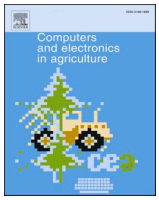
Computers and Electronics in Agriculture 225 (2024) 109329



Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag





Corn yield prediction in site-specific management zones using proximal soil sensing, remote sensing, and machine learning approach

Bere Benjamin Bantchina [a,*], Muhammad Qaswar [b], Selçuk Arslan [c], Yahya Ulusoy [d], Kemal Sulhi Gündoğdu [c], Yücel Tekin [d], Abdul Mounem Mouazen [b]

[a] Department of Biosystems Engineering, Institute of Natural and Applied Sciences, Bursa Uludağ University, 16059, Bursa, Turkey
[b] Department of Environment, Faculty of Biosystems Engineering, Ghent University, 9000, Ghent, Belgium
[c] Department of Biosystems Engineering, Faculty of Agriculture, Bursa Uludağ University, 16059, Bursa, Turkey
[d] Department of Machine and Metal Technologies, Vocational School of Technical Sciences, Bursa Uludağ University, 16059, Bursa, Turkey

ARTICLE INFO

Keywords:
Variable rate fertilization
Corn yield prediction
Proximal soil sensing
Remote sensing
Management zones
Machine learning

ABSTRACT

The integration of advanced technologies, such as soil proximal sensing, remote sensing, and machine learning, has revolutionized agricultural practices, particularly for corn yield prediction. This interdisciplinary approach harnesses the power of cutting-edge sensors to gather high-resolution data on soil conditions coupled with remote sensing technologies that provide a comprehensive view of crop health and environmental factors. This study aimed to evaluate the feasibility of accurately predicting corn (*Zea mays*) yield at the management zones (MZs) level using the fusion of visible and near-infrared spectroscopy (Vis-NIRS)-derived soil properties, remote sensing-derived crop spectral indices, and machine learning algorithms. Clustering analysis was used to develop MZs to implement variable-rate nitrogen fertilization (VRNF) in a drip-irrigated corn field. Site-specific models to forecast corn yield at the MZs level were developed using Sentinel 2A-derived spectral indices and machine learning regression algorithms. Partial least squares Vis-NIR spectral regression modelling for MZs development achieved high accuracy in terms of the coefficient of determination (R^2) which was ranged from 0.60 to 0.99 in cross-validation and from 0.52 to 0.78 in online validation. The developed corn yield prediction models demonstrated moderate efficacy, as evidenced by the R^2 values ranging from 0.50 to 0.71. Further research should include supplementary spectral crop canopy indices and the application of alternative deep and machine learning approaches to improve the accuracy of the prediction models.

1. Introduction

Variable rate fertilization (VRF) plays a key role in precision agriculture (PA) by offering tailored and efficient fertilizer applications and nutrient management to address in-field variability. Variable rate technology (VRT) has the ability to improve profitability by optimizing crop inputs based on site-specific management zones (Koch et al., 2004; Nawar et al., 2017). This approach offers an efficient strategy for addressing the requirements of contemporary agriculture, including "green planting," on-demand input, and non-uniform fertilization, thus functioning as the fundamental principle of PA (Bai et al., 2022). The use of proximal and remote sensors in the agricultural sector has witnessed a notable increase, as they are employed to effectively regulate and oversee crop inputs. In recent years, these sensors have been primarily utilized in precision agriculture to identify appropriate targets and determine the specific requirements of crops to facilitate variable-rate applications (Lee et al. 2010; Mouazen et al., 2019). Data collection using proximal sensors holds significant value for producers and land managers in creating site-specific databases. These databases provide insights into the relationships among nutrients in soil concentrations, plant responses, crop growth over the growing period, and subsequent yield (Viscarra Rossel and Bouma, 2016). Therefore, for the successful implementation of VRT, high-resolution data on yield-limiting factors are necessary, for which proximal and remote sensing technologies have become available and overcome the shortcomings of traditional laboratory analyses.

Proximal soil sensors provide a more detailed assessment of crucial soil characteristics, enabling quantification of spatial and temporal variations within a field. This capability is essential for PA applications (Guerrero et al., 2021a). The accurate characterization of soil properties

* Corresponding author.
E-mail address: ben.bantchina@gmail.com (B.B. Bantchina).

https://doi.org/10.1016/j.compag.2024.109329
Received 19 December 2023; Received in revised form 3 June 2024; Accepted 7 August 2024
Available online 17 August 2024
0168-1699/© 2024 Elsevier B.V. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

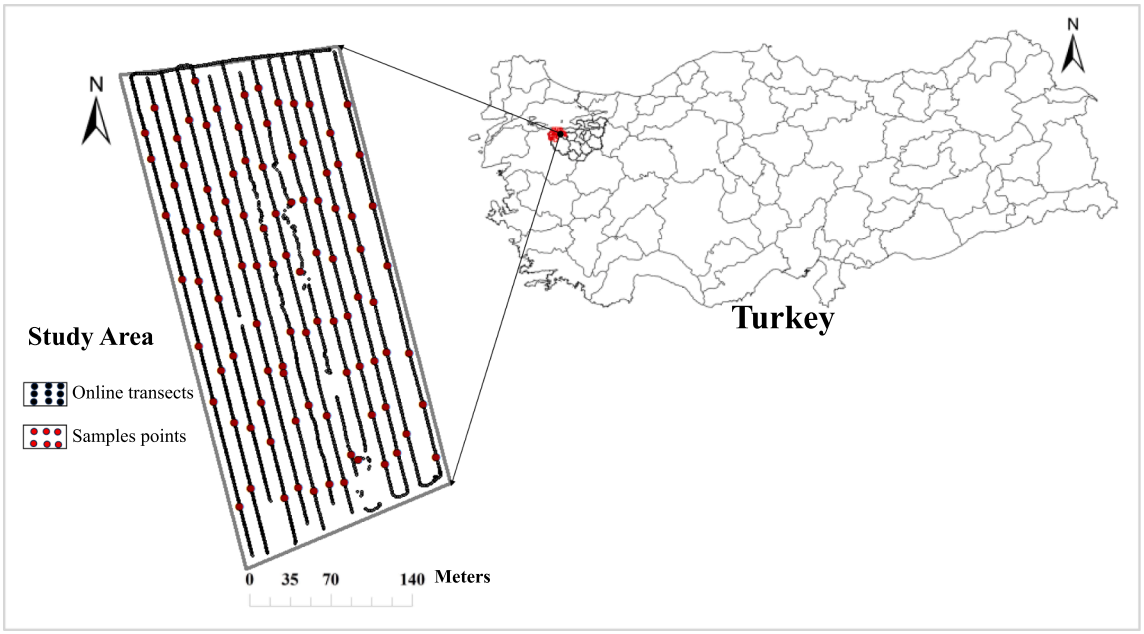


Fig. 1. Experimental site location, showing the Vis-NIRS sensing transects and random soil sampling points.

is a crucial determinant in ensuring the successful implementation of VRF technologies. In this context, the utilization of visible and near-infrared spectroscopy (Vis-NIRS) has proven to be effective in the development of models and mapping of soil attributes (Stenberg et al., 2010; Kuang et al., 2012). This has been achieved through the analysis of soil spectra collected online (the sensor collects data while pulled by a tractor), specifically for PA applications (Guerrero et al., 2021b). Vis-NIR spectroscopy has emerged as an appealing technique in the field of chemometrics (Mouazen et al., 2007; Viscarra-Rossel and Chen, 2011; Tekin et al., 2013; Kodaira and Shibusawa, 2013; Guerrero et al., 2021b). These studies demonstrated that Vis-NIR spectroscopy offers precise quantification of key physical and chemical soil properties, thereby proving its utility in the context of digital soil mapping. Similarly, remote sensing (RS) is also crucial in PA because it provides valuable data for crop monitoring and management. Technologies such as unmanned aerial vehicles (UAVs) and satellite imagery have been widely used to gather data for precision agriculture applications (Yue et al., 2012; Khanal et al., 2020; del Cerro et al., 2021). These technologies have enabled the development of advanced monitoring and decision-making tools to aid in optimizing agricultural practices (Shafi et al., 2019; Yang, 2020). The integration of RS with geographic information systems (GIS) has facilitated the mapping of spatial variability, which is a critical step in implementing PA practices (Varella et al., 2015; Toscano et al., 2019). The continuous development of modern RS technology offers new opportunities for the rapid monitoring of crop pests and diseases, and crop yield prediction, further enhancing precision agricultural practices (Yue et al., 2012; Segara et al., 2020; Yang, 2020). However, the integration of remote and proximal soil sensing for PA applications and yield prediction has received less attention than the individual use of each of the two technologies.

Crop yield prediction is a crucial aspect of PA, enabling farmers to optimize resource allocation, improve crop productivity, and minimize environmental impact. Traditional yield estimation methods such as field sampling and manual observation are often time-consuming, labor-intensive, and prone to inaccuracies. There is a growing trend among researchers to employ modern technologies such as machine learning (ML) and RS to enhance the precision and effectiveness of corn yield prediction models (Geipel et al., 2014; Tagarakis and Ketterings, 2017; Kayad et al., 2019; Hunt et al., 2019; Shahi et al., 2022; Holtorf et al., 2023). Vegetation indices derived from satellite and aerial imagery have shown promise for assessing corn yield variability and predicting grain yields. These indices, along with ML techniques, have been shown to be useful for predicting corn yield at different levels from the field to the country scale (Peng et al., 2018; Hao et al., 2019; Barzin et al., 2020; Zhao et al., 2020; Ji et al., 2021; Radočaj et al., 2023). ML techniques, such as support vector machines, random forests, and artificial neural networks, have demonstrated their efficacy in analyzing large datasets and identifying complex relationships between soil properties, spectral indices, and yield data (Chlingaryan et al., 2018). By training these algorithms on historical yield records and the corresponding soil and remote sensing data, researchers can develop predictive models that can estimate corn yields with high accuracy (Jeong et al., 2022; Mohapatra and Chaudhary, 2022; Hady and Kusumo, 2023). The use of these ML algorithms has shown tangible results in predicting crop yield compared with other methods (Sabitov et al., 2023).

The combination of soil proximal sensing and RS with advanced technologies such as ML and GIS has significantly improved the accuracy and efficiency of precision agriculture practices, offering new opportunities for sustainable and optimized crop production. Using these new tools, farmers can now find and study how crop fields change over time and space, which helps them make the best use of their resources (Duffera et al., 2007; Li et al., 2008). By considering the diversity within a specific field, it is possible to identify subregions that share comparable characteristics and can be categorized into distinct MZs with varying agricultural techniques (Nawar et al., 2017). A model based on a single MZ may exhibit higher accuracy than a model designed for an entire field because of the training and validation of the MZ concept using data from a region characterized by lower spatial variability than the entire field. Previous approaches to agricultural yield prediction using remote sensing have emphasized the amalgamation of many datasets, including crop-related information, weather patterns, pesticide and chemical usage, 3D point cloud data, and hyperspectral and multispectral data (de Oliveira et al., 2022). Only one study has combined a remote-sensing-derived crop growth index with online Vis-NIR spectroscopy soil fertility data to predict crop yield (Pantazi et al., 2016). To the best of our knowledge, no study is available on the data fusion of the two sensing tools for the prediction of crop yield at the MZ level under different rates of N fertilization.

Further research is required to develop effective methods for integrating high-resolution RS data with corn yield prediction models at the

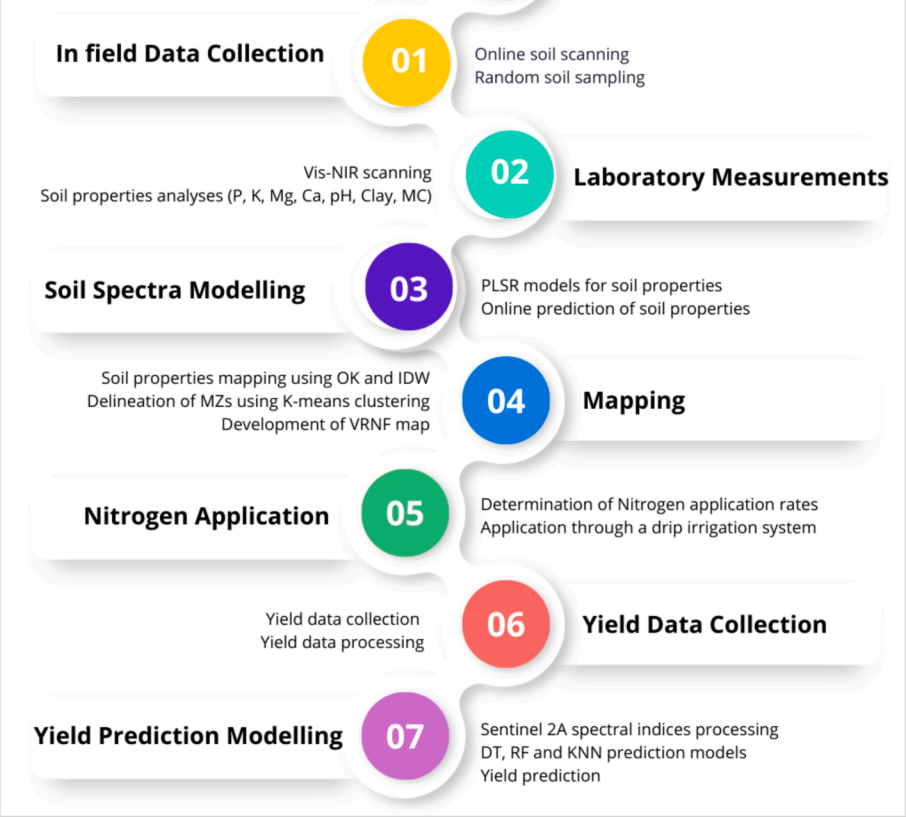


Fig. 2. Study methodology flowchart. Vis-NIR, visible and near infrared; P, Phosphorus; K, Potassium; Mg, Magnesium; Ca, Calcium; MC, Moisture Content; VRNF, Variable Rate Nitrogen Fertilization; OK, ordinary kriging; IDW, inverse distance weighting; PLSR, partial least squares regression; DT, Decision Tree; KNN, K-Nearest Neighbors; RF, Random Forest.

management zone level. This study aimed to evaluate the feasibility of accurately predicting corn (*Zea mays*) yield within specific management zones (MZs) using a combination of online collected Vis-NIR soil spectroscopy data, remote sensing-derived spectral indices, and machine learning algorithms.

2. Materials and methods

2.1. Experimental site

A commercial farm situated in Karacabey, Bursa Province, Turkey (40°08′22.61″N, 28°22′58.75″E) was used for this study (Fig. 1). The cultivated field covered an area of 8.1 ha, with corn as the planted crop, during the 2022 growing season. The soil type in this field varies between sandy clay and loamy clay. The field is typically used for the rotation of wheat and corn crops. The climate in the study area aligns with the Marmara transition-type Mediterranean climate. Precipitation primarily occurs during spring and winter, often in the form of snow during winter months. February tends to be the coldest month, whereas July is typically the hottest. The region experiences an annual average temperature of 14.6 °C (Anonymous, 2023).

2.2. Methodology

In this study, management zones (MZs) were delineated using an online Vis-NIRS sensing of key soil fertility attributes, with the aim of applying variable-rate nitrogen in a drip-irrigated corn field. Sentinel-2A-derived spectral indices included the leaf area vegetation index (LAI), normalized difference red edge vegetation index (NDRE), normalized difference vegetation index (NDVI), normalized difference water index (NDWI), and soil-adjusted vegetation index (SAVI). To assess the feasibility of accurate yield prediction within MZs, machine-learning regression algorithms, specifically Decision Tree (DT), Random Forest (RF), and K-Nearest Neighbors (KNN), were employed. The overall methodology and steps used in this study are illustrated in Fig. 2.

2.3. Online measurements and soil sampling

An online multi-sensor platform utilizing a Vis-NIR spectrophotometer (Tec5 Technology for Spectroscopy, Germany) was used for soil spectra acquisition (Mouazen, 2006). This platform encompasses a subsoiler mechanism designed to penetrate the soil and establish a trench at a depth of 15–25 cm. Affixed to the rear of the subsoiler chisel is an optical probe linked to a Vis-NIR spectroscopy sensor, with a measurement range spanning from 305 to 1700 nm, utilized for the acquisition of soil spectra in the diffuse reflectance mode. A 100 % ceramic disc served as a white reference and was used at intervals of 30 min to calibrate the measurement device. The multisensor platform incorporated a differential global positioning system (DGPS) (Trimble, ver. CFX-750, USA) on a semi-rugged laptop computer (Toughbook, Panasonic, Belgium). This arrangement facilitated the storage of the collected soil spectra and DGPS readings at a frequency of 1 Hz using the MultiSpec pro-II software (Tec5 Technology for spectroscopy, Germany). The platform was affixed to a three-point linkage of the tractor.

The online soil measurements were performed in June 2021. Throughout the measurement process, the subsoiler was drawn along parallel lines spaced 12 m apart, while maintaining an average speed of 3 km/h. In conjunction with real-time measurements, 110 random soil samples were collected from the undisturbed bottom of the trench formed by the subsoiler, situated at an approximate depth of 15 cm. These samples were used for the measurement of soil attributes using

Table 1
Descriptive statistics of measured soil properties.

Soil property	Min	Max	Mean	Median	Std	SC
pH	6.60	7.90	7.277	7.250	0.32	0.031
P (cmol/kg)	5.63	42.37	18.85	18.27	6.89	0.226
K (cmol/kg)	78.80	258.90	161.15	153.15	43.12	0.377
Mg (cmol/kg)	132.30	531.50	310.66	320.95	87.24	0.003
Ca (cmol/kg)	542.90	3366.70	28.29	1488.90	766.47	0.482
MC (%)	7.74	28.29	17.27	18.73	4.49	−0.395
Clay (%)	24.80	44.80	36.18	37.50	5.27	−0.570

MC, moisture content; Min, minimum value; Max, maximum value; Std, standard deviation; SC, Skewness coefficient.

laboratory methods, which are necessary to calibrate the Vis-NIRS sensor.

2.4. Laboratory soil Vis-NIR scanning and chemical analysis

After thorough mixing and elimination of visible roots, stones, or crop residues, each soil sample was subjected to a reduction process using the quartering method (Mukhopadhyay and Maiti, 2018). The samples were subsequently reduced to subsamples weighing 150 g and then subjected to Vis-NIR scanning in a laboratory setting. In the laboratory Vis-NIR scanning procedure, each sample was evenly distributed across three Petri dishes, each measuring 5.5 cm in diameter and 1.5 cm in depth. Following the distribution, the soil surface was compacted using a spatula to optimize the signal-to-noise ratio, according to the methodology outlined by Mouazen et al. (2005). Spectral measurements were conducted in a dark environment using the same Vis-NIR spectroscopy sensor used for the online measurements. Each Petri dish underwent ten consecutive scans, and a consolidation of 30 soil spectra was averaged to derive a singular spectrum representative of each soil sample.

Soil samples collected during online scanning were analyzed for physical and chemical properties in a commercial accredited laboratory (Laben, Antalya, Turkey). Ammonium acetate-ICP–the exchangeable ion content of the soils, potassium (K), calcium (Ca), and magnesium (Mg) were determined using an Eppendorf Elex 6361 model flame photometer (Pratt, 1965). Soil pH was determined using the saturation method and the true pH values were determined using the method described by Thomas (1996). The particle size distribution (Clay, Silt, Sand) was determined using the Bouyoucos hydrometer method (Gee et al., 1986). Descriptive statistics of the collected soil samples are presented in Table 1.

2.5. Vis-NIR spectra modelling and online soil properties prediction

Fresh soil samples scanned in the laboratory using Vis-NIRS were used for modelling and validation. To fortify the Vis-NIR models, an extra 92 samples from the existing laboratory spectral library were incorporated, resulting in a comprehensive dataset of 202 samples for utilization in partial least square regression (PLSR) analyses. These extra samples were collected from the same region of Karacabey and had similar soil types and textures. Different spectra preprocessing steps were used for the soil property modelling. Pre-processing of the Vis-NIR spectra involved a series of steps, including the sequential correction of the spectral jump between the two detectors of the tec5 spectrometer (Mouazen et al., 2009), moving average to mitigate spectral noise, and maximum normalization to standardize the spectral data within a 0 – 1 scale, thereby ensuring an even distribution of data variance. Additional steps include the Savitzky-Golay first derivative to enhance the signal-to-noise ratio and Savitzky-Golay smoothing to diminish the noise in the spectral data (Barnes et al., 1989).

The pre-processed spectra, along with the measured soil reference values, were merged into a data matrix. This matrix was subdivided into calibration (80 %) and validation (20 %) datasets using the Kennard–-Stone algorithm (Kennard and Stone, 1969). Partial Least Squares Regression (PLSR) models for soil properties were formulated using the calibration set through leave-one-out cross validation. Upon attainment of satisfactory model performance, the complete online Vis-NIR spectra obtained from the field were employed to predict soil properties. Both spectra pretreatment and model establishment were performed with RStudio software (RStudio Inc., Boston, MA, USA), using open-source libraries (Stevens and Ramirez Lopez, 2014). Prior to PLSR

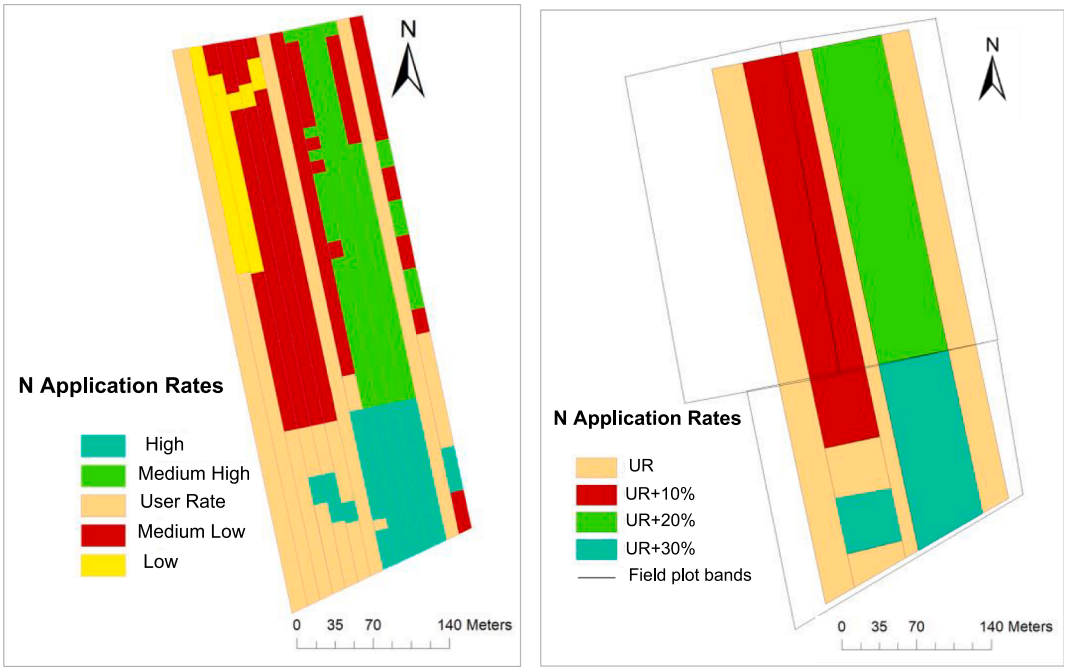


Fig. 3. (a) Theoretical variable rate nitrogen fertilization (VRNF) application map consisting of five N rates (representing five management zones), according to the Robin Hood approach (feeding more the poor) and (b) actual VRNF application map.

Table 2
Sentinel 2A derived crop spectral vegetation indices used in this study.

Index	Formula	Source
NDVI	(Band8 – Band4 / (Band8 + Band4)	Tucker (1979)
NDWI	(Band3 – Band8) / (Band3 + Band8)	Gao (1996)
SAVI	((Band8 – Band4) / (Band8 + Band4 + 0.5)) * (1 + 0.5)	Huete (1988)
LAI	3.618*[2.5 *((Band8 – Band4) / ((Band8) + (6*Band4) – (7.5*Band2) + 1))] – 0.118	Boegh et al. (2002)
NDRE	(Band8 – Band5) / (Band8 + Band5)	Hardisky et al. (1983)

NDVI, normalized difference vegetation index; NDWI, normalized difference water index; LAI, leaf area index; SAVI, soil adjusted vegetation index; NDRE, normalized difference red-edge vegetation index.

Table 3
Hyperparameters used for modelling in this study.

Models	Hyperparameters
Random Forest	{criterion='squared_error',max_depth = None, max_features = 1.0, max_leaf_nodes = None,max_samples = None, min_samples_leaf = 1, min_samples_split = 2, min_weight_fraction_leaf = 0.0,n_estimators = 100,n_jobs = None, oob_score = False, random_state = 7}
K-Nearest Neighbors	{criterion='squared_error', max_depth = None,max_features = None, max_leaf_nodes = None, min_impurity_decrease = 0.0, min_samples_leaf = 1, min_samples_split = 2, min_weight_fraction_leaf = 0.0, random_state = None, splitter='best'}
Decision Tree	{algorithm='auto', leaf_size = 30, metric='minkowski', metric_params = None, n_jobs = None, n_neighbors = 5,p = 2, weights='uniform'}

application, principal component analysis (PCA) was conducted to explore the similarities or distinctions among the soil samples. The performance of the resulting models was validated using samples scanned in the laboratory. Model accuracy was evaluated based on the coefficient of determination (R^2), root mean square error (RMSE), residual prediction error (RPD), mean absolute error (MAE), ratio of performance to inter-quartile distance (RPIQ), and latent variable (LV).

2.6. Delineation of management zones and nitrogen (N) recommendation maps

For the delineation of management zones (MZs) and creating recommendation maps for VRNF, the online collected high-resolution spatial data for predicted soil pH, P, Mg, K, Clay, and Ca were utilized. These predictions were derived from online-collected Vis-NIR spectra. The high-resolution data for the soil properties were normalized and fused using k-means clustering with RStudio software (RStudio Inc., Boston, MA, USA) to divide the field into a limited number of clusters, each having similar soil and crop characteristics. The fused data were then classified into three, four, and five MZs classes for each fertility map to facilitate the application of VRNF.

Subsequently, recommendation maps for N fertilization were generated based on the information associated with the delineated MZs. The fertilizer recommendation maps, derived from the five-class fertility management zones (Fig. 5c), aligned with the five-class MZs (for example, low (L), medium low (ML) medium (M), medium high (MH), and high (H) fertility zones for the VRNF. The development of the recommendation maps involved dividing the field into parallel strips of 12.5 m wide, accommodating UR and VRNF treatments randomly. Each of the VRNF strip was divided into plots representing the different MZs. Given that the partner farmer had already applied K and P fertilizers during corn seeding, no variable P or K fertilizers were applied. In the case of N application maps, the Robin Hood fertilization scheme was considered. This scheme involves applying more fertilizer to the least fertile zones (e.g., L and ML) and less fertilizer to the most fertile zones (e.g., H and MH), according to the principles outlined by Guerrero and Mouazen (2021). The M fertility zone receives the farmer application rate, which is the UR. In this study, it was assumed that the fluctuation in N rates ranged between +30 % and −30 % of the farmer prescribed the UR treatment. The user's N application rate was determined based on the farmer's experience, traditional soil analysis, and nitrate directive guidelines.

The VRNF application map in Fig. 3a was developed based on the available data, using the most appropriate procedures for developing recommendations. In fact, the area of the entire field was approximately 20 ha, of which 8.1 ha were used for the experiment. The drip irrigation method was adopted by dividing the 8.1 ha area into three sections. Accordingly, the longitudinal resolution for irrigation and fertilizer application was the length of irrigation laterals (approximately 130 m). The farmer found it risky to apply lower fertilizer rates than the user rate (310 kg nitrogen per ha) because the potential for high yield could be jeopardized. Therefore, the application of variable-rate fertilizers had been modified accordingly. As a result of collaboration with the farmer, three VRNF rates of UR+10 %, UR+20 %, and UR+30 % kg/ha were applied in addition to the farmer rate (UR) to test the effects of varying fertilizer on yield. The final VRNF map obtained for the actual field application is shown in Fig. 3b. These application rates were manually added because the longitudinal resolution of the laterals could not be changed in drip irrigation lines.

2.7. Data processing and dataset building

2.7.1. Yield data

Corn cultivation commenced on May 16, 2022, with harvesting conducted on November 15, 2022, employing a New Holland CX6.8 80 combine harvester equipped with a New Holland yield monitoring system. Before harvesting in the experimental plots, the corn in the full tanks was weighed in a commercial enterprise and compared with the yield values in three replicates. No calibration was performed on the yield monitor, because the measurement error was approximately 0.5 %. The raw dry yield data showed large positioning errors of up to 4.2 m

Table 4
Prediction results of the models obtained by partial least squares regression (PLSR) analyses in cross-validation and online validation presented for the soil properties [e.g., pH, phosphorous (P), calcium (Ca), moisture content (MC), magnesium (Mg), clay and potassium (K)].

Property	Cross-Validation					Online Validation					LV
	R^2	RMSE	MAE	RPD	RPIQ	R^2	RMSE	MAE	RPD	RPIQ	
pH	**0.60**	0.27	0.21	1.59	2.23	**0.69**	0.25	0.20	1.81	1.83	3
P (cmol/kg)	**0.99**	0.44	0.35	16.30	23.24	**0.66**	4.05	3.00	1.74	2.23	10
Ca (cmol/kg)	**0.85**	383.08	302.9	2.55	4.41	**0.73**	366.78	277.7	1.96	1.61	5
MC (%)	**0.94**	0.81	0.64	4.14	4.73	**0.78**	2.37	1.88	2.17	3.67	8
Mg (cmol/kg)	**0.67**	45.05	35.41	1.76	2.16	**0.66**	52.08	40.93	1.73	2.62	7
Clay (%)	**0.66**	3.03	2.37	1.72	2.54	**0.52**	3.57	3.03	1.47	2.63	2
K (cmol/kg)	**0.79**	28.12	22.75	2.19	3.22	**0.75**	37.90	29.54	2.04	3.74	3

R^2, coefficient of determination; RMSE, root mean square error; MAE, mean square error; RPD, residual prediction error; RPIQ, the ratio of performance to inter-quartile distance; LV, latent variable.

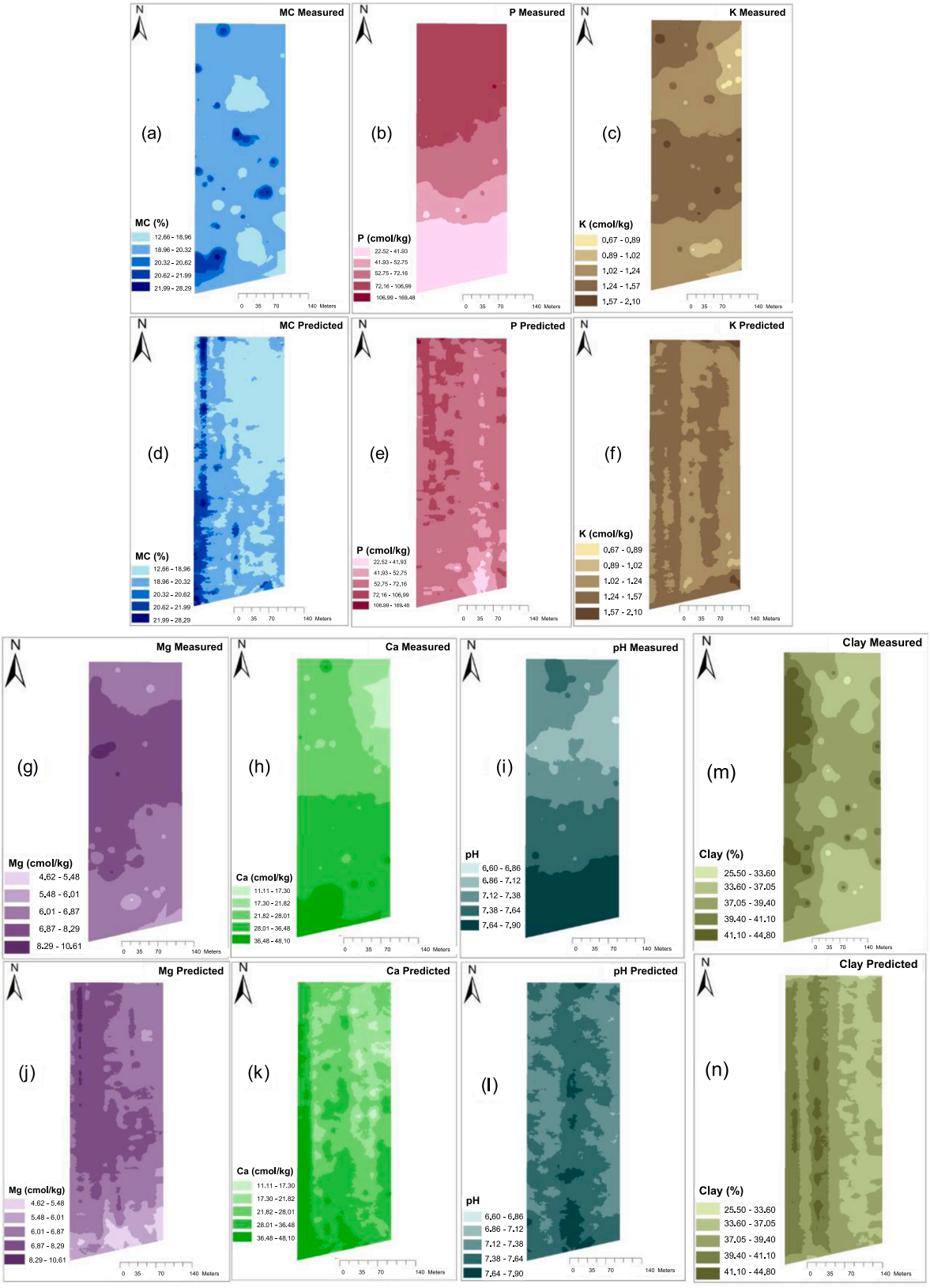


Fig. 4. Comparison between laboratory measured (top maps) and online visible and near infrared (Vis-NIR) predicted (bottom maps) soil properties [e.g., pH, phosphorous (P), calcium (Ca), moisture content (MC), magnesium (Mg), clay and potassium (K)] maps.

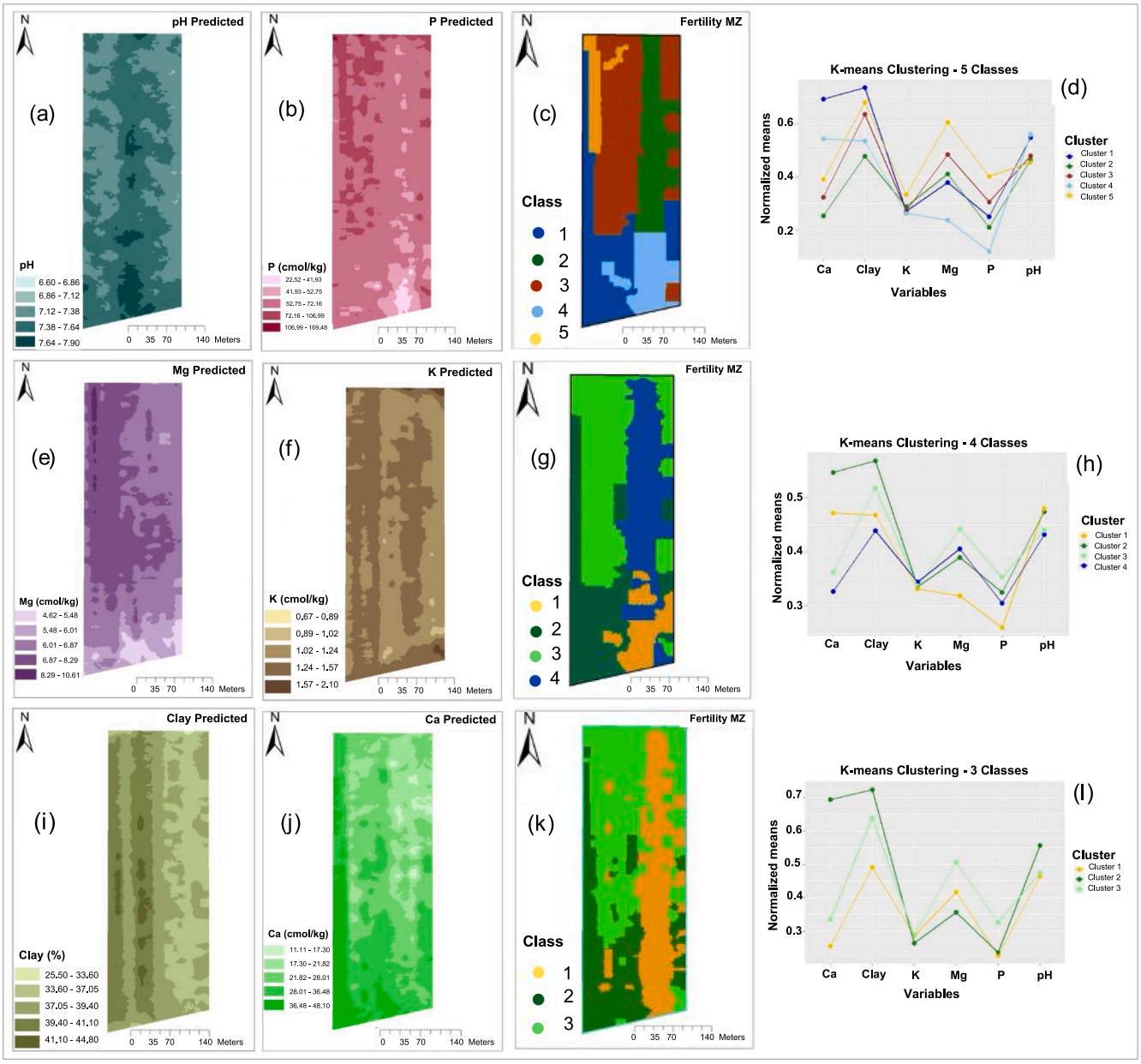


Fig. 5. Delineation of management zones (MZs) by K-means clustering, showing normalized means plot (d, h, l) and corresponding soil fertility MZs map (c, g, k) with 5, 4, and 3 fertility classes respectively.

because only RXT GPS signals were available during data collection. Post-processing was necessary to correct the coordinates of the point data and trim the data around the headlands of the trial plots. The yield data were filtered to eliminate zero (0) values. The final yield dataset had a total of 7878 points for the entire field, with 2945, 1773, 1881, and 1279 points for the N treatments applied, e.g., UR, MZ1 (UR+10 %), MZ2 (UR+20 %), and MZ3 (UR+30 %), respectively, as shown in Fig. 3b. All procedures pertaining to this section were conducted using the ArcGIS ArcMap 10.8 (ESRI, Redlands, CA, USA) software.

2.7.2. *Sentinel 2A-derived spectral indices*

Sentinel-2A-derived vegetation indices have a significant utility in agricultural and environmental monitoring. These indices serve as valuable tools for evaluating the health, growth, and overall productivity of the vegetation. In this study, Sentinel-2A imagery was used to derive vegetation indices. The images were downloaded from the Sentinel 2 products official platform in bottom-of-atmosphere L2A format. Five images from June 18, July 26, August 27, September 26, and October 01, 2022, covering the study area were atmospherically corrected. For each S2 scene, the blue, green, red, NIR, and red-edge spectral bands were selected. Table 2 provides a summary of the spectral indices used in this study along with their corresponding formulae.

To maintain consistency, bands with distinct 20 m resolutions were resampled to a uniform 10 m resolution for the analyses. Concurrently, the yield values retrieved from the combine harvester, recorded at an approximate resolution of 2 m, were stored in a file, along with their corresponding coordinates. To align the resolution, the Sentinel-2A-derived spectral indices maps were resampled to a 2 m resolution, ensuring compatibility with the yield map.

2.8. *Machine learning algorithms*

The final dataset had 7878 rows for the whole field, with 2945, 1773, 1881, and 1279 rows for UR, MZ1, MZ2, and MZ3, respectively,

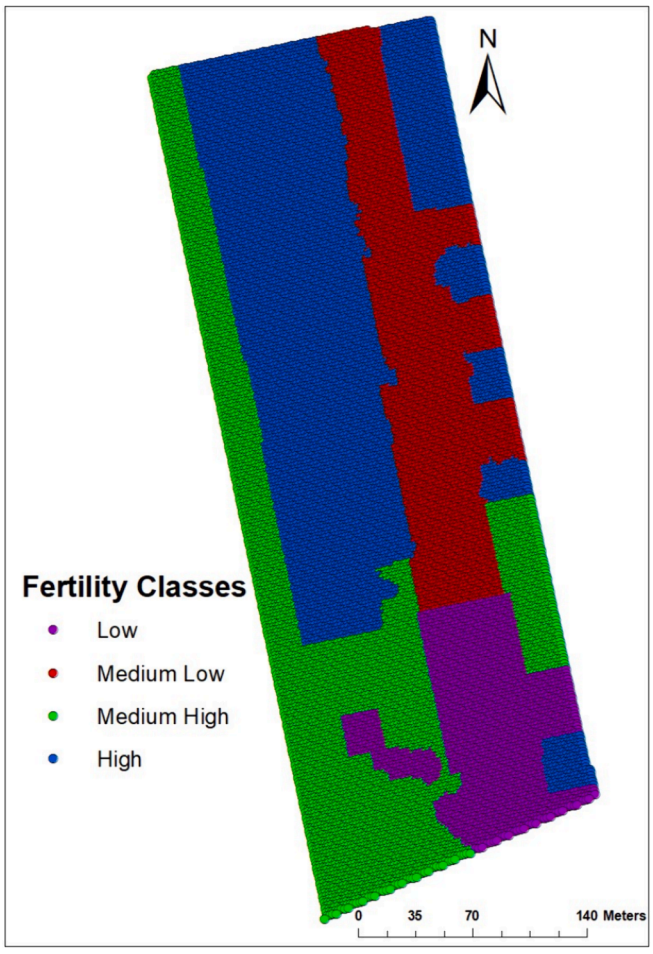


Fig. 6. Management zone map.

containing the crop spectral indices values as the input, and the corresponding yield data as the output was curated for training and testing. Three machine learning regression algorithms, namely Random Forest (RF), K-Nearest Neighbors (KNN), and Decision Tree (DT), were employed in this study. *Random Forest Regressor (RF):* Random Forest regressor is an ensemble learning method that combines the predictions of multiple decision trees to improve the accuracy and robustness of the model. It has been extensively validated and utilized for accurate crop yield prediction (Jeong et al., 2016; Everingham et al., 2016; Prasad et al., 2020; Masjkur & Tan, 2020; Shahhosseini et al., 2021), outperforming traditional methods such as multiple linear regression. *K-Nearest Neighbors regressor (KNN):* The machine learning regressor KNN has been recommended and used in crop yield prediction, demonstrating its potential in this application (Huang et al., 2022; Mupangwa et al., 2020; Wei, 2017). *Decision Tree regressor (DT):* DT has been extensively studied and applied in crop yield prediction. It has demonstrated high accuracy and potential for crop yield prediction, making it a valuable tool in agricultural decision making and precision agriculture (Ed-Daoudi et al., 2023; Kalpana et al., 2023; Patil & R, 2017).

2.8.1. Prediction models implementation

The prediction models were trained using Python programming on Google Colaboratory. The dataset was split into training and test sets by using an 80:20 partition. Hyperparameters were set before the learning process began. The hyperparameters for each algorithm used in this study are listed in Table 3. Model accuracy was evaluated in terms of the coefficient of determination (R^2), root mean square error of prediction (RMSE) and mean absolute error (MAE).

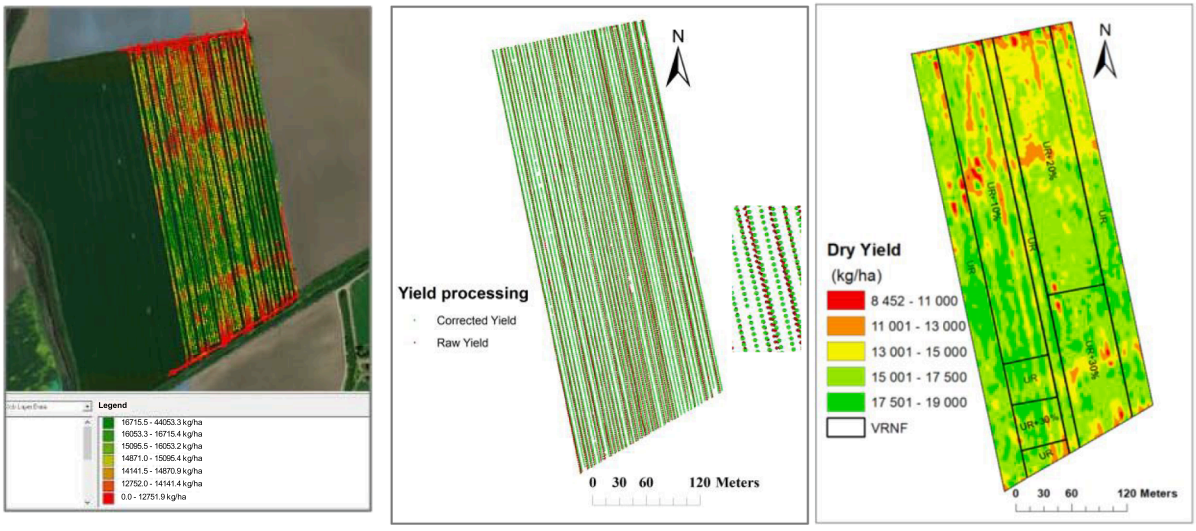
2.9. Mapping

Maps for both the measured and predicted soil properties and yield data were developed using the ArcGIS ArcMap 10.8(ESRI, Redlands, CA, USA) software. The interpolation methods employed for this purpose included inverse distance weighting (IDW) and ordinary kriging (OK) after the semi-variogram analyses.

3. Results and discussion

3.1. Vis-NIR soil spectra models performance in cross-validation and independent validation

The outcomes of the soil Vis-NIR spectra calibration and validation models established through PLSR for the studied soil properties are outlined in Table 4. Models achieved good accuracy values in terms of



(a) (b) (c)

Fig. 7. (a) Raw yield data map with positioning errors, (b) Corrected and non-corrected point data, (c) Reconstructed yield map.

Table 5
Summary of measured corn yield values in kg/ha by management zone (MZ).

MZs	VRNF	Area (ha)	Min	Max	Range	Mean	Std
UR	UR	3.33	8929.43	18362.50	9433.08	15392.26	1312.30
MZ1	UR+10 %	1.75	8452.30	18725.13	10272.83	15129.71	1535.10
MZ2	UR+20 %	1.79	9378.95	18964.62	9585.67	14722.52	1163.30
MZ3	UR+30 %	1.23	9447.83	18504.77	9056.94	15483.43	1323.38

UR, uniform rate; MZ, management zone; VRNF, variable rate of nitrogen fertilization; Min, minimum value; Max, maximum value; Std, standard deviation

Table 6
Descriptive statistics of the normalized corn vegetation indices values used in this study.

Index/Date	Mean	Std	Min	Q1	Q2	Q3	Max
LAI_June18	0.3408	0.0227	0	0.3402	0.3432	0.3445	1
LAI_July 26	0.7569	0.1184	0	0.6993	0.7786	0.8379	1
LAI_August27	0.6807	0.1410	0	0.5919	0.6932	0.7725	1
LAI_September26	0.8424	0.0438	0	0.8442	0.8447	0.8459	1
LAI_October01	0.6820	0.0295	0	0.6801	0.6816	0.6826	1
NDRE_June18	0.5049	0.1763	0	0.4022	0.5090	0.6340	1
NDRE_July 26	0.7033	0.1512	0	0.6079	0.7313	0.8146	1
NDRE_August27	0.5689	0.2012	0	0.4258	0.6159	0.7232	1
NDRE_September26	0.4359	0.2029	0	0.2879	0.4562	0.5721	1
NDRE_October01	0.3772	0.1779	0	0.2535	0.3770	0.4633	1
NDVI_June18	0.4755	0.1633	0	0.3525	0.4805	0.5866	1
NDVI_July 26	0.7792	0.1255	0	0.7080	0.7972	0.8709	1
NDVI_August27	0.4918	0.1618	0	0.3668	0.5284	0.6099	1
NDVI_September26	0.4089	0.1714	0	0.2987	0.4196	0.5210	1
NDVI_October01	0.3308	0.1553	0	0.2282	0.3294	0.4189	1
NDWI_June18	0.4841	0.1852	0	0.3598	0.4734	0.6179	1
NDWI_July 26	0.2662	0.1375	0	0.1664	0.2442	0.3430	1
NDWI_August27	0.4992	0.1553	0	0.3856	0.4688	0.6165	1
NDWI_September26	0.6318	0.1640	0	0.5445	0.6312	0.7358	1
NDWI_October01	0.7007	0.1512	0	0.6286	0.7153	0.8002	1
SAVI_June18	0.4755	0.1633	0	0.3526	0.4805	0.5867	1
SAVI_July 26	0.7792	0.1255	0	0.7080	0.7972	0.8709	1
SAVI_August27	0.4918	0.1618	0	0.3668	0.5284	0.6099	1
SAVI_September26	0.4089	0.1714	0	0.2987	0.4196	0.5209	1
SAVI_October01	0.3308	0.1553	0	0.2282	0.3294	0.4189	1

Min, minimum value; Max, maximum value; Std, standard deviation; Q1, first quartile; Q2,second quartile; Q3, third quartile.

the coefficient of determination (R^2), which ranged from 0.60 to 0.99 in cross-validation and from 0.52 to 0.78 in online validation. The good performance of the models in the online validation was expressed by RPD and RPIQ, which ranged between 1.47–2.17 and 1.61–3.74, respectively. The accuracy of the models in the online validation was higher for some soil properties, including MC, K, and Ca, as compared to the model accuracy for pH, P, Mg and Clay. There are many factors that may affect the model accuracy. Firstly, MC is considered as one of the primary properties in the NIRS, as it has a direct spectral response in the NIRS range (Kuang et al., 2012; Stenberg et al., 2010). Another possible reason for the high accuracy of some soil properties could be the high spatial variability in this particular soil property, and vice versa. High modelling accuracy also indicates a linear relationship between the online spectral and measured properties, as PLSR is a linear multivariate analysis (Kuang et al., 2012). Moreover, different pre-processing methods and number of samples (Kuang and Mouazen, 2012) used for calibration also affect the modelling accuracy.

3.2. Mapping of soil properties

Maps illustrating the laboratory-measured (top maps) and predicted soil properties using the online collected Vis-NIR spectra (bottom maps) are shown in Fig. 4. These maps not only show accurate predictions and recognizable spatial similarities, but also underscore the quality of online maps, which is crucial for delineating accurate MZs maps.

3.3. Management zone maps

MZs developed based on high-resolution spatial data of the predicted soil pH, P, Mg, K, Clay, and Ca are shown in Fig. 4. Noteworthy similarities were observed in the spatial distributions of the MZs and the corresponding soil property maps. The P, K and Mg normalized means were used to rank the classes in terms of fertility, resulting from the clustering analysis. The class with the highest of these three values was defined as the most fertile. The fertility MZs map with five classes showed that class 5 was the most fertile, as it had the highest normalized means of P, K, and Mg, followed by classes 3, 1, 2, and 4. The fertility MZ maps with four and three classes showed that class 3 was the most fertile and class 1 was the least fertile (Fig. 5).

The k-means cluster analysis that best explained the spatial variability was based on five fertility classes (Fig. 5c), which is the reason why this MZ map was adopted for further analysis in this study. Because of the lack of historical yield data in the study area, the five zones were classified into low, medium–low, medium–high, and high-fertility zones based on soil fertility parameters, in addition to the farmer's knowledge and experience in their field (Fig. 6).

3.4. Measured yield map

The measured yield map using the combine harvester on-board yield sensor delineated low- and high-yielding areas with clear distinctions in the field, as illustrated in Fig. 7. Notably, approximately one-third of the plots along the northern edge consistently exhibited lower yields. There were also some small segments with low yields at the opposite edge, the reasons for which were unclear. Weed infestations (*Xanthium strumarium*) occurred at the very edge of the northern end of the plot, which can explain some of the yield reductions. Some patches were also affected by weeds, particularly in the southwestern part of the plot.

Table 7
Performance of the best independent validation results for machine learning (ML) models used for the prediction of corn yield at the management zone and the whole field area.

ML Models	Whole Field		
	R^2	RMSE (kg/ha)	MAE
K-Nearest Neighbors	0.50	1015.68	734.68
Random Forest	0.53	1000.75	727.90
Decision Tree	0.54	963.02	708.09
	UR		
	R^2	RMSE (kg/ha)	MAE
K-Nearest Neighbors	0.50	970.52	716.67
Random Forest	0.55	921.59	679.11
Decision Tree	0.53	941.77	698.46
	MZ1		
	R^2	RMSE (kg/ha)	MAE
K-Nearest Neighbors	0.51	1096.93	877.31
Random Forest	0.53	1134.73	875.42
Decision Tree	0.51	1125.81	874.73
	MZ2		
	R^2	RMSE (kg/ha)	MAE
K-Nearest Neighbors	0.66	686.38	519.31
Random Forest	0.65	694.85	520.43
Decision Tree	0.71	672.79	506.22
	MZ3		
	R^2	RMSE (kg/ha)	MAE
K-Nearest Neighbors	0.55	957.72	726.43
Random Forest	0.55	967.67	713.63
Decision Tree	0.57	928.24	665.58

UR, uniform rate; MZ, management zone.

However, visual observations of weed pressure at these locations during the growing/development stages of corn did not necessarily agree with yield values. The farmer also confirmed that he expected more low-yielding subregions at the lower end of the map. It can be assumed that the low yield at the southern edge of the field is attributed to low P and Mg, as illustrated by the corresponding maps in Fig. 4.

3.4.1. Analysis of yield data

The descriptive statistics for the measured corn yield associated with each management zone (MZ) are presented in Table 5. Upon comparing corn yield between MZs, it was observed that MZ3 exhibited the highest mean yield (15483.43 kg/ha), followed by UR (15392.26 kg/ha), MZ1 (15129.71 kg/ha), and MZ2 (14722.52 kg/ha). The minimum yield (8452.30 kg/ha) was recorded in MZ1, while the maximum yield (18964.62 kg/ha) was observed in MZ2.

Increasing the amount of fertilizer did not significantly increase the corn yield. Receiving UR+20 %, the MZ2 showed the largest mean yield reduction compared to the yield value in the UR. A similar observation can be made for the MZ1, which received UR+30 %, although the reduction in yield was not as significant as that of the MZ2. As a result, the farmer seemed to have applied the correct amount of fertilizer at the field scale, although adding 10 % extra N to the MZ3 has increased the yield compared to the UR treatment. Previous studies have shown that increasing the N rate over the crop demand in respective to the N response curve could have a diverse effect on the yield of barley and wheat due to logging (Guerrero et al., 2021a,b; Telkar et al., 2017). Applying less than the user rate could have shown the effects of variable-rate fertilization, which could not be done in this study due to farmer's decision.

3.5. Yield prediction models performance

The descriptive statistics of the vegetation indices used in this study are presented in Table 6. The spectral indices had higher values during the growing stage and lower values during the early and late growth stages of corn, which has been typically observed in other studies.

The performance metrics, including R^2, RMSE, and MAE, for the three ML regression algorithms employed in this study to predict corn yield are presented in Table 7. Using the five studied spectral indices (NDVI, NDWI, SAVI, LAI, and NDRE), the ML models showed moderate accuracy in predicting corn yield at the MZs level. The models exhibited the highest accuracy for MZ2 ($R^2 = 0.71$, RMSE = 672.79 kg/ha, MAE = 506.22), which was obtained using the Decision Tree modelling technique. Both the UR and MZ1 site-specific models exhibited accuracy levels that were comparable to those of the whole-field model. The

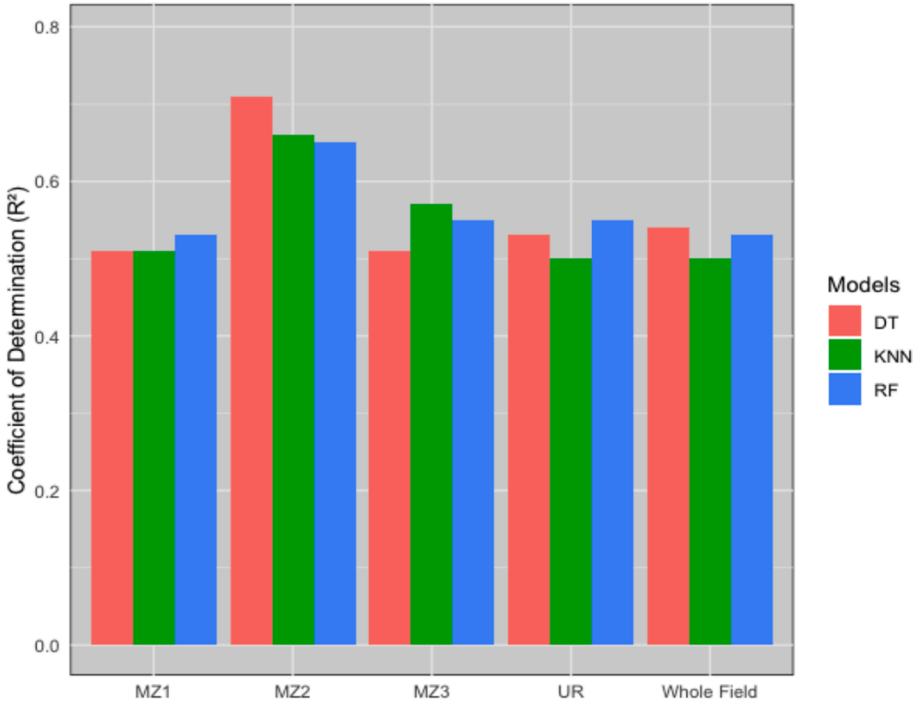


Fig. 8. Comparison of corn yield prediction models by management zone (MZ); DT, Decision Tree; KNN, K-Nearest Neighbors; RF, Random Forest.

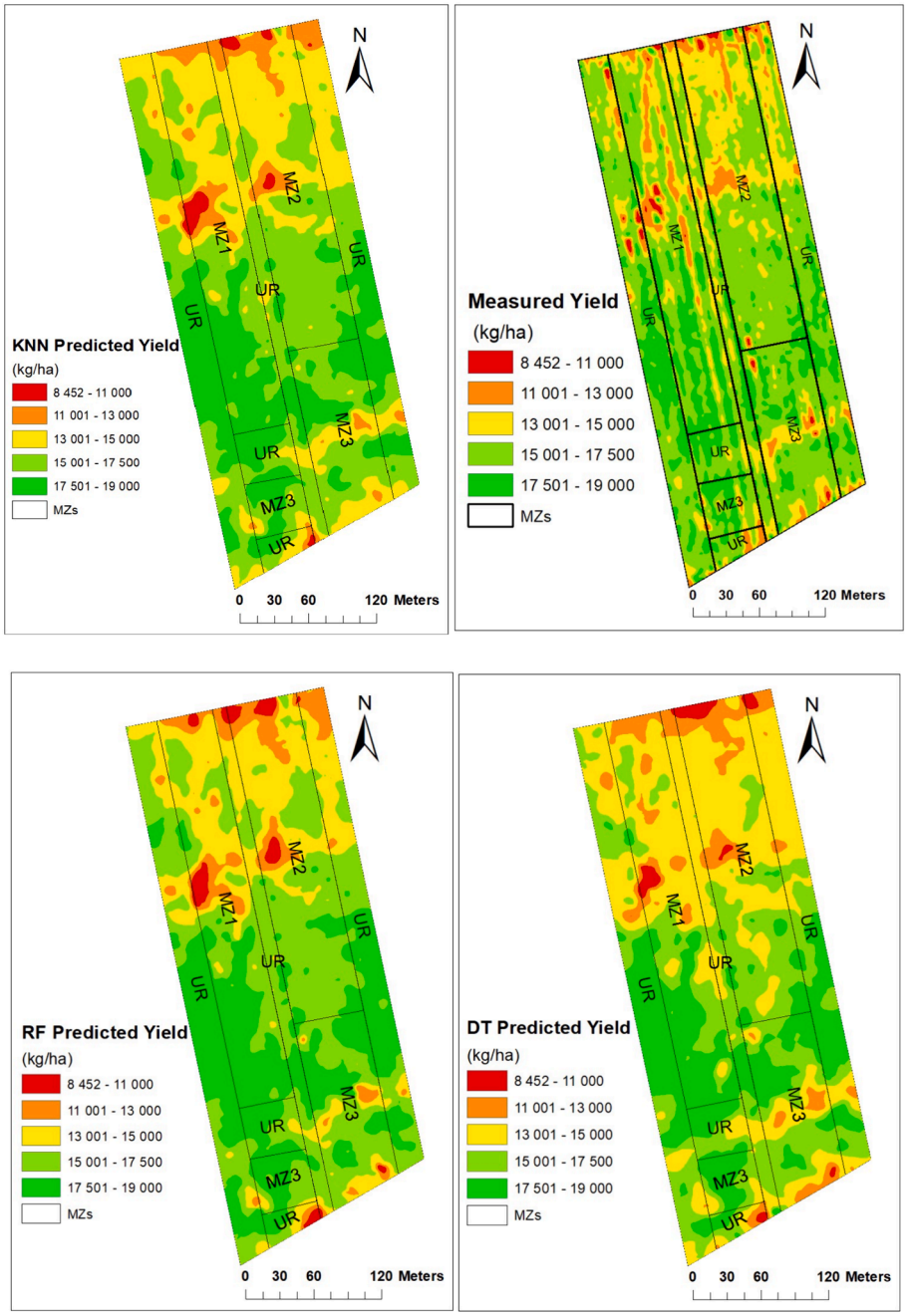


Fig. 9. Comparison between measured yield (a) and machine learning models predicted maps by (b) K-Nearest Neighbors model, (c) Random Forest model, and (c) Decision Tree model.

exception was MZ2, which received an N rate of UR+20 %, where the accuracy was higher than those of the corresponding models for the whole field and the MZ1 and MZ3 models (Fig. 13). This result was consistent with the lowest yield obtained for MZ2. The inclusion of several crop vegetation indices in the ML models could be the reason for the high accuracy obtained for MZ2, as these different indices have captured the influences of different yield limiting factors. The Decision Tree models demonstrated superior accuracy in predicting corn yield (R^2 spanning from 0.51 to 0.71), followed by the random forest models showing moderate accuracy, with R^2 spanning from 0.53 to 0.65, and the K-Nearest Neighbors model performance ranged between 0.50 and 0.66 (Fig. 8).

The measured and predicted yield maps obtained from these models are shown in Fig. 9. Noteworthy similarities are observed between the spatial distributions of the measured and predicted yield maps. The spatial patterns captured in these maps validate the accuracy of our prediction models and highlight the feasibility of leveraging machine learning techniques for yield forecasting at the management zone level derived from the fusion of data on key soil fertility parameters collected by the online Vis-NIR sensor platform (Mouazen, 2006).

Numerous studies have focused on yield predictions using RS data at the field, county, region, or country levels (Jeong et al., 2016; Peng et al., 2018; Gündoğdu and Bantchina, 2018; Hao et al., 2019; Mupangwa et al., 2020; Zhao et al., 2020; Sagan et al., 2021; Shahhosseini et al., 2021). However, the ultimate objective of precision agriculture is to study the spatial and temporal variability within fields. Researchers have also focused on predicting within-field crop yields to improve fertilizer management strategies (Ziliani et al., 2022; Helber

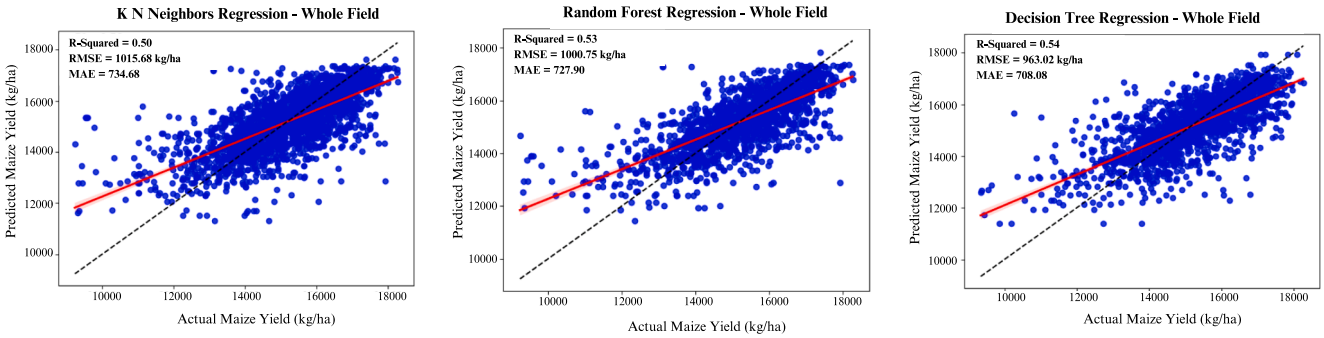


Fig. 10. Scatter plots of predicted versus measured yield using K-Nearest Neighbors regression, Random Forest regression, and Decision Tree regression models independent validation for the whole field.

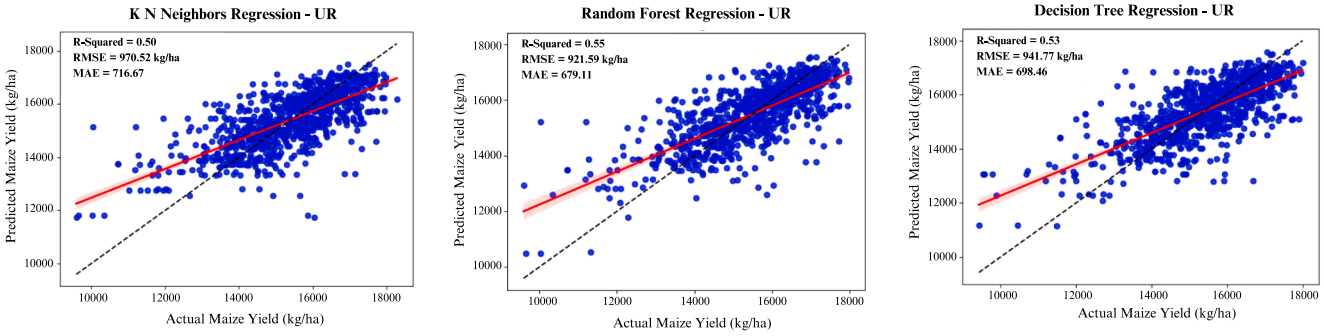


Fig. 11. Scatter plots of predicted versus measured yield using K-Nearest Neighbors regression, Random Forest regression, and Decision Tree regression models independent validation for the UR.

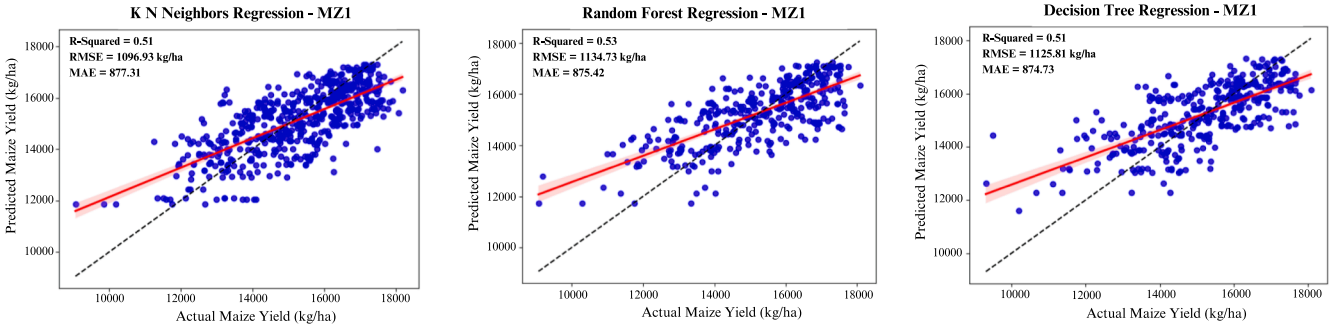


Fig. 12. Scatter plots of predicted versus measured yield using K-Nearest Neighbors regression, Random Forest regression, and Decision Tree regression models independent validation for the MZ1.

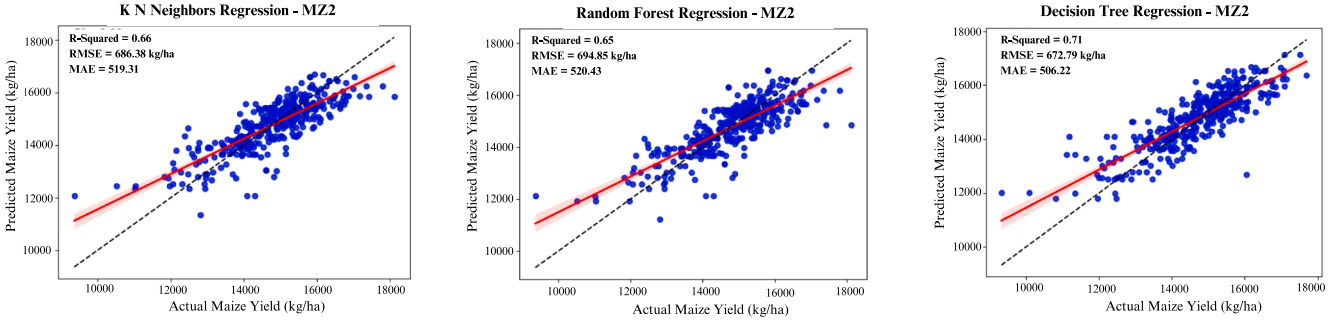


Fig. 13. Scatter plots of predicted versus measured yield using K-Nearest Neighbors regression, Random Forest regression, and Decision Tree regression models independent validation for the MZ2.

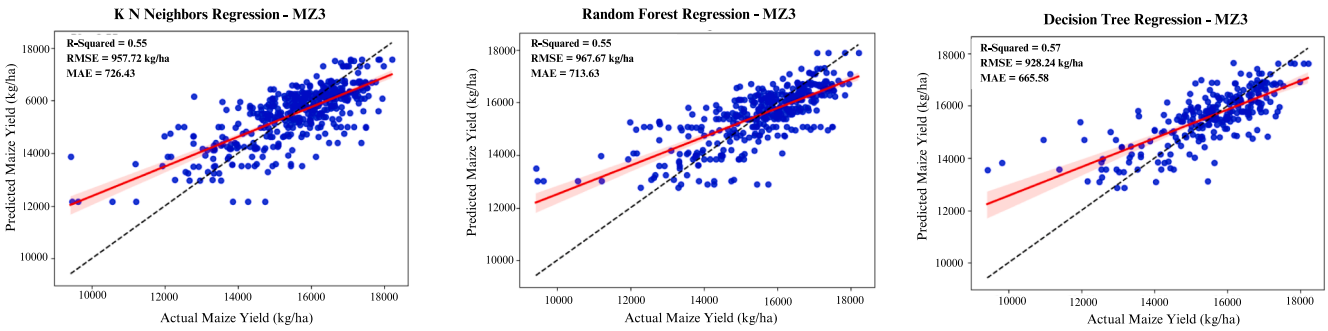


Fig. 14. Scatter plots of predicted versus measured yield using K-Nearest Neighbors regression, Random Forest regression, and Decision Tree regression models independent validation for the MZ3.

et al., 2023). Tagarakis and Ketterings (2017) applied five different nitrogen rates from 0 to 336 kg/ha and concluded that yield prediction models should be applied at the V7 stage in corn because they found the best in-season estimated yield prediction (INSEY) after the V6 stage (R^2 = 0.77) in field-scale evaluations. They stated that INSEY-based equations were better (R^2 = 0.72–0.78) than the raw NDVI measurements (R^2 = 0.56). In our study, a similar R^2 value was found (0.50–0.54) using Random Forest, K-Nearest Neighbors, and Support Vector Machines techniques between the corn vegetative V6 stage and the regenerative R5 stage. Kayad et al. (2019) studied the in-field corm grain yield spatial variability of a corn field using Sentinel-2 imagery (GNDVI) and determined the correlations with the measured yield values using Random Forest and Support Vector Machines, reporting the greatest R^2 of 0.6 obtained using Random Forest. In the current study, a similar result was obtained with R^2 values ranging from 0.53 to 0.65 using the same ML algorithm. Hunt et al. (2019) considered 34 fields to estimate crop yields using Sentinel-2 derived indices and a Random Forest model. They found varying degrees of ability to predict the yield from one field to another, with RMSE values ranging from 0.28 to 1.94 t/ha while the approximate RMSE ranges were 0.96–1.02 t/ha and 0.67–1.13 t/ha for whole field and the MZs, respectively.

To evaluate the performance of the models, the predicted corn yield derived from the resultant model during the validation phase was compared to the observed yield. Figs. 10-14 displays scatter plots depicting the observed and predicted corn yields for various scenarios, including the entire field, UR, MZ1, MZ2, and MZ3.

The findings are deemed acceptable as they align comparably with the findings documented in the relevant literature. de Oliveira et al. (2022) obtained a better accuracy of R^2 between 0.48 and 0.85, by combining some spectral bands, Topographic Wetness Index (TWI), and Topographic Position Index (TPI) using auto-machine learning algorithms for corn yield prediction at the MZ level. Similarly, Shahhosseini et al. (2021) incorporated remote-sensing data and crop phenology to predict corn yield at the county level, thereby demonstrating the potential of ML models for yield prediction. Additionally, Tagarakis and Ketterings (2017) emphasized the importance of proximal sensing for predicting the yield and nitrogen responsiveness of corn, particularly in the context of precision agriculture. The integration of remote sensing data with ML techniques has shown promise in improving yield estimation performance, as demonstrated by Hao et al. (2019), who utilized meteorological indices and phenology-based remote sensing to enhance yield estimation and capture the spatiotemporal variability of environmental conditions for corn growth.

This study cautioned against the application of cross-model techniques as the models were specifically tailored for distinct soil fertility classes and should be confined to the regions for which they were developed. Furthermore, this research suggests exploring alternative spectral indices and employing diverse deep learning or machine learning techniques to improve the precision of the prediction models.

4. Conclusion

The implementation of variable-rate fertilization is crucial for addressing spatial variability in soil characteristics and crop productivity. This approach is essential for informed decision-making regarding the application of soil inputs, particularly variable-rate fertilization, to optimize agricultural outcomes. The developed models showed their ability to effectively forecast corn yield at the MZs level, with acceptable accuracy. Future investigations should prioritize improving precision and rigorously evaluating models across diverse topographic indices to determine whether they exhibit patterns analogous to spectral indices. In the current study, the MZs were based only on soil properties (pH, P, K, Ca, Clay, Mg), whereas other factors, such as soil moisture holding capacity and topography, may also be included during the MZ generation process because such factors may have a significant effect on corn yield.

CRediT authorship contribution statement

Bere Benjamin Bantchina: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Muhammad Qaswar:** Writing – review & editing, Methodology, Investigation, Data curation. **Selçuk Arslan:** Writing – review & editing, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Yahya Ulusoy:** Writing – review & editing, Data curation. **Kemal Sulhi Gündoğdu:** Writing – review & editing, Supervision, Methodology. **Yücel Tekin:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Conceptualization. **Abdul Mounem Mouazen:** Writing – review & editing, Validation, Supervision, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of competing 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.

Data availability

Data will be made available on request.

Acknowledgments

Funding was provided by ICT-AGRI-FOOD 2019 Joint Call of the European Union's Horizon 2020 research and innovation programme, for "A Data-Driven Platform for Site-Specific Fertigation" (ADDFerti) Project under grant agreement no 862665 ICT-AGRI-FOOD. The authors acknowledge the financial support received from the Scientific and

Technological Research Council of Turkey (TÜBITAK), Turkey Contract No. 120N787, from the Fonds Wetenschappelijk Onderzoek (FWO), Belgium Contract No. S007621N, for ADDFerti Project, and from the Wetenschappelijk Onderzoek Vlaanderen (FWO), Belgium for the Post-doctoral Project Contract No. 1282923N. The authors would like to thank Karaca Farm for their cooperation with this project.

References

Anonymous, 2023. MGM, Türkiye İklimi, 2023. Meteoroloji Genel Müdürlüğü. http://www.mgm.gov.tr, Accessed on 10 December 2023.

Bai, J., Tian, M., Li, J., 2022. Control system of liquid fertilizer variable-rate fertilization based on beetle antennae search algorithm. Processes 10 (2), 357. https://doi.org/10.3390/pr10020357.

Barnes, R.J., Dhanoa, M.S., Lister, S.J., 1989. Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. Appl. Spectrosc. https://doi.org/10.1366/0003702894202201.

Barzin, R., Pathak, R., Lotfi, H., Varco, J., Bora, G., 2020. Use of UAS multispectral imagery at different physiological stages for yield prediction and input resource optimization in corn. Remote Sens. (Basel) 12 (15), 2392. https://doi.org/10.3390/rs12152392.

Boegh, E., Soegaard, H., Broge, N., Hasager, C., Jensen, N., Schelde, K., & Thomsen, A., 2002. Airborne Multi-spectral Data for Quantifying Leaf Area Index, Nitrogen Concentration and Photosynthetic Efficiency in Agriculture. Remote Sensing of Environment, 81, no. 2-3.179-193.

Chlingaryan, A., Sukkarieh, S., Whelan, B., 2018. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. Comput. Electron. Agric. 151, 61–69.

de Oliveira, M.F., Ortiz, B.V., Morata, G.T., Jiménez, A.-F., de Souza Rolim, G., da Silva, R.P., 2022. Training machine learning algorithms using remote sensing and topographic indices for corn yield prediction. Remote Sens. 14, 6171. https://doi.org/10.3390/rs14236171.

del Cerro, J., Cruz Ulloa, C., Barrientos, A., de León Rivas, J., 2021. Unmanned aerial vehicles in agriculture: A survey. Agronomy 11 (2), 203.

Duffera, M., White, J.G., Weisz, R., 2007. Spatial variability of Southeastern U.S. Coastal Plain soil physical properties: Implications for site-specific management. Geoderma 137, 327–339.

Ed-Daoudi, R., Alaoui, A., Ettaki, B., & Zerouaoui, J., 2023. Improving crop yield predictions in morocco using machine learning algorithms. Journal of Ecological Engineering, 24(6), 392-400. doi: 10.12911/22998993/162769.

Everingham, Y., Sexton, J., Skocaj, D., Inman-Bamber, N., 2016. Accurate prediction of sugarcane yield using a random forest algorithm. Agron. Sustain. Dev. 36 (2) https://doi.org/10.1007/s13593-016-0364-z.

Gao, B.-C., 1996. NDWI - A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sens. Environ. 58, 257–266.

Gee, G.W., Bauder, J.W., & Klute, A., 1986. Particle-Size Analysis, Methods of Soil Analysis, Part 1. Physical and Mineralogical Methods, Soil Since Society of America. Inc., Madison, WIS, USA.

Geipel, J., Link, J., Claupein, W., 2014. Combined spectral and spatial modeling of corn yield based on aerial images and crop surface models acquired with an unmanned aircraft system. Remote Sens. (Basel) 6 (11), 10335–10355.

Guerrero, A., De Neve, F., Mouazen, A.M., 2021a. Data fusion approach for map-based variable-rate nitrogen fertilization in barley and wheat. Soil Tillage Res. 205, 104789.

Guerrero, A., De Neve, S. and Mouazen, A.M., 2021b. Current sensor technologies for in-situ and online measurement of soil nitrogen for variable rate fertilization – A review, Advances in Agronomy. ISSN 0065-2113. doi: 10.1016/bs.agron.2021.02.001.

Guerrero, A., Mouazen, A.M., 2021. Evaluation of variable rate nitrogen fertilization scenarios in cereal crops from economic, environmental and technical perspective. Soil Tillage Res. 213, 105110 https://doi.org/10.1016/j.still.2021.105110.

Gündoğdu, K.S., Bantchina, B.B., 2018. Landsat Uydu Görüntülerinden NDVI Değer Dağılımının Parsel Bazlı Değerlendirilmesi, Bursa Uludağ Üniversitesi Ziraat Fakültesi Çiftlik Arazisi Örneği. Bursa Uludag Üniv. Ziraat Fak. Derg. 32 (2), 45–53.

Hady, C., Kusumo, D., 2023. Prediction of a sprint delivery's capabilities in iterative-based software development. Jipi (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika), 8(1), 77-83. doi: 10.29100/jipi.v8i1.3292.

Hao, J., Hu, H., Zhong, R., Xu, J., Xu, J., Huang, J., Lin, Z., 2019. A deep learning approach to conflating heterogeneous geospatial data for corn yield estimation: A case study of the us corn belt at the county level. Glob. Chang. Biol. 26 (3), 1754–1766. https://doi.org/10.1111/gcb.14885.

Hardisky, M.A., Klemas, V., Smart, R.M., 1983. The influences of soil salinity, growth form, and leaf moisture on the spectral reflectance of Spartina alterniflora canopies. Photogramm. Eng. Remote Sens. 49, 77–83.

Helber, P., Bischke, B., Habelitz, P., Sanchez, C., Pathak, D., Miranda, M., ... & Dengel, A., 2023. Crop yield prediction: An operational approach to crop yield modeling on field and subfield level with machine learning models. In IGARSS 2023-2023 IEEE International Geoscience and Remote Sensing Symposium (pp. 2763-2766). IEEE.

Holtorf, L., Titov, I., Daschner, F., Gerken, M., 2023. UAV-based Wireless data collection from underground sensor nodes for precision agriculture. AgriEngineering 5 (1), 338–354.

Huang, Y., Zhang, J., Zhang, J., Ye, L., Zhou, X., Xu, X., Yang, G., 2022. Forecasting Alternaria leaf spot in apple with spatial-temporal meteorological and mobile internet-based disease survey data. Agronomy 12 (3), 679. https://doi.org/10.3390/agronomy12030679.

Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). Remote Sens. Environ. 25, 295–309.

Hunt, M.L., Blackburn, G.A., Carrasco, L., Redhead, J.W., Rowland, C.S., 2019. High resolution wheat yield mapping using Sentinel-2. Remote Sens. Environ. 233, 111410.

Jeong, S., Ko, J., Shin, T., Yeom, J., 2022. Incorporation of machine learning and deep neural network approaches into a remote sensing-integrated crop model for the simulation of rice growth. Sci. Rep. 12 (1) https://doi.org/10.1038/s41598-022-13232-y.

Jeong, J., Resop, J., Mueller, N., Fleisher, D., Yun, K., Butler, E., Kim, S., 2016. Random forests for global and regional crop yield predictions. PLoS One 11 (6), e0156571. https://doi.org/10.1371/journal.pone.0156571.

Ji, Z., Pan, Y., Zhu, X., Wang, J., Li, Q., 2021. Prediction of crop yield using phenological information extracted from remote sensing vegetation index. Sensors 21 (4), 1406. https://doi.org/10.3390/s21041406.

Kalpana, P., Prem, I., Mary, S., Rani, A., 2023. Crop yield prediction using machine learning. Jdaai 2 (1), 16–20. https://doi.org/10.46632/jdaai/2/1/3.

Kayad, A., Sozzi, M., Gatto, S., Marinello, F., Pirotti, F., 2019. Monitoring within-field variability of corn yield using sentinel-2 and machine learning techniques. Remote Sens. (Basel) 11 (23), 2873. https://doi.org/10.3390/rs11232873.

Kennard, R.W., Stone, L.A., 1969. Computer-aided design of experiments. Technometrics 11, 137–148. https://doi.org/10.1080/00401706.1969.10490666.

Khanal, S., Kc, K., Fulton, J., Shearer, S., Ozkan, E., 2020. Remote sensing in agriculture—Accomplishments, limitations, and opportunities. Remote Sens. (Basel) 12 (22), 3783. https://doi.org/10.3390/rs12223783.

Koch, B., Khosla, R., Frasier, W., Westfall, D., Inman, D., 2004. Economic feasibility of variable-rate nitrogen application utilizing site-specific management zones. Agron. J. 96 (6), 1572–1580. https://doi.org/10.2134/agronj2004.1572.

Kodaira, M., Shibusawa, S., 2013. Using a mobile real-time soil visible-near infrared sensor for high resolution soil property mapping. Geoderma, 199, Special Issue (SI), 64–79.

Kuang, B.; Mahmood, H.S.; Quraishi, Z.; Hoogmoed, W.B.; Mouazen, A.M.; van Henten, E.J., 2012. Sensing soil properties in the laboratory, in situ, and on-line: a review. In Donald Sparks, editors: Advances in Agronomy, 114, AGRON, UK: Academic Press, 155-224.

Kuang, B., Mouazen, A.M., 2012. Influence of the number of samples on prediction error of visible and near infrared spectroscopy of selected soil properties at the farm scale. Eur. J. Soil Sci. 63 (3), 421–429.

Lee, W.S., Alchanatis, V., Yang, C., Hirafuji, M., Moshou, D., Li, C., 2010. Sensing technologies for precision specialty crop production. Comput. Electron. Agr. 74, 2–33.

Li, Y., Shi, Z., Wu, C., Li, H., Li, F., 2008. Determination of potential management zones from soil electrical conductivity, yield and crop data. J. Zhejiang Univ. Sci. B 9, 68–76.

Masjkur, M., Tan, K., 2020. Comparing decision tree, random forest and boosting in identifying weather index for rice yield prediction. doi: 10.4108/eai.2-8-2019.2290475.

Mohapatra, S., Chaudhary, N., 2022. Statistical analysis and evaluation of feature selection techniques and implementing machine learning algorithms to predict the crop yield using accuracy metrics. Eng. Sci. https://doi.org/10.30919/es8d787.

Mouazen, A.M., De Baerdemaeker, J., Ramon, H., 2005. Towards development of on-line soil moisture content sensor using a fibre-type NIR spectrophotometer. Soil Till. Res. 80, 171–183.

Mouazen, A.M., Alexandridis, T., Buddenbaum, H., Cohen Y., Moshou, D., Mulla, D., Nawar, S., Sudduth, K.A., 2019. Chapter 2: MONITORING. In: A. Castrignanò et al. (Eds.), Agricultural Internet of Things and Decision Support for Precision Smart Farming. ACADEMIC PRESS, Elsevier, 36-138.

Mouazen, A.M., Maleki, M.R., De Baerdemaeker, J., Ramon, H., 2007. On-line measurement of some selected soil properties using a VIS–NIR sensor. Soil Till. Res. 93, 13–27.

Mouazen, A.M., Maleki, M.R., Cockx, L., Van Meirvenne, M., Van Holm, L.H.J., Merckx, R., De Baerdemaeker, J., Ramon, H., 2009. Optimum three-point linkage set up for improving the quality of soil spectra and the accuracy of soil phosphorus measured using an on-line visible and near infrared sensor. Soil Tillage Res. 103, 144–152. https://doi.org/10.1016/J.STILL.2008.10.006.

Mouazen, A.M., 2006. Soil Survey Device. International publication published under the patent cooperation treaty (PCT). World Intellectual Property Organization, International Bureau. International Publication Number: WO2006/015463; PCT/BE2005/000129; IPC: G01N21/00; G01N21/0.

Mukhopadhyay, S., Maiti, S.K., 2018. Techniques for quantative evaluation of mine site reclamation success: Case study, in: Bio-Geotechnologies for Mine Site Rehabilitation. Elsevier, pp. 415–438.

Mupangwa, W., Chipindu, L., Nyagumbo, I., Mkuhlani, S., Sisito, G., 2020. Evaluating machine learning algorithms for predicting maize yield under conservation agriculture in eastern and southern Africa. Sn Appl. Sci. 2 (5) https://doi.org/10.1007/s42452-020-2711-6.

Nawar, S., Corstanje, R., Halcro, G., Mulla, D., Mouazen, A.M., 2017. Delineation of soil management zones for variable-rate fertilization: A review. Adv. Agron. 143, 175–245.

Pantazi, X.-E., Moshou, D., Alexandridis, T., Whetton, R., Mouazen, A.M., 2016. Wheat yield prediction using machine learning and advanced sensing technologies. Comput. Electron. Agric. 121, 57–65.

Patil, S.M., S, R., 2017. Internet of things based smart agriculture system using predictive analytics. Asian Journal of Pharmaceutical and Clinical Research, 10(13), 148. doi: 10.22159/ajpcr.2017.v10s1.19601.

Peng, B., Guan, K., Pan, M., Li, Y., 2018. Benefits of seasonal climate prediction and satellite data for forecasting U.S. maize yield. Geophys. Res. Lett. 45 (18), 9662–9671. https://doi.org/10.1029/2018gl079291.

Prasad, N., Patel, N., Danodia, A., 2020. Crop yield prediction in cotton for regional level using random forest approach. Spat. Inf. Res. 29 (2), 195–206. https://doi.org/10.1007/s41324-020-00346-6.

Pratt C.J., 1965. Influence of soil pH on the relative availability of macronutrients and micronutrients in the soil. Chemical fertilizers. Scientific American, 212: 62-72.

Radočaj, D., Šiljeg, A., Marinović, R., Jurišić, M., 2023. State of major vegetation indices in precision agriculture studies indexed in web of science: A review. Agriculture 13 (3), 707.

Sabitov, B., Kartanova, A., uulu, T., Seitkazieva, N., Dyikanova, A., & Orozobekova, A., 2023. Modeling and forecasting tasks of agriculture based on machine learning. E3s Web of Conferences, 380, 01026. doi: 10.1051/e3sconf/202338001026.

Sagan, V., Maimaitijiang, M., Bhadra, S., Maimaitiyiming, M., Brown, D.R., Sidike, P., Fritschi, F.B., 2021. Field-scale crop yield prediction using multi-temporal WorldView-3 and PlanetScope satellite data and deep learning. ISPRS J. Photogramm. Remote Sens. 174, 265–281.

Segarra, J., Buchaillot, M.L., Araus, J.L., Kefauver, S.C., 2020. Remote sensing for precision agriculture: Sentinel-2 improved features and applications. Agronomy 10 (5), 641.

Shafi, U., Mumtaz, R., García-Nieto, J., Hassan, S., Zaidi, S., Iqbal, N., 2019. Precision agriculture techniques and practices: From considerations to applications. Sensors 19 (17), 3796. https://doi.org/10.3390/s19173796.

Shahhosseini, M., Hu, G., Huber, I., Archontoulis, S., 2021. Coupling machine learning and crop modeling improves crop yield prediction in the US corn belt. Sci. Rep. 11 (1) https://doi.org/10.1038/s41598-020-80820-1.

Shahi, T.B., Xu, C.Y., Neupane, A., Guo, W., 2022. Machine learning methods for precision agriculture with UAV imagery: A review. Electronic Research Archive 30 (12), 4277–4317.

Stenberg, B., Viscarra Rossel, R., Mouazen, A.M., Wetterlind, J., 2010. Visible and near infrared spectroscopy in soil science. Adv. Agron. 107, 163–215.

Stevens, A., Ramirez Lopez, L., 2014. An introduction to the prospectr package. https://www.researchgate.net/publication/255941339_An_introduction_to_the_prospectr_package.

Tagarakis, A.C., Ketterings, Q.M., 2017. In-Season Estimation of Corn Yield Potential Using Proximal Sensing, Agronomy, Soils & Environmental Quality, J. 109: 1323–1330. doi:10.2134/agronj2016.12.0732.

Tekin, Y., Kuang, B., Mouazen, A.M., 2013. Potential of on-line visible and near infrared spectroscopy for measurement of pH for deriving variable rate lime recommendations. Sensors 13, 10177–10190.

Telkar, S., Solanki, S., Chouhan, S., Kumar, R., Nikas, S.B., 2017. Crop Lodging on Cereals: Causes, Effect and Control Introduction. Biomol. Reports.

Thomas, G. W., 1996. Soil pH and soil acidity. Methods of soil analysis: part 3 chemical methods, 5, 475-490.

Toscano, P., Castrignanò, A., Di Gennaro, S.F., Vonella, A.V., Ventrella, D., Matese, A., 2019. A precision agriculture approach for durum wheat yield assessment using remote sensing data and yield mapping. Agronomy 9 (8), 437.

Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens. Environ. 8, 127–150.

Varella, C., Gleriani, J., Santos, R., 2015. Precision agriculture and remote sensing., 185-203. doi: 10.1016/b978-0-12-802239-9.00009-8.

Viscarra Rossel, R.A., Bouma, J., 2016. Soil sensing: A new paradigm for agriculture. Agric. Syst. 148, 71–74. https://doi.org/10.1016/j.agsy.2016.07.001.

Viscarra-Rossel, R.A., Chen, C., 2011. Digitally mapping the information content of visible–near infrared spectra of surficial Australian soils. Remote Sens. Environ. 115, 1443–1455.

Wei, C., 2017. Predictions of surface solar radiation on tilted solar panels using machine learning models: A case study of Tainan city, Taiwan. Energies 10 (10), 1660. https://doi.org/10.3390/en10101660.

Yang, C., 2020. Remote sensing and precision agriculture technologies for crop disease detection and management with a practical application example. Engineering 6 (5), 528–532. https://doi.org/10.1016/j.eng.2019.10.015.

Yue, J., Lei, T., Li, C., Zhu, J., 2012. The application of unmanned aerial vehicle remote sensing in quickly monitoring crop pests. Intell. Automation Soft Comput. 18 (8), 1043–1052. https://doi.org/10.1080/10798587.2008.10643309.

Zhao, Y., Potgieter, A., Zhang, M., Wu, B., Hammer, G., 2020. Predicting wheat yield at the field scale by combining high-resolution sentinel-2 satellite imagery and crop modelling. Remote Sens. (Basel) 12 (6), 1024. https://doi.org/10.3390/rs12061024.

Ziliani, M.G., Altaf, M.U., Aragon, B., Houborg, R., Franz, T.E., Lu, Y., McCabe, M.F., 2022. Early season prediction of within-field crop yield variability by assimilating CubeSat data into a crop model. Agric. For. Meteorol. 313, 108736.