RPIQ} = \frac{\text{IQ}}{\text{RMSE}}$$  \hspace{1cm} (4)

where $y_i$ and $\hat{y}_i$ are the measured and the predicted grain yield, respectively. $\bar{y}$ is the mean of measured grain yield and $n$ is the total number of testing samples. SD is the standard deviation of the measured grain yield. IQ is the interquartile spacing of the measured grain yield.
yield. To limit the effect of chance, the original data were divided into training and test sets 80 times. Each division of the ML model then created five tests during the five-fold CV procedure, for a total of 400 tests after 80 divisions. The average of the accuracy parameters of these test results was used as the final prediction accuracy evaluation metric.

Collinearity analysis

The variance inflation factor (VIF) is the relevant and frequently used quantity that can be consulted to examine individual predictors for potentially strong contributions to multicollinearity (Marcoulides & Raykov, 2019). In this study, VIF was used to represent the strength of collinearity between the individual sensor data output predictions and the individual ML output predictions. The multiple linear regression model assumes a dependent variable (Draper, 1998), denoted as $y$ (grain yield in this study), is related as follows to a given set of independent variables, denoted as $x_1,\ldots, x_k$:

$$ y = a_1 x_1 + a_2 x_2 + \cdots + a_k x_k + l $$

where $a_1, a_2, \ldots, a_i$ are regression coefficients and $l$ is the intercept. When some of the predictors are involved in considerable linear relationships among themselves, standard errors for one or more individual partial regression coefficients can be unduly inflated. This tends to produce findings of possible lack of unique significance for substantively important regressors, in the context of other independent variables. The formula for calculating VIF is as follows:

$$ VIF = 1 / (1 - R^2_i) $$

where $R^2_i$ denotes the $R^2$ index when the $i$th explanatory variable is regressed on the remaining independent variables. In general, a multi-collinearity problem occurs when the VIF value is greater than 10 (Marcoulides & Raykov, 2019).

Results

Performance of models for grain yield prediction

Wheat yield prediction was done using Cubist, SVM, DNN, RR and RF regression algorithms using extracted features from three types of sensors (RGB, MS and TIR) and their different combinations (Table 4, Figs. 8 and 9). For individual sensor-based yield prediction, the prediction values of RF were high when employed on MS ($R^2 = 0.509$) and TIR ($R^2 = 0.599$) data sets. DNN yielded highest with $R^2$ value of 0.606 for RGB data. The prediction accuracy of most ML models has been significantly improved after fusing multi-sensor features. Data fusion results of dual sensors showed that RF performed high for both RGB + TIR ($R^2 = 0.670$) and MS + TIR ($R^2 = 0.629$) data fusion. DNN was the best-performing model for data fusion of RGB + MS with $R^2$ value of 0.648. In data fusion of three sensors, DNN achieved the best predictive accuracy ($R^2 = 0.670$). All four ML algorithms Cubist, SVM, DNN and RR except RF, achieved higher prediction values in data fusion of three sensors than dual-sensor and individual-sensor.
To combine the prediction capacity of individual ML models, RR was utilized as secondary learner of stacking regression. The results (Table 4, Figs. 8 and 9) revealed that the prediction accuracy of ensemble models was observed to be higher than individual ML models (Table 4). Compared to the best-performing individual ML models for individual sensor data, the ensemble learning improved the $R^2$ value from 0.605 to 0.624 for RGB, 0.509 to 0.532 for MS and 0.599 to 0.616 for TIR. In data fusion, the prediction accuracy of ensemble learning was also observed to be higher than each of the individual models (Table 4). In data fusion of all three sensors, ensemble learning showed the highest prediction accuracy ($R^2 = 0.692$, RMSE = 0.916, RPD = 1.771 and RPIQ = 2.602).

| Sensor Metric | Sensor | $R^2$ | Cubist | SVM | DNN | RR | RF | StRR |
|---------------|--------|-------|--------|-----|-----|----|----|------|
| RGB           | $R^2$  | 0.514 | 0.597  | 0.606| 0.556| 0.605| 0.624|
|               | RMSE (t ha$^{-1}$) | 1.149 | 1.046 | 1.045| 1.103| 1.034| 1.016|
|               | RPD    | 1.427 | 1.559 | 1.562| 1.481| 1.580| 1.606|
|               | RPIQ   | 2.081 | 2.275 | 2.279| 2.160| 2.306| 2.345|
| MS            | $R^2$  | 0.498 | 0.502  | 0.489| 0.477| 0.509| 0.532|
|               | RMSE (t ha$^{-1}$) | 1.164 | 1.149 | 1.165| 1.188| 1.140| 1.120|
|               | RPD    | 1.400 | 1.413 | 1.393| 1.368| 1.424| 1.449|
|               | RPIQ   | 2.043 | 2.061 | 2.033| 1.998| 2.078| 2.117|
| TIR           | $R^2$  | 0.529 | 0.553  | 0.578| 0.563| 0.599| 0.617|
|               | RMSE (t ha$^{-1}$) | 1.133 | 1.102 | 1.079| 1.146| 1.038| 1.026|
|               | RPD    | 1.449 | 1.486 | 1.517| 1.440| 1.579| 1.594|
|               | RPIQ   | 2.112 | 2.168 | 2.213| 2.100| 2.303| 2.325|
| RGB + MS      | $R^2$  | 0.541 | 0.638  | 0.648| 0.605| 0.622| 0.662|
|               | RMSE (t ha$^{-1}$) | 1.114 | 0.982 | 0.990| 1.043| 1.009| 0.960|
|               | RPD    | 1.464 | 1.652 | 1.637| 1.556| 1.608| 1.690|
|               | RPIQ   | 2.145 | 2.422 | 2.401| 2.284| 2.360| 2.479|
| RGB + TIR     | $R^2$  | 0.527 | 0.615  | 0.631| 0.603| 0.670| 0.671|
|               | RMSE (t ha$^{-1}$) | 1.143 | 1.028 | 1.015| 1.045| 0.955| 0.951|
|               | RPD    | 1.435 | 1.587 | 1.608| 1.563| 1.711| 1.718|
|               | RPIQ   | 2.093 | 2.315 | 2.346| 2.280| 2.498| 2.508|
| MS + TIR      | $R^2$  | 0.543 | 0.591  | 0.596| 0.592| 0.629| 0.640|
|               | RMSE (t ha$^{-1}$) | 1.116 | 1.048 | 1.052| 1.067| 0.998| 0.991|
|               | RPD    | 1.463 | 1.553 | 1.546| 1.524| 1.633| 1.643|
|               | RPIQ   | 2.137 | 2.269 | 2.259| 2.228| 2.388| 2.402|
| RGB + MS + TIR| $R^2$  | 0.563 | 0.666  | 0.670| 0.630| 0.665| 0.692|
|               | RMSE (t ha$^{-1}$) | 1.092 | 0.949 | 0.964| 1.006| 0.956| 0.916|
|               | RPD    | 1.494 | 1.709 | 1.681| 1.612| 1.698| 1.771|
|               | RPIQ   | 2.193 | 2.511 | 2.470| 2.369| 2.496| 2.602|

$MS$ multi-spectral features, $TIR$ thermal infrared features, $SVM$ support vector machine, $DNN$ deep neural network, $RR$ ridge regression, $RF$ random forest, $StRR$ stacking regression using ridge regression as a secondary learner.
which is a significant increase over the best-performing DNN ($R^2 = 0.670$, RMSE = 0.964, RPD = 1.681 and RPIQ = 2.470) in the individual models. In Figs. 8 and 9, boxplots of $R^2$ and RMSE for 400 tests of various models showed obvious fluctuations in parameter accuracy with a wide range among the individual ML models compared to ensemble learning. Similarly, after data fusion, the fluctuation of parameters accuracy for each ML model was reduced. Figure 10 shows the distribution of the regression coefficient of each ML model within the secondary learner (RR). A larger regression coefficient indicated a higher
Fig. 9 The statistical distributions of the prediction accuracy of individual machine learning and ensemble learning for grain yield prediction using multi-sensor data fusion in the modeling test phase. MS multi-spectral features, TIR thermal infrared features, SVM support vector machine, DNN deep neural network, RR ridge regression, RF random forest, StRR stacking regression using ridge regression as a secondary learner.
Fig. 10 The distribution of coefficients within the level-2 models (ridge regression). MS multi-spectral features, TIR thermal infrared features, SVM support vector machine, DNN deep neural network, RR ridge regression, RF random forest, StRR stacking regression using ridge regression as a secondary learner.
weight within the ensemble procedure. The results showed greater regression coefficients by secondary learner (RR) for individual models with high prediction accuracy. This suggested that the performance of the ensemble model was strongly associated with the high-performance of individual models.

**Effects on accuracy improvement of multi-sensor data fusion and ensemble learning**

Figure 11 shows the comparison of the effects on accuracy improvement of ensemble learning and data fusion based on the best performing individual ML model. For individual sensors, ensemble learning improved prediction accuracy ($R^2$) by up to 5%, while prediction accuracy in multi-sensor fusion improved by up to 31%. The collinearity between the
predicted results of three sensors was modest (VIF < 5) (Fig. 12a). This shows that fusion of multi-sensor data can effectively provide supplementary information. Moreover, the VIF of each individual ML model was high (VIF > 10) (Fig. 12b). This indicates a high degree of collinearity among the predicted values by individual ML models. The results show that there was higher overlapping information between individual ML model predictions as compared to information between different sensors, which indicated different sensor data being able to complement each other.

**Mapping predicted grain yield**

On the test set, 2800 yield prediction values were generated after 80 random divisions of the origin data. The final prediction of the entire experiment was determined by averaging the numerous yield prediction values of each plot. Heat maps (Fig. 13) of yield predictions obtained by StRR and data fusion were shown for the three irrigation treatments,
respectively. The changes in wheat yield under different irrigation treatments can be seen clearly. Grain yield increased with the increase in irrigation water. The t-test results (Fig. S1) demonstrated that the predicted yield values under various irrigation treatments were significantly different, which was consistent with the measured grain yield (Fig. 3). The mapping results indicate that by combining ensemble learning and multi-sensor data fusion, reliable yield prediction values were obtained.

Discussion

Comparison of prediction accuracy between individual sensors

A MS sensor has been successfully applied for assessment of biophysical or biochemical variables in crops (Hassan et al., 2018, 2019; Yang et al., 2020). However, the results in this study showed that the predictions from MS features were not highly accurate (Table 4). It might be because MS VIs readily saturate at high greenness or due to thickness of plants in wheat and rice (Fu et al., 2014; Thenkabail et al., 2000). Texture information extracted from MS images didn’t provide more accurate estimates than VIs for crop parameter evaluation in a previous study (Liu et al., 2019). This is probably the reason that the combination of MS texture and VIs did not yield the desired yield prediction results in this study. Higher accuracy of grain yield prediction was observed using RGB features compared to MS features in all cases (Table 4). It is probably due to the following reasons: first, the spatial resolution of RGB images is higher than that of MS images. Second, MS images cover many mixed pixels of wheat and background (e.g., soil). In addition to the VIs, plant height was also extracted from the RGB images, which has positive effects on yield prediction. In several reports, UAV-based crop height was strongly associated with crop biomass (Han et al., 2019; Li et al., 2020), which are important indicators of crop grain yield. Combining UAV-based crop height with NDVI or normalized difference yellowness index (NDYI) has effectively improved the yield prediction accuracy of rice (Wan et al., 2020).

The use of UAV-based thermal remote sensing in precision agriculture has been shown to be beneficial for assessment of plant growth (Ludovisi et al., 2017; Sheng, et al., 2010). For some ML models, such as Cubist and RR, the predictions based on TIR features are comparable to RGB features. In general, plant water content affects crop yield, and crop canopy temperature is also closely related to water content. This leads to similar changes in crop yield and canopy temperature across irrigation treatments, resulting in high accuracy of yield prediction from TIR features, while TIR features are not subject to the easy saturation that occurs with VIs.

Analysis of multi-sensor data fusion

Data fusion from various sensor combinations (MSI + TIR, RGB + MS, RGB + TIR and RGB + MS + TIR) was used to predict yield and compare with the predictions based on individual sensor data (Table 4). In general, multi-sensor data fusion can yield higher prediction accuracy than individual-sensor data for most ML models. This is because multiple information such as canopy spectral, structural and temperature all contribute to grain yield prediction in unique and complimentary ways (Maimaitijiang et al., 2020). Among the five individual ML algorithms, the RF algorithm performed better with MS + TIR data fusion. The DNN performed best with RGB data, but the RF algorithm did not significantly
underperform the DNN. RF algorithm also performed well in both RGB + TIR and MS + TIR data fusion. The modeling results of RF algorithm were similar to a previous study which reported that RF has higher generalization ability than other algorithms when combining remote sensing data to evaluate plant parameters (Han et al., 2019; Wang et al., 2016). The DNN model also performed best in data fusion of RGB + MS and all three-sensor data. This may be due to the fact that deep learning tends to outperform the most popular ML methods when dealing with larger sample sizes, complex, non-linear and redundant datasets (Kang & Kang, 2017).

To further understand the contribution of each sensor’s information for predicting yield, the relative contribution of each sensor to the modeling performance of fusing all sensors was analyzed as the percent change in the accuracy parameters when corresponding sensor features were excluded from the modeling procedure (Table S1). The results show that in most cases, the accuracy parameters of each ML model varied the most when RGB features were excluded, followed by TIR features. The changes in the accuracy parameters of each ML model were weak when MS features were excluded, especially for RF and StRR. The desired prediction accuracy can be obtained using only RGB and TIR sensors which would contribute to the reduction of sensor costs.

The potential of ensemble learning

Apart from using a variety of sensors to gain a better understanding of plant behavior and response, big data analytics have opened up new possibilities for the creation of innovative methods to extract more information as a result of advance in computing power (Shah et al., 2019). Previous studies have often used an individual ML algorithm to estimate different crop parameters (Matese & Di Gennaro, 2021; Shafiee et al., 2021), but an individual ML algorithm has limitations in parsing different types of data. In precision agriculture management, slight differences in estimation accuracy may lead to different decisions, so it is worthwhile to explore the possibility of obtaining higher prediction accuracy. Ensemble learning results in this study are consistent with previous studies where ensemble models achieved higher prediction accuracy than individual models under various modeling conditions (Feng et al., 2020b; Fu et al., 2019), which demonstrates the reliability of the ensemble approach.

In most studies in estimating plant traits using stacking regression, linear models are often used as secondary learners. Some weight assignment methods, such as decision-level fusion, quadratic programming and Bayesian model averaging (BMA), can also be used as secondary learners for estimating crop parameters, and these methods have been shown to be effective in other fields to improve prediction accuracy in ensemble learning (Wang et al., 2020, 2021; Yin et al., 2021). Ensemble learning requires a complete training of each base model to reach its potential, which inevitably adds a significant amount of model training time compared to the best-performing individual model. How to balance model complexity and accuracy still needs to be explored in future research.

Limitations and implications

For this study, wheat canopy MS, RGB and TIR images were collected rapidly and non-destructively using a UAV platform, and both ensemble learning and multi-sensor data fusion yielded higher prediction accuracy than traditional methods. The more obvious accuracy improvement effect of multi-sensor fusion make it one of the main research
directions for precision agriculture in the future (Fig. 11). Abundant data from widely used sensors such as LiDAR, SARs or crop parameters simulated by crop model (Araya et al., 2010; de Wit et al., 2019), can be fused with the data obtained from the sensors used in this study to enhance the prediction accuracy of crop yield and the stability of the model. The feature extraction process was carried out for each sensor image, and the direct use of raw image information directly as input features for advanced deep learning algorithms has the potential to uncover more potential information that should be explored in the future for multi-sensor image fusion (Sagan et al., 2021).

Thirty varieties were used in this study, which makes the yield data richer and helps to obtain a higher prediction accuracy. However, most of these varieties were from China’s Yellow and Huai Valleys Winter Wheat Zone (YHVWWZ), which may have some regional limitations. Whether the present method can obtain similar results in yield prediction experiments in other wheat regions requires further research.

Conclusion

The present study explored the potential of fusing the data from UAV-based RGB, MS and TIR sensors, as well as of integrating the modeling capabilities of Cubist, SVM, DNN, RR and RF in wheat yield prediction. The results showed that the use of both multi-sensor data fusion and ensemble learning have improved the wheat yield prediction accuracy. The differences in yield under different irrigation treatments was successfully estimated, which is helpful for breeding work to evaluate the performance of each cultivar under different irrigation treatments. To further assess the stability of the proposed method in plant breeding, it should be tested under more developmental stages and growth environments.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11119-022-09938-8.

Funding This research was funded by the Central Public-interest Scientific Institution Basal Research Fund (No.Y2021YJ07), Technology Innovation Program of the Chinese Academy of Agricultural Sciences (CAAS-ZDXT-2019002) and the Key Grant Technology Project of Xinxiang (ZD2020009).

Data availability The datasets generated and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Ahmad, I. S., & Reid, J. F. (1996). Evaluation of colour representations for maize images. Journal of Agricultural Engineering Research, 63, 185–195.

Araya, A., Habtu, S., Hadgu, K. M., Kebede, A., & Dejene, T. (2010). Test of AquaCrop model in simulating biomass and yield of water deficient and irrigated barley (*Hordeum vulgare*). Agricultural Water Management, 97, 1838–1846.

Barnes, E., Clarke, T., Richards, S., Colaizzi, P., Haberland, J., Kostrzewski, M., et al. (2000). Coincident detection of crop water stress, nitrogen status and canopy density using ground based
multispectral data. In *Proceedings of the 5th International Conference on Precision Agriculture* (pp. 16–19). American Society of Agronomy Publishers.

Bendig, J., Yu, K., Aasen, H., Bolten, A., Bennertz, S., Broscheit, J., et al. (2015). Combining UAV-based crop height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. *International Journal of Applied Earth Observation and Geoinformation*, 39, 79–87.

Bhadra, S., Sagan, V., Maimaitijiang, M., Maimaitiyiming, M., Newcomb, M., Shakoor, N., et al. (2020). Quantifying leaf chlorophyll concentration of sorghum from hyperspectral data using derivative calculus and machine learning. *Remote Sensing*, 12, 2082.

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

Broge, N. H., & Leblanc, E. (2001). Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sensing of Environment*, 76, 156–172.

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

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20, 273–297.

Daughtry, C., Walthall, C., Kim, M., De Colstoun, E. B., & McMurtrey III, J. (2000). Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sensing of Environment*, 74, 229–239.

de Wit, A., Boogaard, H., Fumagalli, D., Janssen, S., Knapen, R., van Kraalingen, D., et al. (2019). 25 years of the WOFOST cropping systems model. *Agricultural Systems*, 168, 154–167.

Draper, N. R. (1998). Applied regression analysis. *Technometrics*, 9, 182–183.

Fei, S., Hassan, M. A., He, Z., Chen, Z., Shu, M., Wang, J., et al. (2021). Assessment of ensemble learning to predict wheat grain yield based on UAV-multispectral reflectance. *Remote Sensing*, 13, 2338.

Feng, A., Zhou, J., Vories, E. D., Sudduth, K. A., & Zhang, M. (2020a). Yield estimation in cotton using UAV-based multi-sensor imagery. * Biosystems Engineering*, 193, 101–114.

Feng, L., Zhang, Z., Ma, Y., Du, Q., Williams, P., Drewry, J., et al. (2020b). Alfalfa yield prediction using UAV-Based hyperspectral imagery and ensemble learning. *Remote Sensing*, 12, 2028.

Fu, Y., Yang, G., Wang, J., Song, X., & Feng, H. (2014). Winter wheat biomass estimation based on spectral indices, band depth analysis and partial least squares regression using hyperspectral measurements. *Computers and Electronics in Agriculture*, 100, 51–59.

Fu, P., Meacham-Hensold, K., Guan, K., & Bernacchi, C. J. (2019). Hyperspectral leaf reflectance as proxy for photosynthetic capacities: An ensemble approach based on multiple machine learning algorithms. *Frontiers in Plant Science*, 10, 730.

Gitelson, A. A., Kaufman, Y. J., & Merzlyak, M. N. (1996). Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sensing of Environment*, 58, 289–298.

Gitelson, A. A., Via, A., Arkebauer, T. J., Rundquist, D. C., & Keydan, G. (2003). Remote estimation of leaf area index and green leaf biomass in maize canopies. *Geophysical Research Letters*, 30, 335–343.

Guan, K., Wu, J., Kimball, J. S., Anderson, M. C., Frolking, S., Li, B., et al. (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.

Haboudane, D., Miller, J. R., Tremblay, N., Zarco-Tejada, P. J., & Dextraze, L. (2002). Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sensing of Environment*, 81, 416–426.

Han, L., Yang, G., Dai, H., Xu, B., Yang, H., Feng, H., et al. (2019). Modeling maize above-ground biomass based on machine learning approaches using UAV remote-sensing data. *Plant Methods*, 15, 10.

Haralick, R. M., & Shanmugam, K. (1973). Textural features for image classification. *IEEE Transactions on Systems Man Cybernetics-Systems*. https://doi.org/10.1109/TSMC.1973.4309314

Hassan, M. A., Yang, M., Rasheed, A., Jin, X., Xia, X., Xiao, Y., et al. (2018). Time-series multispectral indices from unmanned aerial vehicle imagery reveal senescence rate in bread wheat. *Remote Sensing*, 10, 809.

Hassan, M. A., Yang, M., Rasheed, A., Yang, G., Reynolds, M., Xia, X., et al. (2019). A rapid monitoring of NDVI across the wheat growth cycle for grain yield prediction using a multi-spectral UAV platform. *Plant Science*, 282, 95–103.

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12, 55–67.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83, 195–213.
Jin, X., Li, Z., Feng, H., Ren, Z., & Li, S. (2020). Deep neural network algorithm for estimating maize biomass based on simulated Sentinel 2A vegetation indices and leaf area index. *The Crop Journal, 8*, 87–97.

Kang, H. W., & Kang, H. B. (2017). Prediction of crime occurrence from multi-modal data using deep learning. *PLoS ONE, 12*, e0176244.

Kawashima, S., & Nakatani, M. (1998). An algorithm for estimating chlorophyll content in leaves using a video camera. *Annals of Botany, 81*, 49–54.

Lee, U., Chang, S., Putra, G. A., Kim, H., & Kim, D. H. (2018). An automated, high-throughput plant phenotyping system using machine learning-based plant segmentation and image analysis. *PLoS ONE, 13*(4), e0196615.

Li, B., Xu, X., Zhang, L., Han, J., Bian, C., Li, G., et al. (2020). Above-ground biomass estimation and yield prediction in potato by using UAV-based RGB and hyperspectral imaging. *ISPRS Journal of Photogrammetry and Remote Sensing, 162*, 161–172.

Liu, Y., Liu, S., Li, J., Guo, X., Wang, S., & Lu, J. (2019). Estimating biomass of winter oilseed rape using vegetation indices and texture metrics derived from UAV multispectral images. *Computers and Electronics in Agriculture, 166*, 105026.

Liu, S., Jin, X., Nie, C., Wang, S., Yu, X., Cheng, M., et al. (2021). Estimating leaf area index using unmanned aerial vehicle data: Shallow vs. deep machine learning algorithms. *Plant Physiology, 187*, 1551–1576.

Ludovisi, R., Tauro, F., Salvari, R., Khoury, S., Mugnozza, G. S., & Harfouche, A. (2017). UAV-based thermal imaging for high-throughput field phenotyping of black poplar response to drought. *Frontiers in Plant Science, 8*, 1681.

Maimaitijiang, M., Ghulam, A., Sidike, P., Hartling, S., Maimaitiyiming, M., Peterson, K., et al. (2017). Unmanned aerial system (UAS)-based phenotyping of soybean using multi-sensor data fusion and extreme learning machine. *ISPRS Journal of Photogrammetry and Remote Sensing, 134*, 43–58.

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.

Marcoulides, K. M., & Raykov, T. (2019). Evaluation of variance inflation factors in regression models using latent variable modeling methods. *Educational and Psychological Measurement, 79*, 874–882.

Matese, A., & Di Gennaro, S. F. (2021). Beyond the traditional NDVI index as a key factor to mainstream the use of UAV in precision viticulture. *Scientific Reports, 11*, 2721.

Merzlyak, M. N., Gitelson, A. A., Chivkunova, O. B., & Rakitin, V. Y. (1999). Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. *Physiologia Plantarum, 106*, 135–141.

Meyer, G. E., & Neto, J. C. (2008). Verification of color vegetation indices for automated crop imaging applications. *Computers and Electronics in Agriculture, 63*, 282–293.

Montesinos-López, O. A., Montesinos-López, A., Crossa, J., Campos, G. D. L., Alvarado, G., Mondal, S., et al. (2017). Predicting grain yield using canopy hyperspectral reflectance in wheat breeding data. *Plant Methods, 13*, 4.

Nigro, D., Giove, S. L., Fortunato, S., Incerti, O., Zacheo, S. A., Blanco, A., et al. (2014). Allelic variation of wheat flour allergens in a collection of wheat genotypes. *Journal of Chemistry, 2014*, 629379.

Pal, M. (2007). Ensemble learning with decision tree for remote sensing classification. In *Proceedings of World Academy of Science Engineering & Technology (pp. 735–737)*. World Academy of Science, Engineering and Technology Publishers.

Penuelas, J., Filella, I., & Gamon, J. A. (1995). Assessment of photosynthetic radiation-use efficiency with spectral reflectance. *New Phytologist, 131*, 291–296.

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.

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.W., Schell, J.A., Deering, D.W., & Harlan, J.C. (1974). Monitoring the vernal advancement and retrogradation (green wave effect) of natural vegetation. *NASA/gsfc type final report*. Greenbelt, USA.
Rutkoski, J., Poland, J., Mondal, S., Autrique, E., Pérez, L. G., Crossa, J., et al. (2016). Canopy temperature and vegetation indices from high-throughput phenotyping improve accuracy of pedigree and genomic selection for grain yield in wheat. *G3-Genes Genomes Genetics, 6*, 2799–2808.

Saberoon, M., Amin, M., Anuar, A., Gholizadeh, A., Wayayok, A., & Khairunniza-Bejo, S. (2014). Assessment of rice leaf chlorophyll content using visible bands at different growth stages at both the leaf and canopy scale. *International Journal of Applied Earth Observation and Geoinformation, 32*, 35–45.

Sagan, V., Maimaitijiang, M., Bhadora, S., Maimaitiyiming, M., Brown, D. R., Sidike, P., et al. (2021). Field-scale crop yield prediction using multi-temporal WorldView-3 and PlanetScope satellite data and deep learning. *ISPRS Journal of Photogrammetry and Remote Sensing, 174*, 265–281.

Schleicher, T. D., Bausch, W. C., Delgado, J. A., & Ayers, P. D. (2001). Evaluation and refinement of the nitrogen reflectance index (NRI) for site-specific fertilizer management. Paper No. 011151. St. Joseph, MI, USA: American Society of Agricultural and Biological Engineers.

Shafiee, S., Lied, L. M., Burud, I., Dieseth, J. A., Alsheikh, M., & Lillemo, M. (2021). Sequential forward selection and support vector regression in comparison to LASSO regression for spring wheat yield prediction based on UAV imagery. *Computers and Electronics in Agriculture, 183*, 06036.

Shah, S. H., Angel, Y., Houborg, R., Ali, S., & McCabe, M. F. (2019). A random forest machine learning approach for the retrieval of leaf chlorophyll content in wheat. *Remote Sensing, 11*, 920.

Sheng, H., Chao, H., Coopmans, C., Han, J., McKee, M., & Chen, Y. (2010). Low-cost UAV-based thermal infrared remote sensing: Platform, calibration and applications. In *Proceedings of the 2010 IEEE/ASME international conference on mechatronic and embedded systems and applications* (pp. 38–43). IEEE Publishers.

Thenkabail, P. S., Smith, R. B., & De Pauw, E. (2000). Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. *Remote Sensing of Environment, 71*, 158–182.

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment, 8*, 127–150.

Wan, L., Cen, H., Zhu, J., Zhang, J., Zhu, Y., Sun, D., et al. (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, 1291*, 108096.

Wang, L., Zhou, X., Zhu, X., Dong, Z., & Guo, W. (2016). Estimation of biomass in wheat using random forest regression algorithm and remote sensing data. *The Crop Journal, 4*, 212–219.

Yang, M., Hassan, M. A., Xu, K., Zheng, C., Rasheed, A., Zhang, Y., et al. (2020). Assessment of water and nitrogen use efficiencies through UAV-based multispectral phenotyping in winter wheat. *Frontiers in Plant Science, 11*, 927.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
Authors and Affiliations

Shuaipeng Fei\textsuperscript{1} · Muhammad Adeel Hassan\textsuperscript{2,5} · Yonggui Xiao\textsuperscript{2} · Xin Su\textsuperscript{3} · Zhen Chen\textsuperscript{1} · Qian Cheng\textsuperscript{1} · Fuyi Duan\textsuperscript{1} · Riqiang Chen\textsuperscript{6} · Yuntao Ma\textsuperscript{4}

\textsuperscript{1} Institute of Farmland Irrigation, Chinese Academy of Agricultural Sciences, Xinxiang 453002, China
\textsuperscript{2} National Wheat Improvement Centre, Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, China
\textsuperscript{3} Water Diversion and Irrigation Engineering Technology Center, Yellow River Institute of Hydraulic Research, Zhengzhou 450003, China
\textsuperscript{4} College of Land Science and Technology, China Agricultural University, Beijing 100193, China
\textsuperscript{5} Dezhou Academy of Agricultural Sciences, Dezhou 253050, China
\textsuperscript{6} School of Information Science and Technology, Beijing Forestry University, Beijing 100083, China