



ARTICLE

Using Chlorophyll Content and Spectral Vegetation Indices for Modeling Wheat Yield Grown under Arid Conditions in Saudi Arabia

Khalid A. Al-Gaadi^{1,2}, Rangaswamy Madugundu¹, Ahmed M. Zeyada^{2,*}, Elkamil Tola¹
and Haroon F. Edrees¹

¹Precision Agriculture Research Chair, Deanship of Scientific Research, King Saud University, Riyadh, Saudi Arabia

²Department of Agricultural Engineering, College of Food and Agriculture Sciences, King Saud University, Riyadh, Saudi Arabia

*Corresponding Author: Ahmed M. Zeyada. Email: azeyada@ksu.edu.sa

Received: 24 June 2026; Accepted: 29 September 2026; Published: 19 October 2026

ABSTRACT: Early prediction of crop yield is a key element in understanding the productivity of agricultural fields. In this context, integrating spectroscopy and plant biophysical parameters with multivariate statistical methods can improve crop growth monitoring and yield prediction, contributing to more effective agricultural management. Therefore, this study was conducted to investigate the effectiveness of leaf chlorophyll content (SPAD readings) and spectral vegetation indices (VIs) in developing yield prediction models at early growth stages of wheat with reference to actual harvested yield. VIs such as the normalized difference vegetation index (NDVI), enhanced vegetation index (EVI), soil-adjusted vegetation index (SAVI), and the normalized difference water index (NDWI) were computed at different crop growth stages. Simple and multiple regression analyses were conducted to assess the relationships among the generated VIs and Y_A for the development of predictive models for early wheat yield estimation. In contrast, the heading stage exhibited a strong and highly significant relationship between wheat yield and SPAD readings in all fields, with R^2 values of 0.96 (field 71), 0.65 (field 77), and 0.92 (field 81), along with the lowest RMSE values and $p \leq 0.001$, demonstrating high predictive accuracy. Using SPAD readings, heading stage provided the best prediction accuracy, with the lowest MBE (0.32 t ha^{-1}) and RMSE (1.59 t ha^{-1}), while the highest R^2 (0.83) occurred at the jointing stage. Using vegetation indices, the booting stage showed the most consistent performance (mean $R^2 = 0.91$; RMSE = 0.34 t ha^{-1}). Overall, combining SPAD readings and vegetation indices, the heading stage provided the best yield prediction performance (mean $R^2 = 0.97$; RMSE = 0.35 t ha^{-1}). Overall, the findings demonstrate that integrating SPAD readings with vegetation indices improves wheat yield prediction accuracy, particularly when using multi-variable models. Among the indices, EVI and SAVI performed better under dense canopy conditions, while NDWI contributed by capturing crop water status. Overall, the results showed that multivariate approaches provide more stable and accurate wheat yield predictions, with optimal performance varying across growth stages.

KEYWORDS: Spectroscopy; yield prediction; multivariate analysis; arid climates

1 Introduction

Wheat (*Triticum aestivum*) is one of the most widely cultivated cereal crops in the world because it is a staple food and the most important food security crop, contributing to global food security and providing a substantial proportion of dietary calories and protein worldwide [1]. Sustainable wheat production can be achieved through the optimal application of crop inputs and the adoption of precision farming practices, which manage different areas of the field according to spatially and temporally varying growing conditions to maximize production potential [2]. Marino [3] also pointed out that remote sensing methods, which use

multispectral satellite imagery and appropriate spectral indices to monitor spatial and temporal variability of crop production traits, are the basis for the proper adoption of precision agriculture. In general, studying spatial variability in agricultural fields is of great importance in precision agriculture, as it is based on the principle that the observations of a spatial variable are correlated over a certain distance and influence the values at nearby unsampled locations that are to be predicted [4].

Variations in crop growth are usually caused by site-specific climate and farm management practices and may arise from sub-field heterogeneities in soil composition, moisture levels, sunlight, and diseases [5]. The cost and time required to obtain crop and soil information to create homogeneous agricultural areas are barriers to progress in precision agriculture practices. As an alternative to traditional methods of measuring biomass, which is important for monitoring plant growth and health, remote sensing techniques provide non-destructive and flexible approaches for measuring useful crop parameters for monitoring crop growth and predicting yields [6,7]. In this context, optical sensing of crop characteristics has proven to be an ideal alternative to serve this task [8]. Remote sensing systems have demonstrated the ability to identify field changes in soil and crops and provide valuable information for site-specific management practices, where hyperspectral imagery can capture accurate spectral responses and, therefore, detect spatial differences in land cover and its changes over time [9,10]. Reflectance in the visible and near-infrared regions has been used to indirectly estimate plant leaf nitrogen concentration because of the strong positive correlation with leaf chlorophyll content and the apparent sensitivity to canopy structures [11]. Furthermore, spectral vegetation indices, used successfully as effective measures of vegetation growth and health, have been used to assess spatial variability in agricultural fields by describing differences in crop canopy characteristics [12]. In addition, Barbanti et al. [2] reported that the assessment of the spatial variability of soil and crop properties provides a fundamental basis for precision agriculture applications. In this context, proximal and remote sensing-based spectral vegetation indices have emerged as effective alternatives for assessing soil and crop properties. Vegetation indices (VIs) enhance the ability and sensitivity to detect green vegetation and have been widely used to characterize crop growth and estimate important crop parameters, including biomass, yield, photosynthetic activity, and leaf area index [13]. However, vegetation indices have been widely used for various applications, including yield prediction [14] and drought monitoring and assessment [15]. They have also been widely applied to assess vegetation health and moisture [16], drought stress evaluation [17], and evapotranspiration mapping [18]. Dash et al. [19] integrated UAV-based thermal and multispectral imagery to evaluate crop water conditions at the field scale. Their findings indicated a strong relationship between UAV-derived land surface temperature and ground-based measurements, highlighting the potential of combining thermal observations with soil moisture data for detecting crop water stress. Similarly, Wang et al. [20] developed a water-stress assessment approach for winter wheat using UAV multispectral and thermal imagery. By integrating vegetation and thermal indices with machine-learning techniques, the study achieved improved estimation of crop water status, demonstrating the potential of UAV remote sensing for monitoring water stress in precision agriculture. The normalized difference vegetation index (NDVI) is considered an effective tool for predicting crop yield [21]. Royo et al. [22] reported a positive correlation between NDVI and grain yield at the flowering stage. The mid-grain filling (milky grain) stage was found to be the most suitable growth stage for yield estimation. Moreover, several studies have demonstrated a significant association between NDVI and grain yield from the heading stage to grain filling in wheat [22,23].

On the other hand, the timing of captured reflectance is one of the key variables that can be adjusted within the model, making it an important factor in crop-yield forecasting. Leukel et al. [24] studied the

timing of satellite observations in various field-scale models, emphasizing the concept of the “prediction horizon,” defined as the period between yield forecasting and harvest.

The study reported substantial differences in prediction horizons, even for the same crops, such as wheat and corn, and within similar geographical regions, including the USA, China, and Australia, with forecasts ranging from several weeks to several months before harvest. While, Fieuzal et al. [25] reported little variation in prediction accuracy across crop growth stages, indicating that satellite images acquired after mid-season may provide limited additional benefits for yield prediction. Overall, these findings indicate that the optimal timing for yield prediction varies according to crop type, modeling approach, and site-specific conditions.

Accurate estimation and early prediction of crop yield are essential for efficient agricultural resource management, food security, and sustainable agricultural planning [26]. The integration of remote sensing data with crop growth models has demonstrated considerable potential for improving the timely and reliable prediction of crop yield. Moreover, accurate information on crop productivity supports decision-making at both farm and national levels by enabling better planning of agricultural inputs and food supplies. In this context, remote sensing provides an effective approach for monitoring crop growth and production over large areas and at different temporal scales [27]. Real-time monitoring and reliable forecasting of crop growth and yield are particularly important for developing timely agricultural strategies and ensuring food security [28]. Such forecasts can support decisions related to food policy, market and price management, import and export planning, and the allocation and storage of food supplies [29].

Crop prediction has proven to be a valuable tool in assisting decision-makers to improve yield and enhance profitability, as early information on possible crop failures and potential yield reductions can help in decision-making aimed at mitigating consequences of serious production shortfalls. In addition, yield modeling not only allows predicting plant production, but also contributes to understanding how crop yields are affected by environmental factors and yield components [30]. However, there are few existing systems capable of accurately predicting yields at field or within-field scales [31]. On the other hand, modern agriculture relies on the control of in-field variability resulting from the interactions of numerous factors such as soil, climate, and crops [32].

Machine learning (ML) is an important technology for developing crop yield prediction models, and predictive modeling using ML is growing rapidly and includes early predictions, pre-harvest predictions, field- or area-level predictions, and predictions for different crop types [24]. Traditional statistical models often struggle to capture nonlinear relationships between crop growth and environmental factors, which makes machine-learning approaches such as Random Forest (RF) and Convolutional Neural Networks (CNN) more effective. Multivariate ML algorithms have been used in agricultural yield prediction and modeling and have proven their capability to improve decision-making and provide reliable results [33]. Examples of ML algorithms integrated with optical remote sensing that have been successfully used in crop growth monitoring and yield estimation were provided by many researchers [34]. These include support vector machine (SVM), artificial neural network (ANN), classification and regression tree (CART), Gaussian process regression (GPR), random forest (RF), decision tree (DT), partial least squares regression (PLSR), least squares linear regression (LSLR), and multilinear regression (MLR). Furthermore, RF is particularly suitable for remote-sensing applications because it effectively handles nonlinear relationships and high-dimensional spectral data, whereas CNNs can automatically extract complex features and patterns from hyperspectral observations. Zhang et al. [35] successfully demonstrated the use of RF and CNN models for estimating winter wheat chlorophyll content and improving prediction accuracy through multimodal data integration. Moreover, vegetation indices such as the Normalized Difference Vegetation Index (NDVI)

and Enhanced Vegetation Index (EVI) are closely associated with crop vigor, biomass accumulation, and yield potential, while moisture-related indices such as the Normalized Difference Water Index (NDWI) provide valuable information on crop water status and drought stress [36]. Therefore, this study aimed to develop and evaluate wheat yield prediction models at different growth stages using leaf chlorophyll measurements (SPAD) and vegetation indices derived from FieldSpec-3 spectral data. Model performance was assessed using independent field datasets to identify the most accurate and robust approach for wheat yield prediction.

2 Materials and Methods

2.1 Study Area

The experimental work was conducted on three winter wheat fields belonging to the Saudi Agricultural Development Company (INMA), located in Wadi Al-Dawasir, the southwestern region of Saudi Arabia, between the latitudes of 19.90° and 20.33° N and the longitudes of 44.81° and 44.95° E (Fig. 1). INMA is a commercial farm established in 1982. Currently, 90 fields are active for cultivation of wheat, barley, alfalfa, and potatoes. The area is characterized by an arid desert climate, with hot summers exceeding 45°C , mild winters ranging from 5 to 25°C , and annual rainfall of less than 100 mm. Water for agricultural activities is extracted from deep aquifers and supplied to crops through drip and center-pivot irrigation systems. The soil is predominantly sandy to loamy sand, with low water-holding capacity, high permeability, and low organic carbon content.

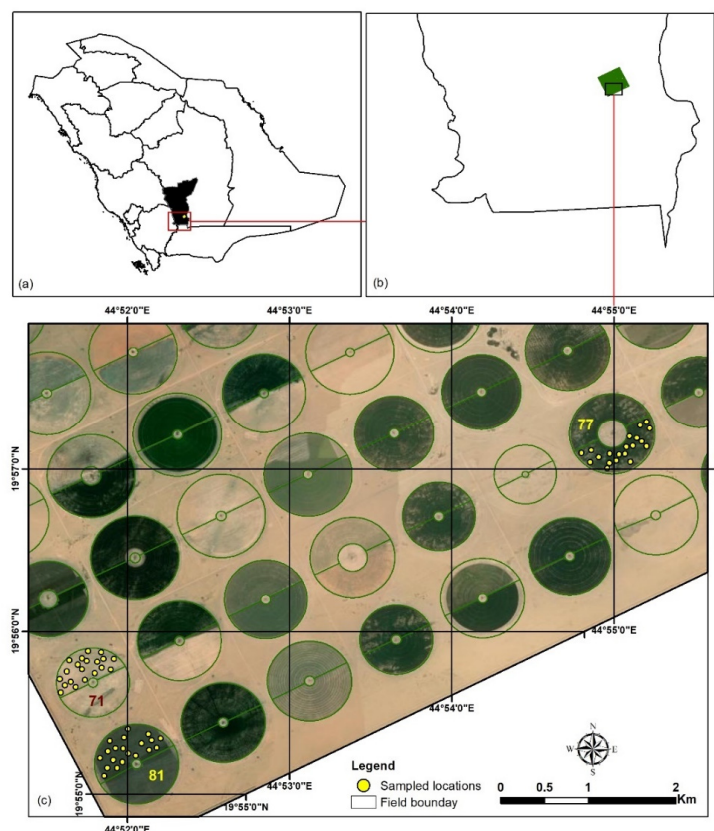


Figure 1: Map of study area and experimental fields: (a) Saudi Arabia; (b) Wadi Al-Dawasir, (c) Experimental farm and fields overlaid with GPS sampled locations.

2.2 Experimental Fields and Crop Details

Of the 90 agricultural fields, three wheat fields (IDs 71, 77, and 81) were earmarked for this study. Approximately 50% of the area of each field, corresponding to an area of about 27 ha, was allocated for sampling activities (Fig. 1). A total of 60 permanent sampling plots (20 for each experimental pivot), corresponding to approximately one sampling point per 1.35 ha, were randomly established and georeferenced using a GeoXH 6000 portable GNSS receiver (Trimble Inc., Westminster, CO, USA). Sampling locations (1 m × 1 m square) were strategically distributed across the field as fixed plots to capture SPAD and hyperspectral measurements throughout the crop growth cycle, enabling consistent assessment of temporal variations in wheat physiological and spectral characteristics. Experimental fields were cultivated from November 2021 to March 2022 with a density of 350 plants m⁻² with spring wheat (*Triticum aestivum* L., cv. Yecora Rojo). Yecora Rojo is a hard red spring wheat cultivar. Seeds were imported by experimental farm INMA, Wadi Al-Dawasir, Saudi Arabia. During land preparation, superphosphate (16% P₂O₅) and potassium sulfate (42% K₂O) were applied at rates of 70 kg ha⁻¹ each. Nitrogen was applied at the rate of 100 kg N ha⁻¹ as ammonium nitrate (33.3% N), split into three equal doses at sowing, tillering, and anthesis. Irrigation was scheduled according to the standard management practices adopted at the INMA Farm, Wadi Al-Dawasir, following the guidelines of the Ministry of Environment, Water and Agriculture, Saudi Arabia. To ensure that spectral and physiological measurements were solely attributable to wheat, sampling was restricted to uniformly managed mono-cropped wheat fields. Experimental plots were free from intercropping systems, mixed crop stands, and significant weed infestation, thereby minimizing interference from non-target vegetation during SPAD readings, hyperspectral data acquisition, and yield assessment. Table 1 summarizes the experimental timeline, including sowing, spectral data collection dates, growth stages, and harvesting.

Table 1: Schedule of spectral data collection, growth stages, and harvesting.

Parameters	Date	Days after Sowing (DAS)	Growth Stage (Feekes)	Description of Growth Stages
Sowing	13 November 2021	1	-	Sowing
Spectral data collection (SPAD and Field Spec-3)	11 December 2021	31	3.0	Tillering
	05 January 2022	56	6.0	Jointing
	25 January 2022	76	9.0	Booting
	14 February 2022	96	10.5	Heading
Harvesting	22 March 2022	138	11.4	Maturity

2.3 SPAD Readings

Leaf chlorophyll content (SPAD readings) was measured using a portable SPAD meter (SPAD-502 Plus Chlorophyll Meter, Spectrum Technologies Inc., USA). The instrument provides an indexed measurement of leaf chlorophyll content by clamping the meter over the leaf tissue non-destructively. Measurements were carried out under clear sky conditions between 09:00 and 11:00 h to minimize the effects of diurnal variation on chlorophyll readings. For each experimental plot, ten representative wheat plants were randomly selected from within the 1 m × 1 m square designated for yield determination. Care was taken to avoid border plants to reduce edge effects and ensure that sampled plants adequately represented the crop stand. Chlorophyll measurements were taken from the flag leaf after heading, or from the uppermost fully expanded leaf prior to heading, as these leaves contribute substantially to photosynthesis and grain filling. Three SPAD readings were recorded along each leaf blade, specifically from the basal, middle, and apical portions while avoiding the midrib and any visibly damaged, diseased, or senescent tissue. The three readings obtained

from each leaf were averaged to represent the chlorophyll status of an individual plant. Consequently, a total of 30 observations (10 plants $\times$ 3 readings) were collected from each quadrat. The mean value of these observations was calculated and used as the representative SPAD value for the respective plot (i.e., sampled location) in subsequent statistical analyses. The same procedure was followed for SPAD readings across the growth stages.

2.4 Spectral Reflectance and Vegetation Indices

Spectral data of wheat plants were acquired using a FieldSpec-3 spectroradiometer (Analytical Spectral Devices Inc., Longmont, Colorado, USA), an advanced hyperspectral sensor capable of capturing reflectance data in the spectral range of 350–2500 nm. For each sampling event, measurements were taken from fully expanded upper canopy leaves, and multiple spectra were recorded from different plants and leaf positions to capture within-plot variability. Prior to each measurement session, the FieldSpec-3 spectroradiometer was calibrated using a certified white reference panel to ensure accurate reflectance measurements. White reference calibration was performed at regular intervals during data collection to minimize the effects of changing illumination conditions, sensor drift, and environmental variability.

To ensure consistency in acquiring spectral reflectance, a sampling protocol similar to that used for SPAD readings was followed. The recorded spectrum from individual samples was averaged to obtain a representative spectral signature for each plot (i.e., sampled location).

Prior to analysis, the collected spectral reflectance dataset was first divided into training (65%) and validation (35%) subsets. To prevent information leakage during model development and validation, normalization parameters and outlier detection thresholds were derived exclusively from the training data and subsequently applied to both subsets. Spectral reflectance values were normalized using the min-max scaling method, which transformed the data to a standardized range between 0 and 1. Outliers were identified using the Z-score method to minimize the influence of extreme values resulting from sensor anomalies, measurement errors, or environmental variability on model performance. Spectral observations with absolute Z-score values exceeding 3 ($|Z| > 3$) were classified as outliers and removed from the dataset. This threshold was selected because observations falling beyond three standard deviations from the mean are generally considered statistically anomalous. Application of this criterion resulted in the removal of approximately 6% of the collected spectral observations, while 94% of the measurements were retained for subsequent analyses.

Vegetation indices (VIs), calculated from specific spectral bands, can provide useful information on vegetation characteristics such as chlorophyll content, biomass, and plant vigor, and are widely used as effective indicators for yield prediction. They are designed to capture various physiological and biochemical characteristics of plants. Hyperspectral reflectance data acquired using a FieldSpec-3 spectroradiometer were utilized to derive four vegetation indices, namely the Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), Soil Adjusted Vegetation Index (SAVI), and Normalized Difference Water Index (NDWI), to characterize wheat canopy vigor, biomass development, and crop water status. Among these, NDWI was calculated using near-infrared (NIR) and 1240 nm water-sensitive NIR reflectance to enhance sensitivity to vegetation water content under the hyper-arid environmental conditions of the study area (Table 2). Furthermore, the classification ranges given in Table 3 were used as thresholds for crop health indication.

Table 2: List of vegetation indices used in the study. L is soil brightness correction factor.

Vegetation Index	Formula	Description	Reference
NDVI	$NDVI = \frac{R_{860} - R_{660}}{R_{860} + R_{660}}$	Assessing vegetation health and biomass	[37] Rouse et al. (1974)
EVI	$EVI = \frac{2.5 \times (R_{860} - R_{660})}{(R_{860} + 6 \times R_{660} - 7.5 \times R_{470} + 1)}$	Improve vegetation monitoring by reducing atmospheric and soil background influences	[38] Huete et al. (2002)
SAVI	$SAVI = (1 + L) \times \frac{(R_{860} - R_{660})}{(R_{860} + R_{660} + L)}$	Minimize soil influence in areas with low vegetation cover	[39] Huete (1988)
NDWI	$NDWI = \frac{R_{860} - R_{1240}}{R_{860} + R_{1240}}$	Monitor water content in vegetation	[40] Gao (1996)

Table 3: List of vegetation indices and their thresholds used for health indication.

Category	SPAD	NDVI	EVI	SAVI	NDWI	Crop Condition
Very Low	<25	<0.20	<0.10	<0.15	<-0.10	Severe stress/poor establishment/low chlorophyll and moisture
Low	25–35	0.20–0.40	0.10–0.25	0.15–0.35	-0.10 to 0.05	Water/nutrient stress, sparse canopy
Medium	35–45	0.40–0.60	0.25–0.45	0.35–0.55	0.05–0.20	Moderate crop vigor
High	45–55	0.60–0.80	0.45–0.65	0.55–0.75	0.20–0.40	Healthy crop with good biomass
Very High	>55	>0.80	>0.65	>0.75	>0.40	Very vigorous canopy, high yield potential

2.5 Harvested Wheat Grain Yield (Y_A)

Harvested wheat grain yield, referred to as actual yield, was determined by harvesting a one-square-meter area (1 m × 1 m) from each designated sampling plot in three experimental fields. The harvested samples were manually collected, labelled separately, weighed, and the fresh yield was recorded. Care was taken to minimize grain losses during harvesting and handling. The harvested plants were threshed to separate grains from the straw, and the obtained grain samples were cleaned to remove impurities and unfilled kernels. Subsequently, they were oven-dried to a moisture content of 13.5%, and the total grains obtained from each 1 m² square were weighed using a digital balance and recorded as dry weight (kg m⁻²). Subsequently expressed on an area basis and converted to the standard unit, tonnes per hectare (t ha⁻¹).

2.6 Yield Prediction (Y_P) Models

A Schematic representation of the methodological framework for prediction and yield estimation using chlorophyll content, hyperspectral data, and machine learning techniques is provided in Fig. 2. Wheat grain yield prediction was performed using both conventional statistical approaches and machine learning (ML) techniques. The collected data, such as chlorophyll content (SPAD values) and vegetation indices (VIs) at different wheat growth stages, were used as independent variables. These stages included tillering (Feekes 3), jointing (Feekes 6), booting (Feekes 9), and heading (Feekes 10.5). Grain yield, referred to as actual yield (Y_A), was considered the dependent variable and was analyzed using multivariate regression techniques. Before performing regression analysis, the obtained dataset (60 samples) was randomly divided into two subsets of 39 samples (65% of the entire dataset) and 21 samples (35% of the entire dataset) for model development and cross-validation, respectively.

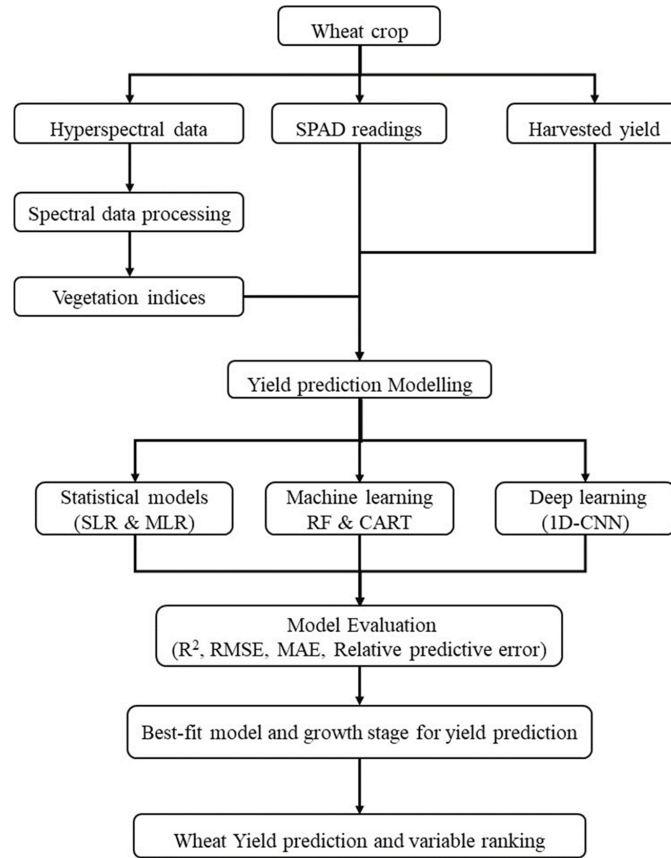


Figure 2: Schematic representation of the methodological framework for prediction and yield estimation using chlorophyll content, hyperspectral data and machine learning techniques.

2.6.1 Simple and Multiple Linear Regression Analysis for Wheat Yield Prediction

Simple linear regression (SLR) was employed to evaluate the relationship between grain yield and individual predictor variables (SPAD, NDVI, EVI, SAVI, and NDWI). Subsequently, Multiple linear regression (MLR) was used to evaluate the combined effects of SPAD readings and vegetation indices on grain yield prediction. The models were expressed in Eqs. (1) and (2), respectively.

$$Y_P = \beta_0 + \beta_1 X + \varepsilon \quad (1)$$

$$Y_P = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n \quad (2)$$

where Y_P represents predicted grain yield, X is the predictor variable, β_0 is the intercept, β_1 is the regression coefficient, and ε is the error term. Model performance was assessed using the coefficient of determination (R^2), root mean square error (RMSE), and significance level ($p < 0.05$).

2.6.2 Machine Learning Models for Wheat Yield Prediction

Machine learning models (Table A1), such as Random Forest (RF), Classification and Regression Trees (CART), and Convolutional Neural Networks (CNN), were applied to predict wheat yield. RF algorithm was employed as described in Breiman [41]. Whereas CART analysis was performed as stated in Carrizosa et al. [42]. On the other hand, CNNs, a deep learning methodology, were used to develop a crop productivity

prediction model based on the NDVI index and RGB color data derived from images taken by unmanned aerial vehicles (UAVs), as reported in the study by Nevavuori et al. [43]. During the execution, the dataset consisting of 60 observations was partitioned into training (65%; $n = 39$) and validation (35%; $n = 21$) subsets using a fixed random seed. The training dataset was used for model development, whereas the validation dataset was used for independent performance assessment. CART, RF, and CNN models were trained using SPAD, NDVI, EVI, SAVI, and NDWI as predictor variables and grain yield as the response variable. Model performance was evaluated using the coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE). Comparison between training and validation statistics was used to assess model robustness and potential overfitting.

Random Forest (RF)

RF is an ensemble learning method that constructs multiple decision trees and improves prediction accuracy by averaging their outputs in regression tasks. It also provides estimates of feature importance, highlighting the contribution of vegetation indices to yield prediction [30]. The dataset, consisting of SPAD readings, vegetation indices, and grain yield measurements, was divided into training and testing subsets using the bagging approach. Subsequently, multiple decision trees were trained using the selected features to construct the Random Forest model. In regression analysis, the final output was derived by averaging the predictions of all individual trees. During model development, bootstrap samples were repeatedly generated from the training dataset using a fixed random seed, and an individual regression tree was constructed for each bootstrap sample. At every node, a random subset of predictor variables was considered for splitting, thereby reducing the correlation among trees and improving model robustness. The RF model was configured using 500 decision trees, while the number of variables considered at each split was optimized during model development. Individual trees were grown without pruning, and final grain-yield predictions were calculated as the average of all tree predictions. Model stability and prediction error were evaluated using out-of-bag (OOB) observations, which were not included in the bootstrap samples used for training. Subsequently, variable importance was assessed using the mean decrease in accuracy and mean decrease in node impurity, enabling identification of the most influential predictors of wheat yield.

Classification and Regression Trees (CART)

CART is a non-linear decision-tree method that partitions data into homogeneous subsets based on spectral features. The CART algorithm was employed to model nonlinear relationships between wheat grain yield and predictor variables, including SPAD, NDVI, EVI, SAVI, and NDWI. CART recursively partitions the dataset into increasingly homogeneous subsets using binary splitting rules. At each node, all predictor variables and candidate thresholds were evaluated, and the split minimizing the residual sum of squares (RSS) was selected. For example, the algorithm may initially partition the dataset using a threshold such as $\text{SPAD} > 45$, thereby separating plots with higher chlorophyll content from those with lower chlorophyll content. Subsequent splits may be based on variables such as $\text{NDWI} > 0.20$ or $\text{NDVI} > 0.70$, resulting in progressively more homogeneous yield groups. These threshold values are data-driven and are automatically determined during model training. The recursive partitioning process continues until predefined stopping criteria, including minimum node size and maximum tree depth, are reached. To reduce overfitting, tree pruning was performed using cross-validation, and the optimal tree structure was selected based on the minimum prediction error. The final CART model generated interpretable decision rules and identified the relative importance of predictor variables for wheat-yield estimation.

Convolutional Neural Networks (CNN)

A one-dimensional CNN was developed to extract spectral patterns from hyperspectral imagery, enabling chlorophyll content analysis and yield prediction through deep learning. Despite their computational intensity, CNNs improve the accuracy of spectral data interpretation. Initially, hyperspectral measurements are structured into multidimensional matrices representing spectral reflectance across different bands. Convolutional filters (kernels) are then applied through convolutional layers to extract spatial-spectral features at multiple scales, enhancing feature representation. The input layer consisted of five features (SPAD, NDVI, EVI, SAVI, NDWI) per sample.

The 1D-CNN architecture comprised a one-dimensional convolutional layer with 32 filters and a kernel size of 2, followed by a Rectified Linear Unit (ReLU) activation function and a max-pooling layer to reduce dimensionality while retaining important feature information. The extracted features were flattened and passed to a fully connected dense layer containing 64 neurons, followed by a dropout layer (dropout rate = 0.20) to reduce overfitting. A final output neuron generated the predicted grain yield value. The model parameters were optimized using the Adam optimizer with a learning rate of 0.001 and a maximum of 100 training epochs. Early stopping was implemented to prevent overfitting and improve model generalization. Model performance was evaluated using the validation dataset based on the coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE). Although 1D-CNN models are commonly applied to high-dimensional hyperspectral datasets, the present study employed the 1D-CNN architecture to investigate whether nonlinear interactions among SPAD readings and vegetation-index variables could improve wheat yield prediction relative to conventional regression and tree-based machine-learning approaches. The inclusion of the 1D-CNN enabled a direct comparison of deep-learning, machine-learning, and statistical methods for wheat yield prediction under hyper-arid conditions.

2.7 Model Evaluation

Developed yield prediction models' robustness was further evaluated using cross-validation procedures, and the average performance metrics were reported. Selected statistical performance indicators included the R^2 , mean bias error (MBE), and root mean square error (RMSE). Furthermore, the predicted wheat yield was evaluated using a split-plot analysis, where wheat fields and growth stages were treated as "main-plot" and "sub-plot" factors, respectively. The main effects of field and growth stage, as well as their interaction, were assessed to compare predicted yields (Y_P) and to identify the stages associated with the highest and lowest predicted yields.

3 Results and Discussion

Table 4 presents the descriptive statistics for actual wheat yield (Y_A), chlorophyll content (SPAD readings), and VIs measured across the growth stages. The actual wheat yields ($t\ ha^{-1}$) for the three experimental fields (IDs: 71, 77, and 81) were relatively similar, with mean values of $7.38 (\pm 1.89)$, $8.03 (\pm 1.33)$, and $7.32 (\pm 1.90) t\ ha^{-1}$, respectively. The mean SPAD readings for the three fields ranged between 30.51 (tillering stage) and 60.93 (heading stage), and an increasing trend in crop development was observed across the growth stages in the three fields studied. Results of the greenness-related VIS (EVI, NDVI, SAVI) showed low levels at the tillering stage, when the crop was still very small, and then gradually increased, reaching peak values at the jointing stage, due to maximum canopy cover and plant turgidity. The NDWI (0.45) showed a high level at the tillering stage.

Table 4: Descriptive statistics (Mean and Standard Deviation “SD”) for the studied parameters at different fields and crop growth stages.

Parameter	Growth Stage	Field ID: 71		Field ID: 77		Field ID: 81		Overall	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Yield (t ha ⁻¹)	Maturity	7.38	1.89	8.03	1.33	7.32	1.90	7.57	2.17
SPAD values	Tillering	31.39	3.25	30.57	3.27	29.56	5.28	30.51	4.05
	Jointing	41.42	5.91	41.21	4.40	41.01	4.59	41.21	4.93
	Booting	52.37	5.56	50.12	5.36	49.86	5.58	50.78	5.52
	Heading	61.65	4.28	60.78	4.36	60.36	4.87	60.93	4.46
EVI	Tillering	0.76	0.14	0.53	0.11	0.29	0.06	0.52	0.22
	Jointing	0.91	0.09	0.79	0.07	0.85	0.07	0.85	0.09
	Booting	0.65	0.09	0.66	0.03	0.79	0.03	0.70	0.09
	Heading	0.64	0.09	0.64	0.05	0.81	0.02	0.70	0.10
NDVI	Tillering	0.67	0.10	0.52	0.10	0.32	0.06	0.50	0.17
	Jointing	0.73	0.05	0.68	0.04	0.70	0.04	0.70	0.05
	Booting	0.58	0.07	0.59	0.02	0.65	0.02	0.61	0.05
	Heading	0.61	0.07	0.63	0.03	0.71	0.02	0.65	0.06
SAVI	Tillering	0.50	0.08	0.37	0.06	0.23	0.04	0.37	0.13
	Jointing	0.58	0.04	0.52	0.04	0.55	0.04	0.55	0.05
	Booting	0.45	0.05	0.45	0.02	0.53	0.02	0.48	0.05
	Heading	0.44	0.05	0.45	0.03	0.54	0.01	0.48	0.06
NDWI	Tillering	0.61	0.10	0.45	0.11	0.29	0.05	0.45	0.16
	Jointing	0.16	0.01	0.14	0.02	0.12	0.02	0.14	0.02
	Booting	0.17	0.02	0.15	0.01	0.14	0.01	0.15	0.02
	Heading	0.15	0.02	0.16	0.02	0.18	0.01	0.16	0.02

3.1 Simple Linear Regression Analysis between Actual Wheat Yield (Y_A) and SPAD Readings

The regression analysis between actual wheat grain yield (Y_A) and chlorophyll content (SPAD readings) across wheat growth stages (Table 5) showed a high correlation at the heading stage ($R^2 = 0.71$), followed by the booting stage ($R^2 = 0.68$). However, moderate ($R^2 = 0.55$) and weak ($R^2 = 0.35$) relationships were observed at the jointing and tillering stages, respectively. Early yield prediction (Y_P) models were developed and cross-validated. The lowest RMSE (1.59 t ha⁻¹; 17.41%) and a very low MBE (0.32 t ha⁻¹; 3.50%) were obtained at the heading stage. After that, the booting stage (RMSE = 2.02 t ha⁻¹) was found to be suitable for early yield prediction using SPAD values alone.

Table 5: Regression analysis for early prediction of wheat yield (Y_P) using the SPAD values.

Growth Stage	Prediction Model	R ²	p-Value			
Tillering	Y _p = 0.3363 * SPAD − 13.5933	0.35	5.73E−11			
Jointing	Y _p = 0.2102 * SPAD − 4.0726	0.55	8.85E−11			
Booting	Y _p = 0.3606 * SPAD − 17.3015	0.68	1.47E−12			
Heading	Y _p = 0.3431 * SPAD − 16.0427	0.71	2.40E−17			
Cross-validation						
Growth Stage	R ²	p-Value	MBE (t ha ^{−1})	RMSE (t ha ^{−1})	MBE (%)	RMSE (%)
Tillering	0.72	3.00E−07	1.67	2.99	18.20	32.67
Jointing	0.83	2.15E−09	1.82	2.89	19.86	31.60
Booting	0.69	6.95E−07	1.17	2.02	12.82	22.07
Heading	0.66	2.22E−06	0.32	1.59	3.50	17.41

Note: R^2 : coefficient of determination; RMSE: root mean square error; MBE mean bias error.

3.2 Simple Linear Regression Analysis between Actual Wheat Yield (Y_A) and Vegetation Indices

Table 6 shows the regression models developed using selected vegetation indices that showed strong potential for predicting wheat yield across growth stages and fields. The coefficient of determination (R^2) values ranged from 0.69 to 0.97, while RMSE values varied between 0.34 t ha^{-1} and 1.20 t ha^{-1} , indicating generally good model performance, although the strength of the relationships varied among growth stages and fields.

At the tillering stage, EVI showed strong predictive performance in fields 71 and 81, with R^2 values of 0.94 and 0.95, respectively. In field 77, NDVI provided a lower predictive performance ($R^2 = 0.82$), suggesting that the relationship between vegetation indices and yield varied among fields at this early stage. During the jointing stage, SAVI was identified as the most suitable predictor across all fields. However, model performance was lower than that observed at other growth stages, with R^2 values ranging from 0.69 to 0.84 and the highest RMSE value (1.20 t ha^{-1}) recorded in field 77. This reduction in predictive accuracy may be associated with rapid canopy expansion during stem elongation. However, Marti et al. [44] reported that NDVI is well correlated with leaf area index (LAI) and is more sensitive to changes in crop canopy at lower LAI during the early growth stages; as the canopy closes, however, the NDVI signal may become saturated, reducing its sensitivity to further changes in crop biomass and potentially weakening its relationship with final grain yield. At booting, SAVI provided relatively consistent and strong predictions across the fields, with R^2 values of 0.86–0.97 and a low RMSE of 0.34 t ha^{-1} in all fields. These results suggest that canopy development and biomass accumulation at booting are strongly associated with total yield, making this stage particularly suitable for early yield forecasting.

At the heading stage, EVI emerged as the best predictor. The highest prediction accuracy was observed in field 71 ($R^2 = 0.97$), followed by fields 77 ($R^2 = 0.89$) and 81 ($R^2 = 0.86$). The strong relationships observed at heading can be attributed to maximum canopy cover, peak chlorophyll content, and high photosynthetic activity, all of which directly influence grain formation and final yield. The effectiveness of EVI at this stage confirms its ability to accurately capture crop vigor under dense canopy conditions. Overall, EVI showed strong predictive performance during tillering and heading, while SAVI provided consistent performance during booting. In contrast, the jointing stage showed comparatively weaker prediction performance. These findings suggest that the reliability of yield prediction was influenced by both growth stage and field conditions, and that the practical usefulness of an index should be evaluated based on its overall prediction consistency rather than small differences in R^2 alone.

The impact of chlorophyll content (SPAD) and vegetation indices (EVI and NDVI) on yield prediction models was evaluated across different growth stages for the three experimental fields (71, 77, and 81). The model performance was evaluated using coefficient of determination (R^2), root mean square error (RMSE, t ha^{-1}), and statistical significance (p -value). The results reveal strong variability in model performance depending on both growth stage and spectral index (Table 7). At the tillering stage, SPAD showed very poor predictive capability ($R^2 = -2.63$ to -0.10), whereas EVI (0.91–0.95), NDVI (0.82–0.90), and SAVI (0.68–0.90) demonstrated strong and significant relationships with yield, indicating their suitability for early prediction. However, during the jointing and booting stages, model performance declined sharply, with several negative R^2 values observed for EVI (-3.80) and NDVI (-1.92), reflecting instability due to rapid canopy development. In contrast, SAVI remained relatively consistent during these stages ($R^2 = 0.69$ –0.97), highlighting its robustness against soil and structural effects. The heading stage exhibited the highest predictive performance across all indices, with R^2 values reaching up to 0.98 (NDVI) and 0.97 (EVI), indicating strong linkage between canopy characteristics and final yield formation. SPAD also performed well only at this stage ($R^2 = 0.65$ –0.96), confirming its limited applicability outside peak growth.

Table 6: Regression analysis of the best model for early prediction of wheat yield (Y_P) using vegetation indices for each field.

Growth Stage	Field ID	Prediction Model	R^2	RMSE (t ha ⁻¹)
Tillering	71	$Y_P = -3.198 + 37.491 * EVI$	0.94	0.37
	77	$Y_P = -2.616 + 24.116 * NDVI$	0.82	0.88
	81	$Y_P = -3.055 + 42.211 * EVI$	0.95	0.49
Jointing	71	$Y_P = -11.175 + 28.159 * SAVI$	0.84	0.72
	77	$Y_P = -11.102 + 29.306 * SAVI$	0.71	1.20
	81	$Y_P = -15.803 + 37.500 * SAVI$	0.69	0.73
Bootling	71	$Y_P = -9.824 + 27.825 * SAVI$	0.86	0.34
	77	$Y_P = -8.480 + 27.368 * SAVI$	0.91	0.34
	81	$Y_P = -10.211 + 30.631 * SAVI$	0.97	0.34
Heading	71	$Y_P = -6.145 + 29.974 * EVI$	0.97	0.43
	77	$Y_P = -10.853 + 27.585 * EVI$	0.89	0.80
	81	$Y_P = -8.840 + 38.817 * EVI$	0.86	0.64

Overall, the results suggest that early yield prediction in wheat is feasible using EVI and NDVI at the tillering stage, while SAVI provides the most stable predictions across all stages. The poor performance during intermediate stages indicates that these phases are less suitable for yield estimation due to canopy transitions.

Table 7: Performance of the vegetation indices used in the yield prediction models.

FieldID	Growth Stage	SPAD			EVI			NDVI			SAVI		
		R^2	RMSE	p	R^2	RMSE	p	R^2	RMSE	p	R^2	RMSE	p
71	Tillering	-2.63	2.85	0.01	0.94	0.37	0.00	0.88	0.52	0.00	0.90	0.47	0.00
	Jointing	-0.13	1.93	0.70	-0.34	2.11	0.32	-0.58	2.28	0.31	0.84	0.72	0.00
	Bootling	-0.19	1.00	0.89	-0.23	1.02	0.87	-0.13	0.98	0.85	0.86	0.34	0.00
	Heading	0.96	0.50	0.00	0.97	0.43	0.00	0.92	0.67	0.00	0.89	0.80	0.00
77	Tillering	-0.24	2.36	0.19	0.91	0.62	0.00	0.82	0.88	0.00	0.68	1.20	0.00
	Jointing	-0.24	2.50	0.21	-0.01	2.26	0.77	-0.02	2.27	0.74	0.71	1.20	0.00
	Bootling	-0.68	1.45	0.96	-1.35	1.71	0.28	-1.12	1.63	0.02	0.91	0.34	0.00
	Heading	0.65	1.26	0.00	0.53	1.46	0.00	0.77	1.02	0.00	0.50	1.50	0.00
81	Tillering	-0.10	2.37	0.86	0.95	0.49	0.00	0.90	0.73	0.00	0.83	0.93	0.00
	Jointing	-0.16	1.42	0.74	-3.80	2.87	0.02	-1.92	2.24	0.25	0.69	0.73	0.00
	Bootling	0.00	2.08	0.95	-0.19	2.27	0.74	-0.03	2.11	0.91	0.97	0.34	0.00
	Heading	0.92	0.48	0.00	0.86	0.64	0.00	0.98	0.25	0.00	0.84	0.70	0.00

3.3 Multiple Linear Regression Analysis SPAD Readings and All Vegetation Indices

To identify the most suitable vegetation index (VI) from machine learning (ML) analysis for predicting wheat yield (Y_P , t ha⁻¹), the observed wheat yield values and corresponding VIs were subjected to linear regression analysis using IBM SPSS Statistics software. Linear regression results showed three models with a significant correlation between VIs and Y_A , with R^2 values ranging from 0.81 to 0.99 (Table 8). Among the studied VIs, EVI, SAVI, and NDVI were identified as the most suitable vegetation indices from the ML analysis for predicting wheat yield.

At the tillering stage, the models already showed high predictive performance, particularly in field 81 ($R^2 = 0.98$, RMSE = 0.32 t ha⁻¹), followed by fields 77 and 71. During the jointing stage, model accuracy decreased slightly, especially in field 77 ($R^2 = 0.81$, RMSE = 0.88 t ha⁻¹), while fields 71 and 81 maintained

relatively strong performance ($R^2 = 0.93$). This reduction may be related to rapid canopy expansion and increased variability in crop growth, which introduces noise into spectral–yield relationships even when multiple predictors are included. While at the booting stage, prediction accuracy improved again across all fields ($R^2 = 0.94$ – 0.95), with relatively stable RMSE values (0.47 – 0.54 t ha^{-1}). This indicates that combining vegetation indices with SPAD values is particularly effective during this stage, when both chlorophyll content and canopy structure are strongly linked to final grain yield formation.

Table 8: Regression analysis for early prediction of wheat yield using the SPAD readings and vegetation indices.

Regression	Stage	Field ID	Prediction Model	R^2	RMSE (t ha^{-1})
Multiple	Tillering	71	$Y_P = -2.892 - 0.027 * \text{SPAD} - 6.556 * \text{NDVI} + 57.170 * \text{EVI} - 7.257 * \text{SAVI} + 10.907 * \text{NDWI}$	0.93	0.59
		77	$Y_P = -0.977 - 0.036 * \text{SPAD} + 8.232 * \text{NDVI} + 32.760 * \text{EVI} - 10.919 * \text{SAVI} + 23.828 * \text{NDWI}$	0.95	0.44
		81	$Y_P = -1.194 + 0.035 * \text{SPAD} - 8.733 * \text{NDVI} + 41.834 * \text{EVI} + 1.166 * \text{SAVI} + 13.032 * \text{NDWI}$	0.98	0.32
	Jointing	71	$Y_P = -15.628 + 0.013 * \text{SPAD} + 10.150 * \text{NDVI} - 9.403 * \text{EVI} + 29.545 * \text{SAVI} + 3.564 * \text{NDWI}$	0.93	0.59
		77	$Y_P = -1.351 + 0.070 * \text{SPAD} - 36.730 * \text{NDVI} + 30.123 * \text{EVI} + 27.409 * \text{SAVI} - 3.595 * \text{NDWI}$	0.81	0.88
		81	$Y_P = -22.572 + 0.069 * \text{SPAD} - 3.380 * \text{NDVI} + 16.852 * \text{EVI} + 41.059 * \text{SAVI} - 27.678 * \text{NDWI}$	0.93	0.64
	Booting	71	$Y_P = -11.368 - 0.003 * \text{SPAD} + 7.606 * \text{NDVI} - 12.695 * \text{EVI} + 27.953 * \text{SAVI} + 7.838 * \text{NDWI}$	0.95	0.47
		77	$Y_P = -1.883 + 0.015 * \text{SPAD} - 8.570 * \text{NDVI} + 1.916 * \text{EVI} + 24.912 * \text{SAVI} + 2.439 * \text{NDWI}$	0.94	0.50
		81	$Y_P = -12.468 + 0.019 * \text{SPAD} - 6.883 * \text{NDVI} + 14.398 * \text{EVI} + 31.633 * \text{SAVI} - 0.412 * \text{NDWI}$	0.95	0.54
	Heading	71	$Y_P = -20.237 + 0.126 * \text{SPAD} + 21.581 * \text{NDVI} + 16.372 * \text{EVI} - 3.746 * \text{SAVI} + 11.920 * \text{NDWI}$	0.98	0.30
		77	$Y_P = -41.092 + 0.816 * \text{SPAD} + 69.670 * \text{NDVI} - 10.843 * \text{EVI} - 9.022 * \text{SAVI} + 9.471 * \text{NDWI}$	0.94	0.49
		81	$Y_P = -30.567 - 0.187 * \text{SPAD} + 61.375 * \text{NDVI} + 9.143 * \text{EVI} - 1.113 * \text{SAVI} - 6.719 * \text{NDWI}$	0.99	0.25

Note: Y_P : Predicted wheat yield, RMSE: Root mean square error.

The heading stage produced the highest overall predictive performance, especially in field 81 ($R^2 = 0.99$, $\text{RMSE} = 0.25 \text{ t ha}^{-1}$), followed by fields 71 and 77. This stage corresponds to peak canopy development and photosynthetic activity, where all spectral indicators and SPAD values collectively reflect crop vigor and yield potential more accurately than earlier stages. Furthermore, the heading stage is a critical stage when winter wheat chlorophyll status, biomass, and photosynthetic activity are closely related to subsequent grain formation. However, the results indicate overfitting, it may be due to greater similarities and high consistency in crop management across the fields. In general, across all stages, the inclusion of multiple predictors; SPAD combined with vegetation indices improved model performance. However, EVI and SAVI contributed strongly during dense canopy stages, while NDWI reflect plant water status, enhancing model sensitivity. Overall, the multi-variable regression models provide more accurate and stable wheat yield predictions, with the highest performance observed during the growth stages.

The slight differences in R^2 values are biologically plausible but appear to have limited practical significance because NDVI, EVI, and SAVI are derived from closely related spectral information and respond to common winter wheat characteristics such as canopy greenness, biomass, and leaf area. Under the hyper-arid conditions of Wadi Al-Dawasir, SAVI may theoretically reduce the influence of bright, dry soil, particularly during early growth when canopy cover is incomplete [39]. Likewise, EVI, which is designed to improve sensitivity under dense canopies, did not consistently outperform NDVI [38]. Overall, the results suggest that growth stage had greater practical importance than the small differences among NDVI, EVI, and SAVI.

Moreover, the results highlighted the advantage of integrating spectral and physiological measurements to better characterize crop variability. Overall, the results showed strong relationships between SPAD measurements and vegetation indices across all growth stages. Furthermore, Kızılgeçi et al. [45] investigated the relationships between NDVI, SPAD readings, chlorophyll meter measurements, and wheat grain yield at different growth stages. They reported that the use of vegetation cover indicators and chlorophyll measurements improves crop condition assessment and contributes to yield estimation. Zhao et al. [31] also demonstrated that combining complementary crop indicators can improve wheat yield prediction accuracy.

As shown in Table 9, the regression model developed from the dataset of each field at each growth stage (training datasets) was validated against independent datasets from the other fields (test datasets). The validation results revealed that the model developed from the dataset of field 77 achieved better predictive accuracy when applied to independent datasets from other fields at the same growth stages. Among the studied wheat fields, the best performance was noticed for field 77, with a low RMSE at the tillering (0.997 ha t^{-1}), the booting (1.291 ha t^{-1}), and the heading stages (0.741 ha t^{-1}). At the jointing stage, the model validation model developed from the dataset of field 71 achieved the highest R^2 (0.915) and the lowest RMSE (1.932 ha t^{-1}).

Table 9: Validation regression analysis for early prediction of wheat yield using the SPAD values and vegetation indices.

Growth Stages	Prediction Model	Trained Field								
		Field ID (77 + 81)			Field ID (71 + 81)			Field ID (71 + 77)		
		R^2	SE	RMSE %	R^2	SE	RMSE %	R^2	SE	RMSE %
Tillering	Model-Field 71	0.943	0.472	1.766						
	Model-Field 77				0.912	0.510	0.997			
	Model-Field 81							0.904	0.665	1.481
Jointing	Model-Field 71	0.915	0.521	1.932						
	Model-Field 77				0.894	0.553	1.967			
	Model-Field 81							0.914	0.644	4.558
Booting	Model-Field 71	0.915	0.521	3.188						
	Model-Field 77				0.894	0.553	1.291			
	Model-Field 81							0.914	0.644	2.498
Heading	Model-Field 71	0.927	0.496	1.591						
	Model-Field 77				0.941	0.470	0.741			
	Model-Field 81							0.930	0.571	1.533

Note: SE: Standard error, RMSE: Root mean square error.

3.4 Machine Learning Models

The predictive performance of the machine-learning models RF, CART, and CNN modules (Table 10) varied across growth stages, with the highest accuracy observed during the booting and heading stages.

The wheat heading stage showed the best yield prediction performance, particularly with the RF model, as indicated by the lowest mean squared error ($MSE = 4.73 \text{ t ha}^{-1}$) and the highest R^2 value (0.75). In contrast, the CNN and CART models performed less effectively at this stage; the CNN model showed lower predictive performance at the heading stage, with an MSE of 8.15 and an R^2 of 0.67. However, during the early growth stages, the jointing stage emerged as the most informative stage. At this stage, the CNN achieved the best yield prediction performance ($MSE = 5.26 \text{ t ha}^{-1}$, $R^2 = 0.64$), followed by the RF model ($MSE = 5.51 \text{ t ha}^{-1}$, $R^2 = 0.62$). These results indicate that the physiological and spectral parameters measured at the jointing stage are reliable early indicators of wheat yield.

On the other hand, the booting stage also provided valuable information for the prediction of wheat yield, with both RF and CNN models showing similar results ($MSE = 5.54\text{--}5.49 \text{ t ha}^{-1}$, $R^2 = 0.72\text{--}0.73$), respectively. However, the tillering stage showed the poorest yield prediction results for all models, particularly for the CART which showed the highest prediction errors ($MSE = 13.63 \text{ t ha}^{-1}$, $R^2 = -1.83$). Overall, the RF and CNN models consistently outperformed the CART model across all growth stages considered.

Table 10: Performance and accuracy of studied machine learning models for wheat yield prediction. (ML: Machine learning, MSE: Mean square error).

Wheat Growth Stages	ML Technique	RMSE (t ha^{-1})	R^2
Tillering	RF	7.08	-0.47
	CART	13.63	-1.83
	CNN	8.75	-0.82
Jointing	RF	5.51	0.62
	CART	6.03	0.58
	CNN	5.26	0.64
Booting	RF	5.54	0.72
	CART	6.61	0.74
	CNN	5.49	0.73
Heading	RF	4.73	0.75
	CART	5.28	0.73
	CNN	8.15	0.67

For the ML analysis, the results indicated that CART and CNN were the most reliable models, showing a strong relationship between VIs, SPAD values, and actual yield (Y_A). However, Xu et al. [46] demonstrated the effectiveness of RF and deep-learning approaches for winter wheat yield prediction using hyperspectral and multispectral data. For example, RF and CNN models have been successfully applied to winter wheat yield predictions using UAV-based multispectral and hyperspectral data, with model performance varying according to growth stage and input features [46]. To evaluate their accuracy in predicting wheat yield (Y_P), the selected models, including RF and CNN were subjected to cross-validation using independent samples that were not used in model development (35% of the total dataset). CART and CNN exhibited high predictive accuracy, as indicated by the correlation between predicted yield (Y_P) and SPAD values at the tillering stage. However, the RF model showed lower prediction accuracy compared with CART and CNN, as confirmed by RMSE and MBE values. For the NDVI+VIS dataset, the CNN model performed best in predicting wheat yield.

In general, the most predictive vegetation indices (VIs) for wheat were NDVI and SAVI in the ML analysis at the tillering and jointing stages, respectively. For the NDVI+VIS dataset, the most predictive VIs included NDVI and EVI at the jointing and booting stages, respectively. Similar results were reported

in previous studies. For example, Kaya et al. [47] reported a moderate linear correlation between wheat yield and NDVI at early growth stages and a high correlation at the flowering stage. In addition, Jurecka et al. [48] reported a strong correlation between wheat yield and NDVI ($R^2 = 0.78\text{--}0.82$) and EVI ($R^2 = 0.71$).

Despite the promising results, models exhibited high coefficients of determination, especially with multiple linear regression analysis, these results should be interpreted with caution. The dataset consisted of a limited number of observations collected during a single growing season under relatively homogeneous environmental and management conditions. Under such circumstances, model performance metrics may be optimistic and may not fully reflect predictive performance under independent conditions. Furthermore, the vegetation indices used in this study are derived from related spectral bands and therefore contain partially overlapping information, which may contribute to high model fitness. While internal validation indicated satisfactory predictive performance, additional validation across multiple growing seasons, locations, and environmental conditions is required to evaluate model robustness and transferability. Consequently, the developed models should be considered preliminary yield-prediction tools applicable to conditions similar to those represented in the present study.

3.5 LSD Analysis

A split-plot analysis was performed on predicted wheat yield obtained from multiple regression models across three fields and growth stages. The statistical results revealed that no significant differences were observed between the predicted wheat yields of field 81 (8.614 t ha^{-1}) and field 77 (8.560 t ha^{-1}). However, the predicted yield from the dataset of field 77 showed better model performance and higher prediction accuracy during the validation analysis across most growth stages. This may attribute greater field homogeneity, which likely contributed to the improved reliability and transferability of the model developed from the field of 77 datasets for yield prediction in independent fields. But both fields had significantly higher predicted yields than field 71 (8.364 t ha^{-1}). The observed variation in crop yield among fields may be attributed to differences in soil fertility and environmental stress conditions.

Also, the results revealed significant differences in predicted yield among growth stages (Table 11). The booting stage showed the highest mean yield (9.278 t ha^{-1}) and was significantly different from all other stages, while the jointing stage showed the lowest mean yield (7.552 t ha^{-1}). Notably, similar predictive performance of NDVI, EVI, and SAVI across growth stages indicates that these indices captured largely overlapping information related to canopy biomass, chlorophyll content, and photosynthetic activity. Since all three indices are primarily derived from red and near-infrared reflectance, their strong correlation with vegetation vigor resulted in comparable R^2 and RMSE values. Although EVI and SAVI were developed to improve performance under dense canopies and variable soil backgrounds, these advantages were minimal under the relatively uniform wheat-canopy conditions of this study. In contrast, NDWI provided additional information associated with canopy water status, suggesting that moisture-related spectral responses contributed complementary information for yield prediction beyond that contained in the greenness indices alone.

Table 11: Statistical comparison of predicted wheat yield across different growth stages (LSD = 0.134).

Growth Stages	Predicted Yield (t ha^{-1})
Booting	9.278 ^a
Tillering	8.525 ^b
Heading	8.379 ^b
Jointing	7.552 ^c

Note: Means followed by the same letter are not significantly different, whereas means followed by different letters are significantly different according to the LSD test at $p < 0.05$. LSD: Least significant difference.

Furthermore, the interaction analysis was conducted, and the results revealed significant differences among fields and growth stage combinations (Table 12). The highest predicted yield was recorded in field 77 at the booting stage (9.858 t ha⁻¹), which was significantly higher than all other treatments. While the lowest predicted yield was recorded at the jointing stage in both field 77 (7.392 t ha⁻¹) and field 71 (7.376 t ha⁻¹), indicating consistently poor performance at this growth stage.

Table 12: Statistical analysis of the interaction between fields and growth stages for predicted wheat yield (LSD = 0.233).

Field ID	Growth Stages	Predicted Yield (t ha ⁻¹)
77	Booting	9.858 ^a
81	Booting	9.068 ^b
71	Booting	8.908 ^{bc}
81	Tillering	8.644 ^{cd}
77	Tillering	8.524 ^d
77	Heading	8.517 ^d
81	Heading	8.452 ^{de}
71	Tillering	8.407 ^{de}
71	Heading	8.168 ^{ef}
81	Jointing	7.889 ^f
77	Jointing	7.392 ^g
71	Jointing	7.376 ^g

Note: Means followed by the same letter are not significantly different, whereas means followed by different letters are significantly different according to the LSD test at $p < 0.05$. LSD: Least significant difference.

4 Limitations of the Study

The developed statistical, machine learning, and deep learning models demonstrated promising performance for early wheat yield prediction under hyper-arid conditions. However, several limitations of this study are: (i) the models were developed using data collected from three wheat fields pertaining to a single growing season; therefore, model performance may be influenced by field management practices, soil heterogeneity, genotype-specific responses, and climatic variability; (ii) the study relied primarily on proximal sensing measurements, including hyperspectral reflectance and SPAD-derived chlorophyll observations, without explicitly incorporating meteorological, irrigation, or soil nutrient variables that may further affect yield variability. In addition, the predictive relationships identified in this study are specific to the environmental and management conditions under which the data were collected and may require recalibration before application to other production regions or cropping systems.

Despite these limitations, the developed models provide a practical framework for early-season wheat yield forecasting and demonstrate the potential of integrating hyperspectral sensing and chlorophyll measurements for precision agriculture applications in water-limited environments. The generated yield predictions can support irrigation scheduling, nutrient management, and decision-making aimed at improving resource-use efficiency and crop productivity. Future research should focus on validating the developed models across multiple growing seasons, locations, and environmental conditions. Furthermore, integrating hyperspectral observations with meteorological data, soil properties, crop management information, and satellite-derived metrics may enhance prediction accuracy and facilitate the development of scalable regional yield forecasting systems. The application of advanced deep learning architectures, feature selection techniques, and data-fusion approaches also warrants further investigation to strengthen early-season yield prediction capabilities in arid and hyper-arid agroecosystems.

5 Conclusions

A field study was conducted to employ remote sensing techniques in order to evaluate the performance of wheat yield and design crop yield prediction models. The experiment was conducted during one season across three fields within one commercial farm. The specific conclusions of the study include the following:

The interaction between wheat yield, plant vigor, and the SPAD values high accuracy models at the jointing growth stage ($R^2 = 0.83$; $RMSE = 2.89 \text{ t ha}^{-1}$). However, the lowest RMSE (1.59 t ha^{-1} ; 17.41%) and very low MBE (0.32 t ha^{-1} ; 3.5%) values were obtained at the heading stage. Therefore, the heading stage was thought to be suitable for early yield prediction using the SPAD values.

However, enhanced yield prediction was achieved at the booting stage using multivariate models (RF: $R^2 = 0.72$, CART: $R^2 = 0.74$, CNN: $R^2 = 0.73$). The study also confirmed moderate early forecasting accuracy at the jointing stage ($R^2 = 0.63$), when multivariate combinations such as SPAD + NDVI, SPAD + EVI, SPAD + SAVI, and SPAD + NDWI were evaluated.

In general, across all stages, the inclusion of multiple predictors, SPAD combined with vegetation indices, improved model performance. However, EVI and SAVI contributed strongly during dense canopy stages, while NDWI reflected the plant's water status, enhancing model sensitivity. Overall, the multi-variable regression models provide more accurate and stable wheat yield predictions, with the highest performance observed during the growth stages.

Acknowledgement: The authors are grateful to the Deanship of Scientific Research, King Saud University, for funding this study through the Ongoing Research Funding program (ORF-2026-1740), King Saud University, Riyadh, Saudi Arabia.

Funding Statement: Ongoing Research Funding program (ORF-2026-1740), King Saud University, Riyadh, Saudi Arabia.

Author Contributions: Conceptualization, Khalid A. Al-Gaadi and Rangaswamy Madugundu; data curation, Ahmed M. Zeyada and Elkamil Tola; formal analysis, Khalid A. Al-Gaadi, Rangaswamy Madugundu and Ahmed M. Zeyada; investigation, Elkamil Tola and Haroon F. Edrees; methodology, Rangaswamy Madugundu, Ahmed M. Zeyada, Elkamil Tola and Haroon F. Edrees; resources, Khalid A. Al-Gaadi and Rangaswamy Madugundu; supervision, Khalid A. Al-Gaadi; writing—original draft, Rangaswamy Madugundu; writing—review & editing, Ahmed M. Zeyada, Elkamil Tola and Haroon F. Edrees. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: Data available within the article and its Appendix A. Any additional data required to support the findings of this study will be provided upon reasonable request from the authors.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Table A1: Hyperparameter configuration and architectural settings of the Classification and Regression Tree (CART), Random Forest (RF), and one-dimensional Convolutional Neural Network (1D-CNN) models used for wheat grain yield prediction based on SPAD measurements and vegetation indices (NDVI, EVI, SAVI, and NDWI).

Parameter	CART	Random Forest (RF)	1D-CNN
Model type	Decision tree	Ensemble of decision trees	Deep learning neural network
Input variables	SPAD, NDVI, EVI, SAVI, NDWI	SPAD, NDVI, EVI, SAVI, NDWI	SPAD, NDVI, EVI, SAVI, NDWI

Table A1: Cont.

Parameter	CART	Random Forest (RF)	1D-CNN
Training/Validation split	65%/35%	65%/35%	65%/35%
Random seed	42	42	42
Tree construction	Recursive binary splitting	Bootstrap aggregation (bagging)	Convolutional feature extraction
Maximum tree depth	5	Unlimited	Not applicable
Minimum samples for split	2	2	Not applicable
Minimum samples per leaf	1	1	Not applicable
Pruning	Cost-complexity pruning	Not required	Not applicable
Number of trees (<i>ntree</i>)	1	500	Not applicable
Number of variables per split (<i>mtry</i>)	All variables evaluated	Optimized ($\approx \sqrt{p}$)	Not applicable
Bootstrap sampling	No	Yes	No
Out-of-bag (OOB) estimation	No	Yes	No
Convolution layers	No	No	1
Number of filters	No	No	32
Kernel size	No	No	2
Activation function	No	No	ReLU
Pooling layer	No	No	Max Pooling
Dense layer neurons	No	No	64
Dropout rate	No	No	0.2
Optimizer	No	No	Adam
Learning rate	No	No	0.001
Batch size	No	No	16
Maximum epochs	No	No	100
Early stopping	No	No	Yes
Variable importance	Yes	Yes	No
Model output	Grain yield (t ha ⁻¹)	Grain yield (t ha ⁻¹)	Grain yield (t ha ⁻¹)
Evaluation metrics	R ² , RMSE, MAE	R ² , RMSE, MAE	R ² , RMSE, MAE

References

- Shiferaw B, Smale M, Braun HJ, Duveiller E, Reynolds M, Muricho G. Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. *Food Secur.* 2013;5(3):291–317. [CrossRef].
- Barbanti L, Adroher J, Damian JM, Di Virgilio N, Falsone G, Zucchelli M, et al. Assessing wheat spatial variation based on proximal and remote spectral vegetation indices and soil properties. *Ital J Agron.* 2018;13(1):1086. [CrossRef].
- Marino S. Understanding the spatio-temporal behavior of crop yield, yield components and weed pressure using time series Sentinel-2-data in an organic farming system. *Eur J Agron.* 2023;145:126785. [CrossRef].
- Uribe-Opazo MA, Dalposso GH, Galea M, Johann JA, De Bastiani F, Moyano ENC, et al. Spatial variability of wheat yield using the Gaussian spatial linear model. *Aust J Crop Sci.* 2023;17(2):179–89. [CrossRef].
- Nduku L, Munghemezulu C, Mashaba-Munghemezulu Z, Masiza W, Ratshiedana PE, Kalumba AM, et al. Field-scale winter wheat growth prediction applying machine learning methods with unmanned aerial vehicle imagery and soil properties. *Land.* 2024;13(3):299. [CrossRef].
- Perry E, Sheffield K, Crawford D, Akpa S, Clancy A, Clark R. Spatial and temporal biomass and growth for grain crops using NDVI time series. *Remote Sens.* 2022;14(13):3071. [CrossRef].
- Yu H, Kong B, Hou Y, Xu X, Chen T, Liu X. A critical review on applications of hyperspectral remote sensing in crop monitoring. *Exp Agric.* 2022;58:e26. [CrossRef].
- Mulla DJ. Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosyst Eng.* 2013;114(4):358–71. [CrossRef].
- Lu B, Dao P, Liu J, He Y, Shang J. Recent advances of hyperspectral imaging technology and applications in agriculture. *Remote Sens.* 2020;12(16):2659. [CrossRef].

10. Ma J, Zheng B, He Y. Applications of a hyperspectral imaging system used to estimate wheat grain protein: A review. *Front Plant Sci.* 2022;13:837200. [[CrossRef](#)].
11. Li L, Lin D, Wang J, Yang L, Wang Y. Multivariate analysis models based on full spectra range and effective wavelengths using different transformation techniques for rapid estimation of leaf nitrogen concentration in winter wheat. *Front Plant Sci.* 2020;11:755. [[CrossRef](#)].
12. Al-Gaadi. In-season assessment of wheat crop health using vegetation indices based on ground measured hyper spectral data. *Am J Agric Biol Sci.* 2014;9(2):138–46. [[CrossRef](#)].
13. Polivova M, Brook A. Detailed investigation of spectral vegetation indices for fine field-scale phenotyping. In: *Vegetation index and dynamics*. London, UK: IntechOpen; 2022. [[CrossRef](#)].
14. Hunt ML, Blackburn GA, Carrasco L, Redhead JW, Rowland CS. High resolution wheat yield mapping using Sentinel-2. *Remote Sens Environ.* 2019;233:111410. [[CrossRef](#)].
15. West H, Quinn N, Horswell M. Remote sensing for drought monitoring & impact assessment: Progress, past challenges and future opportunities. *Remote Sens Environ.* 2019;232:111291. [[CrossRef](#)].
16. Jackson TJ, Chen D, Cosh M, Li F, Anderson M, Walthall C, et al. Vegetation water content mapping using Landsat data derived normalized difference water index for corn and soybeans. *Remote Sens Environ.* 2004;92(4):475–82. [[CrossRef](#)].
17. Thapa S, Rudd JC, Xue Q, Bhandari M, Reddy SK, Jessup KE, et al. Use of NDVI for characterizing winter wheat response to water stress in a semi-arid environment. *J Crop Improv.* 2019;33(5):633–48. [[CrossRef](#)].
18. Consoli S, Vanella D. Mapping crop evapotranspiration by integrating vegetation indices into a soil water balance model. *Agric Water Manag.* 2014;143:71–81. [[CrossRef](#)].
19. Dash SK, Sembhi H, Langsdale M, Wooster M, Dodd E, Ghent D, et al. Assessing the field-scale crop water condition over an intensive agricultural plain using UAV-based thermal and multispectral imagery. *J Hydrol.* 2025;655:132966. [[CrossRef](#)].
20. Wang J, Lou Y, Wang W, Liu S, Zhang H, Hui X, et al. A robust model for diagnosing water stress of winter wheat by combining UAV multispectral and thermal remote sensing. *Agric Water Manag.* 2024;291:108616. [[CrossRef](#)].
21. Zsebő S, Bede L, Kukorelli G, Kulmány IM, Milics G, Stencinger D, et al. Yield prediction using NDVI values from GreenSeeker and MicaSense cameras at different stages of winter wheat phenology. *Drones.* 2024;8(3):88. [[CrossRef](#)].
22. Royo C, Aparicio N, Villegas D, Casadesus J, Monneveux P, Araus JL. Usefulness of spectral reflectance indices as durum wheat yield predictors under contrasting Mediterranean conditions. *Int J Remote Sens.* 2003;24(22):4403–19. [[CrossRef](#)].
23. Prasad B, Carver BF, Stone ML, Babar MA, Raun WR, Klatt AR. Potential use of spectral reflectance indices as a selection tool for grain yield in winter wheat under great Plains conditions. *Crop Sci.* 2007;47(4):1426–40. [[CrossRef](#)].
24. Leukel J, Zimpel T, Stumpe C. Machine learning technology for early prediction of grain yield at the field scale: A systematic review. *Comput Electron Agric.* 2023;207:107721. [[CrossRef](#)].
25. Fieuzal R, Bustillo V, Collado D, Dedieu G. Combined use of multi-temporal landsat-8 and sentinel-2 images for wheat yield estimates at the intra-plot spatial scale. *Agronomy.* 2020;10(3):327. [[CrossRef](#)].
26. Abu Javed M, Azmi Murad MA. Crop yield prediction in agriculture: A comprehensive review of machine learning and deep learning approaches, with insights for future research and sustainability. *Heliyon.* 2024;10(24):e40836. [[CrossRef](#)].
27. Gumma MK, Nukala RM, Panjala P, Bellam PK, Gajjala S, Dubey SK, et al. Optimizing crop yield estimation through geospatial technology: A comparative analysis of a semi-physical model, crop simulation, and machine learning algorithms. *AgriEngineering.* 2024;6(1):786–802. [[CrossRef](#)].
28. Wang J, Wang Y, Qi Z. Remote sensing data assimilation in crop growth modeling from an agricultural perspective: New insights on challenges and prospects. *Agronomy.* 2024;14(9):1920. [[CrossRef](#)].
29. Bharadiya JP, Tzenios NT, Reddy M. Forecasting of crop yield using remote sensing data, agrarian factors and machine learning approaches. *J Eng Res Rep.* 2023;24(12):29–44.
30. Behpouri A, Farokhzadeh S, Zinati Z, Khosravi Z. Use of multivariate analysis and machine learning methods to characterize traits contributing to wheat yield diversity. *Span J Agric Res.* 2023;21(1):e0901. [[CrossRef](#)].

31. Zhao Y, Potgieter AB, Zhang M, Wu B, Hammer GL. Predicting wheat yield at the field scale by combining high-resolution sentinel-2 satellite imagery and crop modelling. *Remote Sens.* 2020;12(6):1024. [[CrossRef](#)].
32. Dalla Marta A, Grifoni D, Mancini M, Orlando F, Guasconi F, Orlandini S. Durum wheat in-field monitoring and early-yield prediction: Assessment of potential use of high resolution satellite imagery in a hilly area of Tuscany, Central Italy. *J Agric Sci.* 2015;153(1):68–77. [[CrossRef](#)].
33. Choudhury MR, Das S, Christopher J, Apan A, Chapman S, Menzies NW, et al. Improving biomass and grain yield prediction of wheat genotypes on sodic soil using integrated high-resolution multispectral, hyperspectral, 3D point cloud, and machine learning techniques. *Remote Sens.* 2021;13(17):3482. [[CrossRef](#)].
34. Guo BB, Feng YL, Ma C, Zhang J, Song X, Wang MY, et al. Suitability of different multivariate analysis methods for monitoring leaf N accumulation in winter wheat using *in situ* hyperspectral data. *Comput Electron Agric.* 2022;198:107115. [[CrossRef](#)].
35. Zhang Y, Fan Z, Li S, Liu L, Feng S, Wang Y, et al. Multimodal data fusion and attention-based deep learning for estimating winter wheat chlorophyll content. *Comput Electron Agric.* 2026;245:111536. [[CrossRef](#)].
36. El-Hendawy SE, Alotaibi M, Al-Suhaibani N, Al-Gaadi K, Hassan W, Dewir YH, et al. Comparative performance of spectral reflectance indices and multivariate modeling for assessing agronomic parameters in advanced spring wheat lines under two contrasting irrigation regimes. *Front Plant Sci.* 2019;10:1537. [[CrossRef](#)].
37. Rouse JW, Haas RH, Schell JA, Deering DW. Monitoring vegetation systems in the great Plains with ERTS. In: *Proceedings of the Goddard Space Flight Center Third Earth Resources Technology Satellite-1 Symposium*; 1973 Dec 10–14; Washington, DC, USA. Washington, DC, USA: NASA; 1973. p. 309–17.
38. Huete A, Didan K, Miura T, Rodriguez EP, Gao X, Ferreira LG. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens Environ.* 2002;83(1–2):195–213. [[CrossRef](#)].
39. Huete AR. A soil-adjusted vegetation index (SAVI). *Remote Sens Environ.* 1988;25(3):295–309. [[CrossRef](#)].
40. Gao BC. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sens Environ.* 1996;58(3):257–66. [[CrossRef](#)].
41. Breiman L. Random forests. *Mach Learn.* 2001;45(1):5–32. [[CrossRef](#)].
42. Carrizosa E, Molero-Río C, Romero Morales D. Mathematical optimization in classification and regression trees. *TOP.* 2021;29(1):5–33. [[CrossRef](#)].
43. Nevavuori P, Narra N, Lipping T. Crop yield prediction with deep convolutional neural networks. *Comput Electron Agric.* 2019;163:104859. [[CrossRef](#)].
44. Marti J, Bort J, Slafer GA, Araus JL. Can wheat yield be assessed by early measurements of Normalized Difference Vegetation Index? *Ann Appl Biol.* 2007;150(2):253–7. [[CrossRef](#)].
45. Kızılgeçi F, Tebrizli N, Eliş S, Özkan R, Bayhan M, Yıldırım M. Non-Destructive chlorophyll meters: A comparison of three types of meters for grain yield estimation of durum wheat under semi-arid environments. *Int J Agric Environ Food Sci.* 2025;9(1):115–22. [[CrossRef](#)].
46. Xu H, Yin H, Liu Y, Wang B, Song H, Zheng Z, et al. Regional winter wheat yield prediction and variable importance analysis based on multisource environmental data. *Agronomy.* 2024;14(8):1623. [[CrossRef](#)].
47. Kaya Y, Polat N. A linear approach for wheat yield prediction by using different spectral vegetation indices. *Int J Eng Geosci.* 2023;8(1):52–62. [[CrossRef](#)].
48. Jurecka F, Hlavinka P, Lukas V, Trnka M, Zalud Z. Crop yield estimation in the field level using vegetation indices. In: *Proceedings of the International PhD Students Conference, MendelNet 2016*; 2016 Nov 9–10; Brno, Czech Republic.