Check for updates

Determining leaf nutrient concentrations in citrus trees using UAV imagery and machine learning

Lucas Costa[1] · Sudip Kunwar[2] · Yiannis Ampatzidis[1] · Ute Albrecht[2]

© The Author(s) 2021

Abstract

Nutrient assessment of plants, a key aspect of agricultural crop management and varietal development programs, traditionally is time demanding and labor-intensive. This study proposes a novel methodology to determine leaf nutrient concentrations of citrus trees by using unmanned aerial vehicle (UAV) multispectral imagery and artificial intelligence (AI). The study was conducted in four different citrus field trials, located in Highlands County and in Polk County, Florida, USA. In each location, trials contained either 'Hamlin' or 'Valencia' sweet orange scion grafted on more than 30 different rootstocks. Leaves were collected and analyzed in the laboratory to determine macro- and micronutrient concentration using traditional chemical methods. Spectral data from tree canopies were obtained in five different bands (red, green, blue, red edge and near-infrared wavelengths) using a UAV equipped with a multispectral camera. The estimation model was developed using a gradient boosting regression tree and evaluated using several metrics including mean absolute percentage error (MAPE), root mean square error, MAPE-coefficient of variance (CV) ratio and difference plot. This novel model determined macronutrients (nitrogen, phosphorus, potassium, magnesium, calcium and sulfur) with high precision (less than 9% and 17% average error for the 'Hamlin' and 'Valencia' trials, respectively) and micro-nutrients with moderate precision (less than 16% and 30% average error for 'Hamlin' and 'Valencia' trials, respectively). Overall, this UAV- and AI-based methodology was efficient to determine nutrient concentrations and generate nutrient maps in commercial citrus orchards and could be applied to other crop species.

Lucas Costa and Sudip Kunwar have contributed equally.

✉ Yiannis Ampatzidis
  i.ampatzidis@ufl.edu

✉ Ute Albrecht
  ualbrecht@ufl.edu

[1] Agricultural and Biological Engineering Department, Southwest Florida Research and Education Center, University of Florida, IFAS, 2685 SR 29 North, Immokalee, FL 34142, USA

[2] Horticultural Sciences Department, Southwest Florida Research and Education Center, University of Florida, IFAS, 2685 SR 29 North, Immokalee, FL 34142, USA

Introduction

Adoption of best management practices and development of superior food crop cultivars are necessary to cope with pressures imposed by different biotic and abiotic factors (e.g., pests, diseases, drought, water logging, salinity, nutrient deficiencies, extreme temperatures, etc.) and secure food production. General agricultural management practices include nutrient management, pest and disease management, irrigation and drainage (Boman, 2012; Vincent et al., 2021). These practices require regular field monitoring to identify problems and examine crop responses to management. This is often accompanied by specialized laboratory analyses to assess crop physiological responses, which are costly and time-consuming.

Recent advances in plant breeding have accelerated the development of new crop cultivars to cope with a rapidly changing production environment affected by diseases and other plant stresses (Lenaerts et al., 2019). These new cultivars require evaluation on a large-scale in a commercial production setting to assess their horticultural traits, physiological needs and economic potential before being released for widespread commercial adoption. Identification of plant physiological needs through field scouting and hand sampling of tissues for laboratory analysis has been the traditional way to evaluate new cultivars and fine-tune management practices. As these processes are laborious, costly and prone to human error, new phenotyping techniques are needed to advance crop selection and improve crop production (Li et al., 2014). New high-throughput phenotyping techniques are now available that utilize unmanned aerial vehicle (UAVs) (Abdulridha et al., 2020a, b; Costa et al., 2020a) or ground-based remote sensing and artificial intelligence (Burud et al., 2017; Cruz et al., 2019; Reynolds et al., 2019). These techniques have been used to determine various horticultural traits in several crops. For example, the correlation of different models and crop indices based on spectral reflectance data has been studied to assess crop yield in wheat (Hassan et al., 2019; Mirasi et al., 2019), biomass in oat (Coelho et al., 2018), leaf area index in wheat (Xie et al., 2014), plant nutrient content in citrus and grapevine (Osco et al., 2019; Moghimi et al., 2020), detection of pests and diseases in citrus and avocado (Abdulridha et al., 2019a; Partel et al., 2019a, b).

In citrus, remote sensing and AI have also been used to determine crop yield (Vijayakumar et al., 2021; Ye et al., 2007) and canopy volume (Ampatzidis et al., 2019), to count trees (Csillik et al., 2018; Ampatzidis et al., 2020), determine leaf stomatal properties (Costa et al., 2021) and to detect diseases such as huanglongbing (Cerreta et al., 2018; Garza et al., 2020), canker (Abdulridha et al., 2019b) and foot rot (Garza et al., 2020). In this study, citrus was used as a model system for developing novel high-throughput phenotyping techniques utilizing UAV imagery and machine learning to accelerate cultivar selection and improve crop production.

Nutrient management is one of the most important factors in citrus production as it directly influences tree health and productivity (Galvez-Sola et al., 2015; Morgan & Graham, 2019). The plant nutrient status is influenced by the nutrient uptake efficiency of the plant, which is affected by many factors including rootstock cultivar (Uygur & Yetisir, 2009; Toplu et al., 2011; Yilmaz et al., 2018), soil type, season and plant developmental stage (Scagel et al., 2007) as well as soil-borne and other diseases such as huanglongbing (Cao et al., 2015; Morgan & Graham, 2019). To determine the nutrient status of a plant and correct potential deficiencies, regular analysis of leaf nutrient concentrations is essential (Galvez-Sola et al., 2015; Stammer & Mallarino, 2018). Because

the nutrient status of a plant is variable, multi-monthly and multi-annual analyses may be necessary to assess plant responses to environmental factors and management practices accurately.

Plant nutrient analysis requires the chemical analysis of leaf samples in a specialized laboratory, which is expensive and uses toxic chemicals with negative impacts on the environment (Galvez-Sola et al., 2015). Moreover, it is prone to human error resulting from inconsistencies and bias during leaf sampling and during the analysis process, which may compromise interpretation and significance of the data. Faster, cheaper and more environmentally friendly alternatives to conventional nutrient analysis methods are being developed at a rapid pace. Qamar-uz-Zaman & Schumann (2006) calculated NDVI using aerial photography and correlated it with citrus leaf macro and micronutrient concentrations. Machine learning regression models were used with close-range spectroscopy devices to scan leaves and determine concentrations of leaf nitrogen and other leaf nutrients in citrus (Galvez-Sola et al., 2015; Osco et al., 2019, 2020). These methods may help overcome some of the limitations of the traditional method of leaf nutrient analysis or complement it. However, some of these methodologies were developed for specific nutrients and are therefore not applicable to assess all nutrients simultaneously. Moreover, some approaches (e.g., Osco et al., 2020), although precise, use expensive ground-based sensors (e.g., spectroradiometer) for data collection, which requires more resources (e.g., driver, sensor operator) and time to scan large orchards than a UAV-based sensing system. Developing a more efficient methodology using novel machine learning algorithms and UAV-based sensing may help improve the overall consistency and data collection speed.

The objectives of this study were to: (i) develop a novel high-throughput method to determine leaf nutrient concentrations in citrus (as a case study), (ii) utilize this method to identify nutrient deficient zones and (iii) create fertility maps compatible with variable rate technologies (e.g., variable rate fertilizers) for zone-based management. The proposed technique and model can be applied to different crops and production systems.

Materials and methods

Study site

Four field trials with 5-year-old grafted sweet orange (*Citrus sinensis*) trees growing in a commercial citrus production system in Florida, USA (Lykes Bros. Inc.) were used for nutrient analysis and model development. Trees were composed of 2 different scion cultivars and more than 30 different rootstock cultivars. The scion cultivars were 'Hamlin' orange, an early maturing cultivar, and 'Valencia' orange, a late maturing cultivar. The rootstocks included more than 30 commercial and experimental cultivars with different taxonomic backgrounds (Kunwar et al., 2021; Ampatzidis et al., 2019).

Trials A (10.76 ha) and B (11.25 ha) were located in Southeast Florida near Fort Basinger, Highlands County (27° 22′ 16.0″ N 81° 08′ 08.0″ W), and trials C (8.78 ha) and D (two areas with 6.10 ha and 3.43 ha) were located in Central Florida near Lake Wales, Polk County (27° 56′ 07.4″ N 81° 30′ 00.1″ W) (Fig. 1). The soil type in trials A and B is a poorly drained sandy Entisol with Spodosol-like properties (Mylavarapu et al., 2016), and trees were planted in double-rows on raised beds separated by furrows at a spacing of 2.4 m along the rows and 7.6 m between the rows. The soil in trials C and D is a well-drained sandy Entisol, and trees were planted in single rows without beds at a spacing of 2.4 m
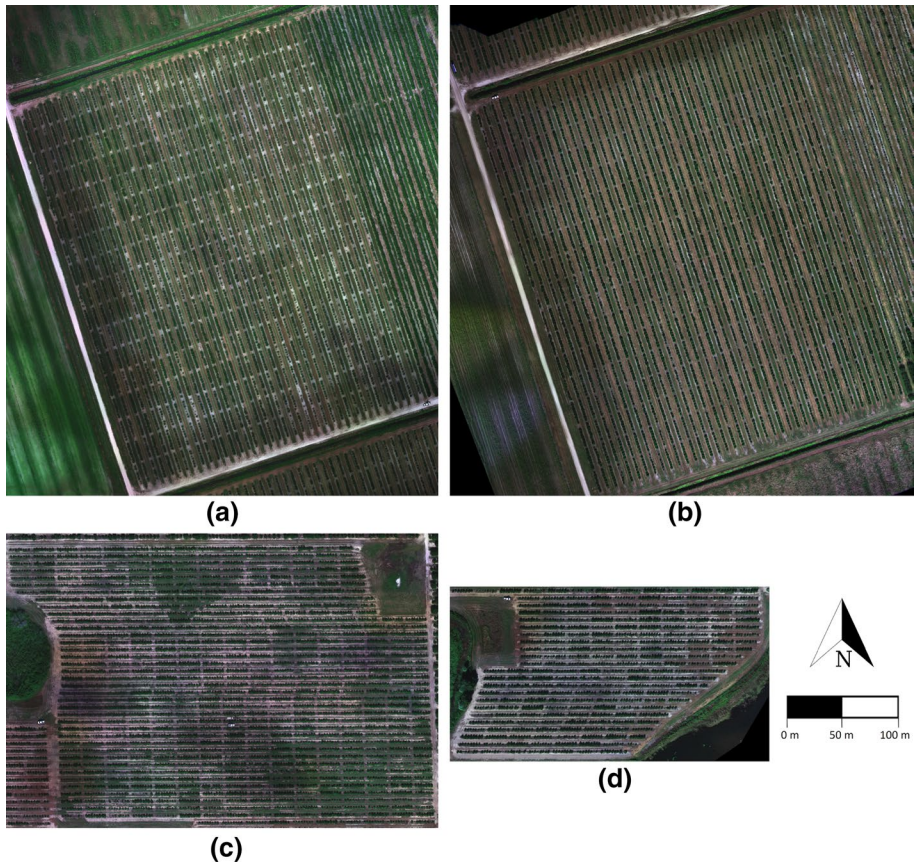
Fig. 1 Aerial views of **a** trial A, **b** trial B, **c** trial C, and **d** trial D-east

along the rows and 6.7 m between the rows. Trees in all trials were produced in a commercial citrus nursery (Lykes Bros. Inc., Basinger, FL, USA) and were planted in 2015. Trees in trials A and C were composed of 'Hamlin' scion on 32 and 35 different rootstock cultivars, and trees in trials B and D were composed of 'Valencia' scion on 32 and 35 rootstock cultivars (Table 1). The experimental design was randomized with rootstocks replicated as linear plots of eight trees. Irrigation was by under-tree micro-sprinklers. Nutrient, disease and weed management was per the grower's standards, as outlined in Kunwar et al., (2021) and was similar for all the trials.

Study workflow

The framework of this study was divided into two main phases (Fig. 2). In the data acquisition phase (Fig. 2, green), the spectral measurements of the canopy reflectance were acquired using a UAV-based multispectral camera (red, green, blue, red edge and near-infrared wavelengths), and leaf samples were collected and analyzed in a laboratory as described in the next paragraph to generate the dataset. The second phase (Fig. 2, blue) consisted of: (i) pre-analysis to evaluate the dataset for each nutrient; (ii) model

Table 1 Details of study sites used for model development

	Trial A	Trial B	Trial C	Trial D
Location	Highlands County, FL, USA		Polk County, FL, USA	
Scion cultivars	'Hamlin'	'Valencia'	'Hamlin'	'Valencia'
No of rootstocks	32	32	35	35
No of samples[1]	192	192	210	210

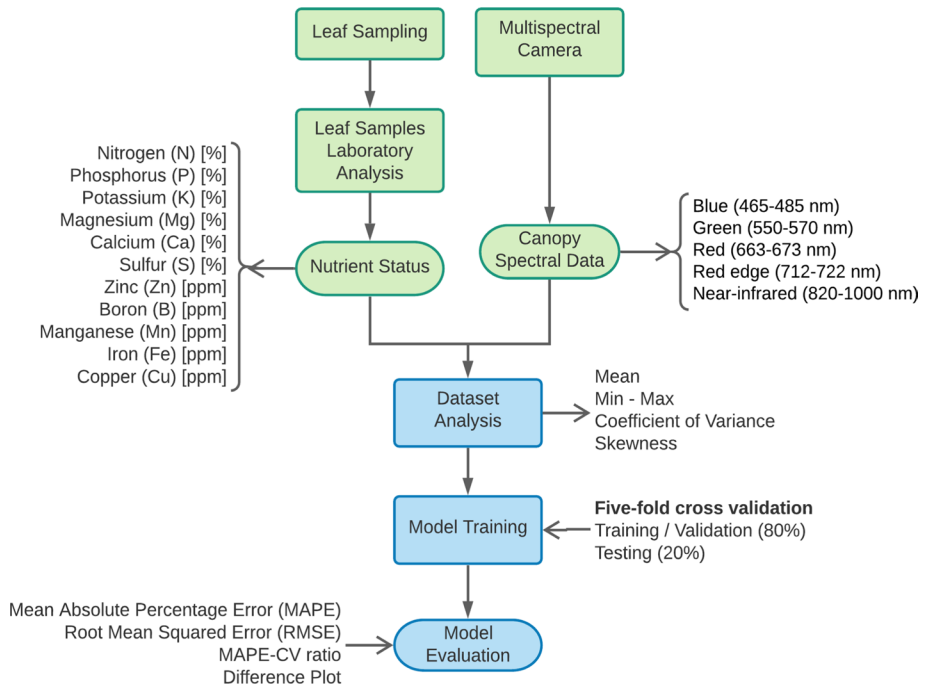[1]A sample was composed of leaves from two trees in each linear plot

Fig. 2 Workflow of the development of the estimation model for nutrient analysis that includes dataset collection and analysis, and model development and evaluation (Color figure online)

development using the multispectral bands as inputs and 5-fold cross-validation (80% of samples for training and validation of the model, 20% for testing) to ensure the repeatability of the methodology used; and (iii) with the model generated, the evaluation metrics of mean absolute percentage error (MAPE), root mean square error (RMSE), MAPE-CV ratio and difference plot were applied to evaluate the results.

Leaf nutrient analysis

Six replications per scion/rootstock combination in each trial were used for leaf nutrient analysis using traditional procedures. A random sample of 16 mature leaves (4 from each cardinal direction) from the most recent spring flush was collected from the 3rd and the 6th trees from each replicated plot in September 2019. Because there were more than six replications of each scion/rootstock combination in each trial, not all areas at each trial site were equally sampled. Leaf samples from the two trees were pooled to generate one leaf sample for analysis. In total, 804 leaf samples were collected, 192 from each trial at the Polk County location and 210 from each trial at the Highlands County location (402 leaf samples each from 'Hamlin' and 'Valencia' scion). Leaf nutrient analysis was conducted by a commercial analysis service (Waters Agricultural Laboratories, Georgia, USA). Total leaf nitrogen (N) was determined using the Dumas combustion method slightly modified by Sweeny (1989), which measures all forms of nitrogen (ammonium, nitrate, heterocyclic and protein). For the other nutrients—phosphorus (P), sulfur (S), potassium (K), zinc (Zn), magnesium (Mg), calcium (Ca), iron (Fe), manganese (Mn), copper (Cu) and boron (B)

- leaf samples were first digested using nitric acid and hydrogen peroxide solution followed by inductively coupled argon plasma (ICAP) analysis (Havlin & Soltanpour, 1980).

Spectral data acquisition

The canopy reflectance of trees was measured with a 5-band multispectral camera (Micasense Altum, Micasense, USA) [blue (465–485 nm), green (550–570 nm), red (663–673 nm), red edge (712–722 nm) and near-infrared (820–1000 nm)] mounted on a quadcopter UAV (Matrice 210, DJI, Shenzehen, China). For UAV flight planning and mission control, the Pix4DCapture (Pix4D S.A., Prilly, Switzerland) software app was used on an iPad (Apple, Cupertino, CA, USA) connected to the remote controller of the UAV. All flights were conducted around solar noontime with a clear sky to minimize atmospheric interference. The UAV flew at 122 m height. The generated ortho-mosaic map had a resolution of 50 mm/pixel. Figure 3 presents the workflow of data collection to generate the study's dataset. The dataset was divided into (1) 'Hamlin' scion (containing all rootstock cultivars), and (2) 'Valencia' scion (containing all rootstock cultivars).

5-Fold cross-validation

A 5-fold cross-validation scheme was used to validate the model's ability to estimate the objectives. It is a popular strategy since it is straightforward to understand and produces a less biased or optimistic estimate of model competence than other approaches, such as
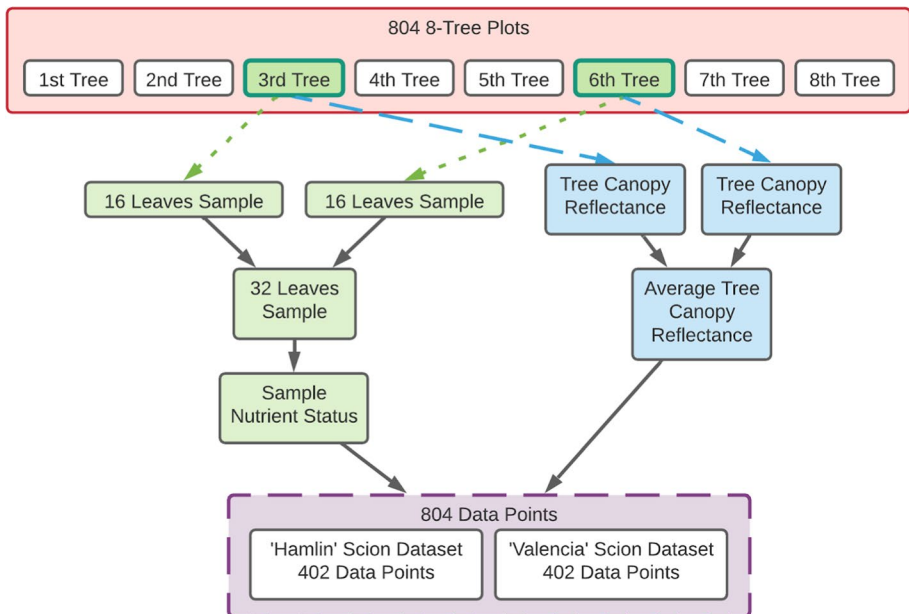

Fig. 3 Data collection and dataset generation workflow

Fig. 4 General procedure for a 5-fold cross-validation

Randomly shuffle the dataset
Split the data into 5 groups, each containing 20% of the dataset
For each unique group, do:
    Use chosen group as a test dataset (20%)
    Use remaining groups as a training dataset (80%)
    Fit a model to the training set and test it against the test set
    Keep the evaluation score and discard the model
Use all five scores to summarize the model estimation skill

Table 2 Guidelines for interpretation of leaf analysis based on 4-6 month-old spring flush from non-fruiting twigs (based on Kadyampakeni & Morgan 2020)

Nutrients	Deficient	Low	Optimum	High	Excess
N [%]	< 2.2	2.2–2.4	2.5–2.7	2.8–3.0	> 3.0
P [%]	< 0.09	0.09–0.11	0.12–0.16	0.17–0.30	> 0.30
K [%]	< 0.7	0.7–1.1	1.2–1.7	1.8–2.4	> 2.4
Mg [%]	< 0.20	0.20–0.29	0.30–0.49	0.50–0.70	> 0.70
S [%]	< 0.14	0.14–0.20	0.20–0.40	0.40–0.60	> 0.60
Ca [%]	< 1.5	1.5–2.9	3.0–4.9	5.0–7.0	> 7.0
B [ppm]	< 20	20–35	36–100	101–200	> 200
Zn [ppm]	< 18	18–24	25–100	101–300	> 300
Mn [ppm]	< 18	18–24	25–100	101–300	> 300
Fe [ppm]	< 35	35–59	60–120	121–200	> 200
Cu [ppm]	< 3	3–4	5–16	17–20	> 20

a simple train/test split. The general procedure of the 5-fold cross-validation used is presented in Fig. 4.

Data analysis

The dataset was analyzed using standard statistical procedures to obtain the mean, standard deviation, skewness and coefficient of variation (CV) values. The skewness evaluates the distribution of the data, where zero corresponds to a symmetric distribution. Negative and positive skewness presents a distribution of values clustered on smaller values (negative) or higher values (positive). The coefficient of variation is presented in percent and shows the variability of the dataset concerning its mean value.

The recommended ranges for each nutrient based on the University of Florida's (UF) Institute of Food and Agricultural Sciences (IFAS) guidelines for citrus nutrient management (Kadyampakeni & Morgan, 2020) are presented in Table 2. These values were used as guidelines for data interpretation in this study.

Regression model

The collected spectral reflectance (five bands) from the multispectral camera was used to generate a regression model capable of using these five inputs to estimate the nutrient concentrations of the crop. For this study, multiple regression algorithms such as ElasticNet, Lasso regression, Linear SVM, PLSR, Random Forest and Ridge regression were

evaluated on their ability to generate a model in this dataset without overfitting or with acceptable error. The gradient boosting regression model proved to have the best results and therefore is used in this study.

Gradient boosting regression tree

The gradient boosting regression tree is an ensemble model. Ensemble-based models consist of multiple base models (in this case, regression trees) where each base model provides a solution to the problem, whose outputs are combined in some way (typically by weighted or unweighted voting or averaging) to produce the final ensemble model output (Zhang & Haghani, 2015). The success of the ensemble comes from the reduced total error by correcting the mistakes of individual models. Combining individual base models with different errors can reduce the final error of the ensemble model. Trees are one type of base model that is commonly used for ensembles. They are sensitive to small variability in the training data. This unique property makes them good candidates for ensembles (Zhang & Haghani, 2015).

A single tree regression model is partitioned in nodes occupied each by a simple constant model. In these nodes, a decision is made based on the value of the variable being analyzed. For each subsequent node, a new decision is based on an input variable until it reaches the end of the tree and gives the output. Figure 5a presents a schematic plot for a simple regression tree with decision nodes, where X1 and X2 are the variables being analyzed, b1, b2, b3 and b4 are values of the simple constant function, and Ya, Yb, Yc, Yd, Ye are the possible outputs of the model. Figure 5b presents a schematic example of an ensemble model composed of multiple regression trees where B1, B2, B3, B4 and B5 represent each a base model (single regression tree) and the output Y is the weighted (w1, w2, w3, w4, w5) vote of them.

The gradient boosting methodology generates base models sequentially. Estimation accuracy is improved by emphasizing on the training cases that are more difficult to estimate. In the boosting process, examples that are harder to determine appear more often in the training phase. Each new base model is designed to correct the mistakes made by its previous base models.
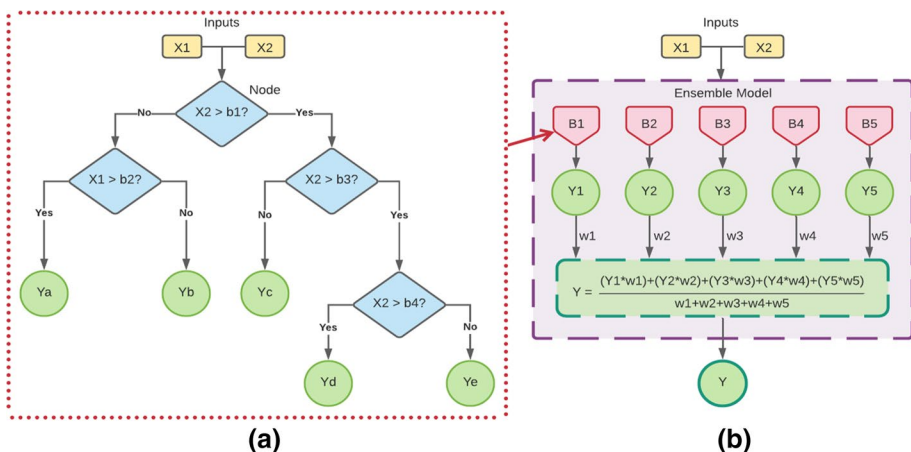

Fig. 5 **a** Single regression tree example; and **b** ensemble model example

The purpose of resampling of the training data is to provide the most useful information for each consecutive model. The adjusted distribution for each step of training is based on the error produced by the previous models. Figure 6 presents the steps of the gradient boosting algorithm.

Evaluation metrics

Mean absolute percentage error (MAPE), root mean squared error (RMSE), and MAPE-CV ratio

The mean absolute percentage error (MAPE) is the computed average of absolute percentage errors by which outputs of a model differ from actual values of what is being estimated. The MAPE is given by the sum of the absolute difference between both the ground truth (Gt) and the model output (Y) divided by the ground truth, which is then divided by the number of individuals (n). The formula for the MAPE is presented in Eq. 1.

$$\text{MAPE} = 1/n \sum \left| \frac{Gt - Y}{Gt} \right| \tag{1}$$

The root-mean-squared error (RMSE) is the root of the squared difference between the ground truth (Gt) and output value (Y), divided by the number of individuals (n) (Eq. 2).

$$\text{RMSE} = \sqrt{\left(\sum (Gt - Y)^2 / n \right)} \tag{2}$$

When analyzing the performance of different algorithms and models, the quality of the dataset must also be considered. Datasets with high CV values present a greater challenge for regression algorithms, as the possible outputs are spread into a larger range. In contrast, datasets with low CV values present a lesser challenge for regression algorithms. In these cases, a good analysis is the MAPE-CV ratio given in Eq. 3, which is a ratio of error by challenge level, an indicator of performance where smaller values correspond to better performance.

$$MAPE - CVRatio = MAPE/CV \tag{3}$$

Measure of agreement The agreement between measurements refers to the degree of concordance between two sets of measurements of the same individual by two different methodologies. The Pearson correlation coefficient is often inappropriately used to evaluate for agreement since it is an incorrect measure of reproducibility or repeatability (Watson & Petrie, 2010). This invalidates the approach for comparison between other works with different datasets. Statistical methods to test agreement are used to decide

Fig. 6 Work flow of the gradient boosting algorithm

Given a dataset *D*. Total number of base models as *M*:
The initial training resample is *D1 = D*
For *m*=1 to *M* do:
    Train a base model *Bm* using the sample *Dm*
    Compute the error of the model
    Output the developed model *Bm*
    Adjust the new resample *Dm+1* to evidence the error of model *Bm*
End
Output the ensemble trees

whether one technique for measuring a variable can substitute another (Ranganathan et al., 2017).

The difference plot is a display of the pattern and agreement of one variable being measured by two different methodologies (Watson & Petrie, 2010). The diagram plots the difference between a measurement pair on the vertical axis and the mean of the pair on the horizontal axis. To determine the repeatability of the proposed approach, the method assumes a normal distribution of differences, where 95% of them are expected to lie between $d \pm 1.96\,s$, where d is the mean of observed differences and s is the standard deviation, generating an error in a confidence range (95%).

Interpretation

Given all evaluation metrics, an important step is understanding how to interpret these results. MAPE presents the absolute regression error in percentage, being suited to compare the results when the datasets have different units or magnitudes. RMSE presents the same absolute regression error in the original units used in the dataset, being suited for comparing different regressions in the same dataset. The MAPE-CV ratio is a score of regression performance, being the ratio between efficacy and challenge.

Lastly, the difference plot presents the agreement between two sets of variables. This also generates an error range in a confidence level (95% in this study). Essentially, this is the maximum error expected for each regression with a 95% confidence. The key interpretation here is the difference between this range and the MAPE, where the first is the worst-case scenario for a single estimation, while the latter is the average error expected for a large number of estimations. These metrics are suited for specifically scaled comparisons; the error range should be used as the error of a small-scale estimation (for example, when comparing the estimation of nutrient content between two citrus trees), and the MAPE should be used on a large-scale estimation (for example, when comparing the average nutrient content of two groves, each containing more than 5,000 citrus trees).

Results

Dataset analysis

'Hamlin' dataset

Table 3 presents the leaf nutrient data (lab chemical analysis) for the 'Hamlin' trials. In this case, N, P and S returned low variability, but the data are still viable as these are the normal ranges for each of the nutrients for these trees. As this low variability could force the model to overfit, to evaluate this possibility, the MAPE of the model created was compared to the CV of the data itself. This was used to verify that the model was aiming to determine the correct values.

The larger values of CV, namely for Fe and Cu, present a dataset with larger variability, which could mean that these nutrients present a more challenging problem for the model to determine the correct value. Based on the nutrient recommendations for citrus

Table 3 Mean, minimum and maximum, CV and skewness of the 'Hamlin' dataset

Nutrient	Mean	Min–Max	CV	Skewness
Nitrogen (N) [%]	2.88	1.53–3.47	9.34%	− 2.14
Phosphorus (P) [%]	0.18	0.13–0.38	9.28%	2.47
Potassium (K) [%]	1.77	1.05–3.43	18.03%	− 0.08
Magnesium (Mg) [%]	0.34	0.25–0.66	13.76%	1.59
Calcium (Ca) [%]	2.82	1.93–4.82	13.25%	0.44
Sulfur (S) [%]	0.35	0.27–0.64	10.87%	1.20
Boron (B) [ppm]	100.87	62.00–173.72	13.95%	0.41
Zinc (Zn) [ppm]	27.99	12.00–69.00	32.08%	0.61
Manganese (Mn) [ppm]	50.98	22.58–99.00	23.16%	0.67
Iron (Fe) [ppm]	101.15	43.00–530.00	56.01%	2.78
Copper (Cu) [ppm]	186.76	22.46–456.80	51.94%	0.30

(Kadyampakeni & Morgan, 2020), the average leaf nutrient concentrations for the 'Hamlin' trees were in the optimal to high range, except for Ca, which was low, and Cu, which was in excess.

'Valencia' dataset

Table 4 presents the leaf nutrient data for the 'Valencia' trials. Although the mean, minimum and maximum values are similar to the 'Hamlin' dataset, the data were more spread (larger CV). This can produce a larger error in the model, which had to be investigated further in the model evaluation. Based on Kadyampakeni and Morgan (2020), the average leaf nutrient concentrations for the 'Valencia' trees were in the optimal to high range, except for Cu, which was in excess.

Model: 'Hamlin' dataset (UAV imagery)

The models were generated using the open-source library scikit-learn for Python. The five-fold cross validation ensures that all data points were used for training and testing, where the average error for all five folds is considered the overall error of the model and

Table 4 Mean, minimum and maximum, CV and skewness of the 'Valencia' dataset

Nutrient	Mean	Min–Max	CV	Skewness
Nitrogen (N)	2.78	1.53–3.47	11.70%	− 1.66
Phosphorus (P)	0.18	0.14–0.24	8.86%	0.29
Potassium (K)	1.68	1.05–2.54	23.77%	0.31
Magnesium (Mg)	0.37	0.25–0.59	13.54%	1.09
Calcium (Ca)	3.04	2.03–4.15	10.79%	− 0.11
Sulfur (S)	0.37	0.30–0.49	8.65%	0.43
Boron (B)	97.02	62.00–150.00	16.39%	0.72
Zinc (Zn)	27.01	12.00–69.00	39.92%	0.84
Manganese (Mn)	50.86	26.00–99.00	24.96%	0.86
Iron (Fe)	109.33	43.00–530.00	69.06%	2.06
Copper (Cu)	128.37	32.00–455.00	59.79%	0.89

Table 5 Mean absolute percentage error (MAPE), root mean squared error (RMSE) and MAPE-CV ratio for the estimation models (five-fold validation)

| Nutrient	MAPE Group 1	MAPE Group 2	MAPE Group 3	MAPE Group 4	MAPE Group 5	Average MAPE	RMSE	MAPE-CV Ratio
Nitrogen (N)	4.67%	3.54%	2.96%	3.78%	4.14%	3.82%	0.1454	0.41
Phosphorus (P)	8.46%	5.45%	5.72%	5.56%	5.23%	6.09%	0.0224	0.66
Potassium (K)	8.13%	8.00%	8.06%	8.27%	7.78%	8.05%	0.2008	0.45
Magnesium (Mg)	6.95%	7.61%	8.10%	6.52%	8.79%	7.59%	0.0350	0.55
Calcium (Ca)	7.81%	8.52%	8.44%	8.23%	9.50%	8.50%	0.2930	0.64
Sulfur (S)	4.42%	4.87%	4.71%	5.56%	4.95%	4.90%	0.0261	0.45
Boron (B)	7.47%	7.68%	8.47%	8.97%	6.96%	7.91%	11.0564	0.57
Zinc (Zn)	13.82%	11.97%	13.88%	10.76%	12.37%	12.56%	5.0776	0.39
Manganese (Mn)	17.34%	15.35%	15.35%	14.34%	15.59%	15.59%	9.8151	0.67
Iron (Fe)	13.72%	15.57%	14.06%	14.84%	13.00%	14.24%	17.9639	0.25
Copper (Cu)	27.77%	22.26%	20.13%	20.41%	21.33%	22.38%	62.0127	0.43
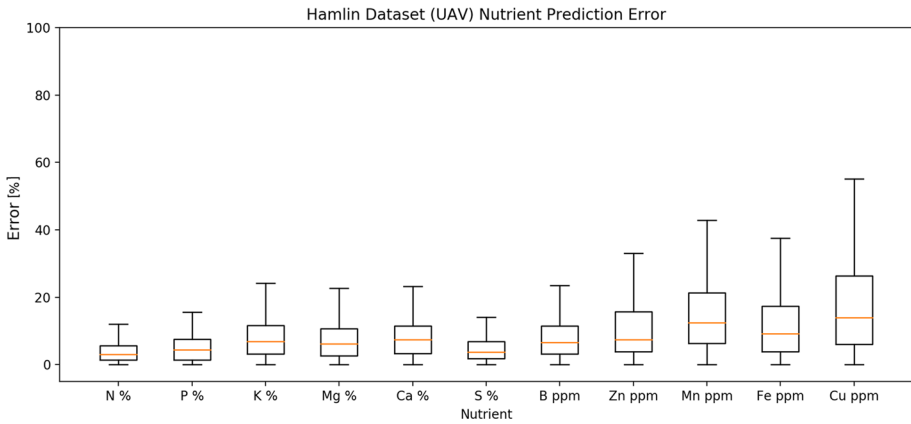
Fig. 7 Whisker plot of the distribution of errors for each nutrient in the 'Hamlin' dataset

Fig. 8 Difference plot for nitrogen measurements in the 'Hamlin' (UAV) dataset
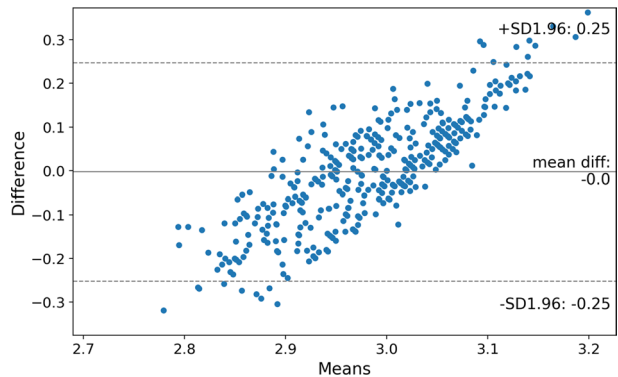
algorithm. Table 5 shows the MAPE values for all nutrients for each validation (five-fold), the average MAPE, the average RMSE and MAPE-CV ratio.

Nutrients like N, P and S, which had a lower variability in the dataset (lower CV in Table 3), also had lower MAPE values (Table 5), while nutrients with higher variability in the dataset (higher CV in Table 3) had higher MAPE values (Table 5). A lower variability in the dataset will usually generate models with higher precision, as these values are spread in smaller ranges and therefore generate smaller errors. In such cases with different variabilities, the MAPE-CV ratio is useful to compare results, with a smaller ratio indicating a better performance.

Overall, the MAPE shows a good regression for most of the nutrients, with errors under 10%. For Zn, Mn, Fe and Cu, which returned larger errors (Table 5), the model was less accurate, but satisfactory, considering that data acquisition was done by UAV (and not by time-consuming and laborious laboratory analysis). The error distribution is shown in Fig. 7. For N and S, the spread of errors is under 10% of the absolute error. For P, K, Mg, Ca and B, the model resulted in errors of less than 20%. These are encouraging results for a procedure that uses UAV imagery.

The difference plots were used to compare the standard method of measuring nutrient concentrations (lab chemical analysis) with the new proposed UAV- and AI-based method. Figure 8 shows the difference plot for N as an example. It shows that the operating range for N determination by spectral canopy reflectance analysis works on a ± 0.25 range of error for the percentage of the nutrient in the plant. This is equivalent to a ± 8.39% range of error in operation.

The difference plot was applied to all nutrients. Table 6 shows the upper and bottom range of operation in their nutrient concentrations and equivalent values in percentage. The table shows that the maximum error for almost half of the nutrients was less than 20%.

Model: 'Valencia' dataset (UAV imagery)

This model was generated following the same procedures described for the 'Hamlin' model. Table 7 shows the MAPE for all nutrients for all five-fold validations, the average MAPE, RMSE and MAPE-CV ratio. It shows a larger error for the estimation as expected from a dataset with more variability, but the average MAPE for most nutrients can be considered precise enough for a UAV-based measurement. The same pattern was found for the distribution of errors, where the larger CV in the dataset analysis shows a higher average MAPE.

The error distributions are presented in Fig. 9. Although errors are larger compared with the 'Hamlin' model, the errors for N, P and K are still less than 20%. The error distribution for Fe and Cu are shown as infinite, possibly due to outliers in the dataset. The MAPE-CV ratio for this dataset is larger, showing a poorer performance.

The difference plot was again applied to all nutrient data. Table 8 presents the upper and bottom range of operation in their nutrient concentrations and equivalent values in percentage. This model presents a larger error range of operation, expected from a wider spread of errors. When comparing to the other UAV model ('Hamlin' dataset), these ranges show that the presence of outliers can be highly influential. This can be minimized by collecting a larger dataset to better represent these outliers.

	Nutrient	Upper range	Bottom range	Range in percentage	
Table 6 Upper and bottom range for the difference plot for each nutrient in the 'Hamlin' dataset	Nitrogen (N)	0.25	− 0.25	8.39%	− 8.39%
	Phosphorus (P)	0.03	− 0.03	17.06%	− 17.06%
	Potassium (K)	0.34	− 0.33	18.31%	− 17.77%
	Magnesium (Mg)	0.06	− 0.07	18.60%	− 21.70%
	Calcium (Ca)	0.51	− 0.50	19.61%	− 19.22%
	Sulfur (S)	0.05	− 0.05	15.41%	− 15.41%
	Boron (B)	18.08	− 18.35	17.27%	− 17.53%
	Zinc (Zn)	8.68	− 8.50	29.97%	− 29.34%
	Manganese (Mn)	16.38	− 16.48	32.06%	− 32.25%
	Iron (Fe)	29.03	− 28.99	31.22%	− 31.18%
	Copper (Cu)	106.37	− 105.42	43.41%	− 43.03%

Table 7 Mean absolute percentage error (MAPE), root mean squared error (RMSE) and MAPE-C ratio for the estimation models (five-fold validation)

| Nutrient	MAPE Group 1	MAPE Group 2	MAPE Group 3	MAPE Group 4	MAPE Group 5	Average MAPE	RMSE	MAPE-CV Ratio
Nitrogen (N)	11.27%	9.37%	10.78%	8.13%	9.68%	9.85%	0.3558	0.84
Phosphorus (P)	8.44%	9.03%	7.73%	6.97%	7.27%	7.89%	0.0187	0.89
Potassium (K)	15.61%	17.00%	14.26%	17.55%	16.75%	16.23%	0.3463	0.68
Magnesium (Mg)	11.85%	10.08%	12.06%	10.73%	13.15%	11.57%	0.0575	0.85
Calcium (Ca)	9.49%	9.04%	10.01%	11.15%	10.28%	9.99%	0.3901	0.93
Sulfur (S)	7.36%	7.98%	7.96%	6.78%	7.48%	7.51%	0.036	0.87
Boron (B)	12.61%	12.14%	15.06%	13.49%	13.84%	13.43%	16.6577	0.82
Zinc (Zn)	29.67%	34.67%	29.25%	27.93%	33.96%	31.09%	10.2218	0.78
Manganese (Mn)	25.20%	22.06%	19.22%	20.41%	21.69%	21.72%	13.9967	0.87
Iron (Fe)	61.01%	50.51%	46.99%	67.14%	45.17%	54.17%	86.6101	0.78
Copper (Cu)	54.98%	62.66%	55.92%	50.96%	61.02%	57.11%	78.9178	0.96
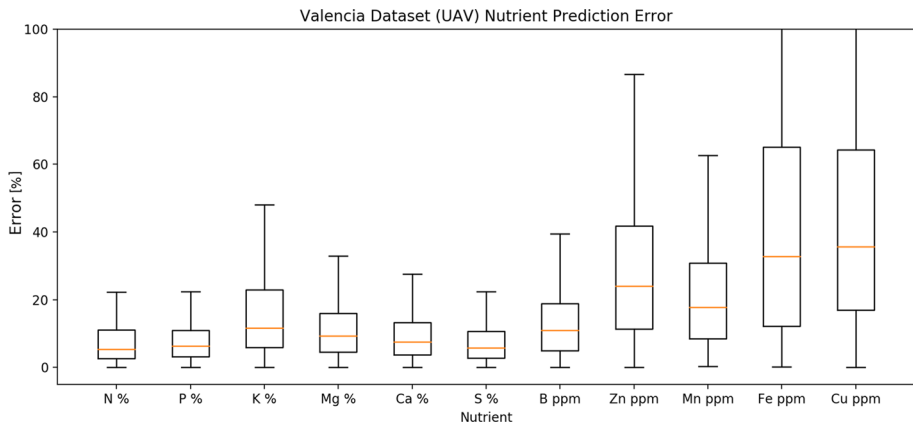
Fig. 9 Whisker plot of the distribution of errors for each nutrient in the 'Valencia' (UAV) dataset

Table 8 Upper and bottom range for the difference plot for each nutrient in the 'Valencia' (UAV) dataset

Nutrient	Upper range	Bottom range	Range in percentage	
Nitrogen (N)	0.71	− 0.68	25.51%	− 23.60%
Phosphorus (P)	0.04	− 0.04	21.72%	− 22.21%
Potassium (K)	0.68	− 0.68	40.38%	− 38.40%
Magnesium (Mg)	0.12	− 0.11	32.52%	− 31.81%
Calcium (Ca)	0.75	− 0.79	24.65%	− 28.00%
Sulfur (S)	0.07	− 0.07	18.69%	− 20.03%
Boron (B)	32.40	− 32.43	33.39%	− 32.15%
Zinc (Zn)	20.03	− 20.03	74.17%	− 71.57%
Manganese (Mn)	27.48	− 26.99	54.03%	− 52.94%
Iron (Fe)	164.1	− 175.46	150.10%	− 173.47%
Copper (Cu)	152.08	− 158.36	118.47%	− 123.36%

Discussion

To evaluate the performance of the proposed model, leaf nutrient data were used from two different citrus scion cultivars grafted on more than 30 different rootstock cultivars and growing in two field locations with different soil and environmental conditions. The best overall fit was found for N, for which the model achieved an 8.39% range of operational error (difference plot range) for the 'Hamlin' dataset, and a 25.51% error for the 'Valencia' dataset. For the other nutrients (P, K, Mg, Ca, S and B), the error ranges were less than 20% for the 'Hamlin' dataset, and up to 40% for the 'Valencia' dataset. The model was not able to achieve acceptable errors for the micro-nutrients Zn, Mn, Fe and Cu. However, on a large-scale analysis, the MAPE better presents the model's estimation accuracy (see "Evaluation metrics" section); the model determined N with a 3.82% MAPE and 9.85% MAPE for the 'Hamlin' and 'Valencia' datasets, respectively.

Assessing leaf nutrient concentrations of citrus trees is important for determining the nutrient status of the plants, which can fluctuate depending on the developmental

stage of the plant, environmental conditions and pest/disease pressure. Regular nutrient analysis is important in crop production to identify and correct potential nutrient deficiencies and prevent production losses (Shaw et al., 2002; Kadyampakeni & Morgan, 2020; Bahtiar et al., 2020). Rather than collecting leaves manually and subjecting them to costly chemical analysis procedures, combining UAV imaging (spectral leaf reflectance measurements) and AI is less laborious, safer for the environment and more cost-efficient. This approach not only allows determination of the crop nutrient status but also development of fertility maps for variable rate fertilizer applications for further cost savings and environmental benefits. The aerial maps produced may also identify other problems. These include differences in the topography or micro-environment of the production site as well as disease zones that may affect nutrient availability, uptake or retention, or technical problems that may have arisen from malfunctions of machinery and other equipment associated with crop management.

To demonstrate the development of fertility maps, and the correlation of the canopy reflectance data and the nutrient concentrations obtained by chemical analysis, nitrogen maps were generated for the 'Hamlin' trees in trial B (Fig. 10). These maps and zones were based on the UF/IFAS guidelines for nutrient ranges of citrus trees (Table 2). One map (Fig. 10a) was generated using the proposed methodology (AI-based estimation model) based on canopy spectral reflectance values for all 4925 trees that are part of the trial in that location, and the second map (Fig. 10b) was generated using the ground truth data (laboratory chemical analysis) from a subset of 192 tree samples (red dots represent individual tree samples in the study site).

The major advantage of the proposed methodology is that large populations of plants can be assessed quickly and at low cost, while also reducing inaccuracies resulting from sampling a limited subset of plants. This is demonstrated in Fig. 10b. It can be seen that there is a larger discrepancy between maps in the areas where fewer trees were sampled (red dots). For example, only three leaf samples were obtained from the northeast area of this trial site, which contained 202 trees. Similarly, only three samples were collected
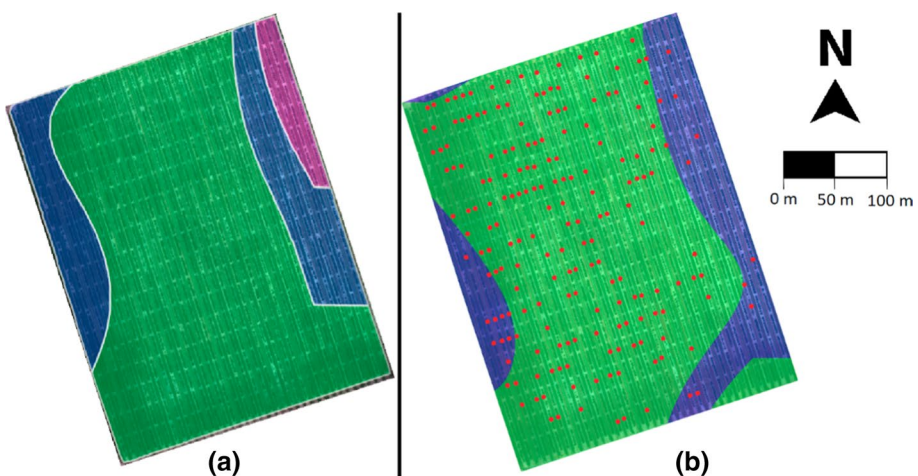

Fig. 10 Nitrogen maps based on UF/IFAS guidelines (Table 2), with optimum (green), high (blue), and excess (pink) zones: **a** map developed by the proposed methodology; and **b** map developed from ground truth samples (chemical analysis; red dots represent sample points) (Color figure online)

from the southeast area, which contained 307 trees. This difference in sampling frequency is likely the major reason for the discrepancy of the two maps in these areas of the trial.

In addition to aiding growers in nutrient management and disease identification, the proposed technology could be used in breeding programs to study cultivar performance and assess nutritional requirements and cultivar adaptation to different environmental conditions and diseases. A recent study by Osco et al., (2020) presented a framework based on machine learning algorithms to assess leaf nutrient concentration based on ground-based hyperspectral imaging. The study used different algorithms such as k-nearest neighbor (KNN), lasso regression, ridge regression, support vector machine (SVM), artificial neural network (ANN), decision tree (DT) and random forest (RF). Leaf nutrient determination with the proposed UAV-based methodology presented here yielded results comparable to those obtained by all algorithms used in Osco et al., (2020). For example, in that study, the best regression achieved for N was with a MAPE of 2.39%. In contrast, the methodology proposed in this study achieved a MAPE of 3.82%. Although most nutrient values in the dataset were similar to values measured by Osco et al., (2020), Cu concentrations differed considerably (157.57 ppm in current dataset, and 72.20 ppm in Osco et al., 2020) as did CVs (44.01% in current dataset, and 36.14% in Osco et al., 2020). The reason for the higher Cu concentrations and variance measured in this study is the frequent foliar sprays of this element to combat the bacterial disease citrus canker (Behlau et al., 2010). Achieving 80–90% accuracy of measurement for most of the other nutrients demonstrates the superiority of the proposed UAV- and AI-based methodology compared to the methodology used by Osco et al. (2020). The UAV based approach is ideal for evaluation of large-scale commercial fields, where larger areas can be covered by UAVs in considerably less time than by ground-based sensing systems.

In this study, the difference plots were incorporated into the analysis to evaluate accuracy of the model. Difference plots better represent accuracy of an application by considering most of the errors and eliminating some outliers to generate a range of operational errors. This is useful when comparing different datasets, as demonstrated in this study.

Conclusions

Herein, a methodology for generating and applying regression models for determining nutrient concentrations of citrus trees was developed. Different datasets including two different scion cultivars ('Valencia' and 'Hamlin') in combination with more than 30 different rootstock cultivars and two different locations were used. The proposed method incorporated a gradient boosting regression tree to determine different macro- and micro-nutrients and was evaluated with a cross validation method for overall errors and by a difference plot. This approach developed a precise estimation model, with less than 15% MAPE for most of the nutrients analyzed. The gradient boosting regression algorithm proved to be suitable for small datasets (less than 400 individuals) with both larger and smaller CV. For this analysis, it surpassed other regression algorithms without overfitting the data. It is a reliable methodology for crop spectral analysis, as larger datasets are difficult to create. Although this model was tested in commercial citrus production systems, it could be adapted to other crop systems. This new technology will allow the generation of prescription maps for variable rate application of fertilizers based on UAV imagery.

References

Abdulridha, J., Ampatzidis, Y., Qureshi, J., & Roberts, P. (2020a). Laboratory and UAV-based identification and classification of tomato yellow leaf curl, bacterial spot, and target spot diseases in tomato utilizing hyperspectral imaging and machine learning. *Remote Sensing*, 12(17), 2732. https://doi.org/10.3390/rs12172732.

Abdulridha, J., Ampatzidis, Y., Roberts, P., & Kakarla, S. C. (2020b). Detecting powdery mildew disease in squash at different stages using UAV-based hyperspectral imaging and artificial intelligence. *Biosystems Engineering*. https://doi.org/10.1016/j.biosystemseng.2020.07.001.

Abdulridha, J., Batuman, O., & Ampatzidis, Y. (2019b). UAV-based remote sensing technique to detect citrus canker disease utilizing hyperspectral imaging and machine learning. *Remote Sensing, 11*(11), 1373. https://doi.org/10.3390/rs11111373.

Abdulridha, J., Ehsani, R., Abd-Elrahman, A., & Ampatzidis, Y. (2019a). A remote sensing technique for detecting laurel wilt disease in avocado in presence of other biotic and abiotic stresses. *Computers and Electronics in Agriculture*, 156, 549–557. https://doi.org/10.1016/j.compag.2018.12.018

Ampatzidis, Y., Partel, V., & Costa, L. (2020). Agroview: Cloud-based application to process, analyze and visualize UAV-collected data for precision agriculture applications utilizing artificial intelligence. *Computers and Electronics in Agriculture, 174,* 105157. https://doi.org/10.1016/j.compag.2020.105457.

Ampatzidis, Y., Partel, V., Meyering, B., & Albrecht, U. (2019). Citrus rootstock evaluation utilizing UAV-based remote sensing and artificial intelligence. *Computers and Electronics in Agriculture*, 164, 104900. https://doi.org/10.1016/j.compag.2019.104900

Bahtiar, A. R., Santoso, A. J., & Juhariah, J. (2020). Deep learning detected nutrient deficiency in chili plant. *8th international conference on information and communication technology (ICoICT)* (pp. 1–4). IEEE.

Behlau, F., Belasque Jr, J., Graham, J., & Leite, R. Jr. (2010). Effect of frequency of copper applications on control of citrus canker and the yield of young bearing sweet orange trees. *Crop Protection*, 29(3), 300–305. https://doi.org/10.1016/j.cropro.2009.12.010

Boman, B. (2012). Citrus best management practices. *Advances in Citrus Nutrition*. https://doi.org/10.1007/978-94-007-4171-3_26.

Burud, I., Lange, G., Morten, L., Bleken, E., Grimstad, L., & From, P. J. (2017). Exploring robots and UAVs as phenotyping tools in plant breeding. *IFAC-PapersOnLine*, 50(1), 11479–11484. https://doi.org/10.1016/j.ifacol.2017.08.1591

Cao, J., Cheng, C., Yang, J., & Wang, Q. (2015). Pathogen infection drives patterns of nutrient resorption in citrus plants. *Scientific Reports*, 5, 14675. https://doi.org/10.1038/srep14675

Cerreta, J., Hanson, A., Martorella, J. E., & Martorella, S. (2018). Using 3 dimensional health vegetation index point clouds to determine HLB infected citrus trees. *Journal of Aviation/Aerospace Education and Research*. https://doi.org/10.15394/jaaer.2018.1776.

Coelho, A., Rosalen, D., & Faria, R. (2018). Vegetation indices in the prediction of biomass and grain yield of white oat under irrigation levels. *Pesquisa Agropecuária Tropical*, 48(2), 109–117. https://doi.org/10.1590/1983-40632018v4851523

Costa, L., Archer, L., Ampatzidi, Y., Casteluci, L., Caurin, G. A. P., & Albrecht, U. (2021). Determining leaf stomatal properties in citrus trees utilizing machine vision and artificial intelligence. *Precision Agriculture, 22,* 1107–1119. https://doi.org/10.1007/s11119-020-09771-x.

Costa, L., Nunes, L., & Ampatzidis, Y. (2020a). A new visible band index (vNDVI) for estimating NDVI values on RGB images utilizing genetic algorithms. *Computers and Electronics in Agriculture*, 172(May), 105334. https://doi.org/10.1016/j.compag.2020.105334

Csillik, O., Cherbini, J., Johnson, R., Lyons, A., & Kelly, M. (2018). Identification of citrus trees from unmanned aerial vehicle imagery using convolutional neural networks. *Drones*, 2(4), 39. https://doi.org/10.3390/drones2040039

Cruz, A., Ampatzidis, Y., Pierro, R., Materazzi, A., Panattoni, A., Bellis, L. D., et al. (2019). Detection of grapevine yellows symptoms in *Vitis vinifera* L. with artificial intelligence. *Computers and Electronics in Agriculture, 157,* 63–76. https://doi.org/10.1016/j.compag.2018.12.028.

Galvez-Sola, L., Garcia-Sanchez, F., Perez-Perez, J., Gimeno, V., Navarro, J. M., Moral, R., et al. (2015). Rapid estimation of nutritional elements on citrus leaves by near infrared reflectance spectroscopy. *Frontiers in Plant Science*, 6, 571. https://doi.org/10.3389/fpls.2015.00571

Garza, B. N., Ancona, V., Enciso, J., Perotto-Baldiviesco, H. L., Kunta, M., & Simpson, C. (2020). Quantifying citrus tree health using true color UAV images. *Remote Sensing*, 12(1), 170. https://doi.org/10.3390/rs12010170

Hassan, M. A., Yang, M., Rasheed, A., Yang, G., Reynolds, M., Xia, X., et al. (2019). A rapid monitoring of NDVI across the wheat growth cycle for grain yield prediction using a multi-spectral UAV platform. *Plant Science*, 282, 95–103. https://doi.org/10.1016/j.plantsci.2018.10.022

Havlin, J., & Soltanpour, P. (1980). A nitric acid plant tissue digest method for use with inductively coupled plasma spectrometry. *Communications in Soil Science and Plant Analysis*, 11(10), 969–980. https://doi.org/10.1080/00103628009367096.

Kadyampakeni, D. M., & Morgan, T. K. (2020). Nutrition of florida citrus trees, third edition: SL253/SS478, Rev. 3/2020. EDIS 2020. https://doi.org/10.32473/edis-ss478-2020

Kunwar, S., Grosser, J., Gmitter, F. G. Jr., Castle, W. S., & Albrecht, U. (2021). Field performance of 'Hamlin' orange trees grown on various rootstocks in HLB-endemic conditions. *HortScience*, 56(2), 244–253. https://doi.org/10.21273/HORTSCI15550-20

Lenaerts, B., Collard, B. C. Y., & Demont, M. (2019). Review: Improving global food security through accelerated plant breeding. *Plant Science*, 287, 110207. https://doi.org/10.1016/j.plantsci.2019.110207

Li, S. X., Wang, Z. H., Miao, Y. F., & Li, S. Q. (2014). Soil organic nitrogen and its contribution to crop production. *Journal of Integrative Agriculture*. https://doi.org/10.1016/S2095-3119(14)60847-9.

Mirasi, A., Mahmoudi, A., Navid, H., Kamran, K., & Asoodar, M. (2019). Evaluation of sum-NDVI values to estimate wheat grain yields using multi-temporal Landsat OLI data. *Geocarto International*, 36(12), 1309–1324. https://doi.org/10.1080/10106049.2019.1641561

Moghimi, A., Pourreza, A., Zuniga-Ramirez, G., Williams, L. E., & Fidelibus, M. W. (2020). A novel machine learning approach to estimate grapevine leaf nitrogen concentration using aerial multispectral imagery. *Remote Sensing, 12*(21), 3515. https://doi.org/10.3390/rs12213515.

Morgan, K. T., & Graham, J. H. (2019). Nutrient status and root density of Huanglongbing-affected trees: Consequences of irrigation water bicarbonate and soil pH mitigation with acidification. *Agronomy*, 9(11), 746. https://doi.org/10.3390/agronomy9110746

Mylavarapu, R. S., Harris, W. G., & Hochmuth, G. J. (2016). Agricultural soils of Florida. EDIS, SL441. Retrieved October 26, 2021, from https://edis.ifas.ufl.edu/publication/SS655.

Osco, L., Ramos, A. P. M., Pereira, D. R., Moriya, Ã. A. S., Imai, N. N., Matsubara, E. T., et al. (2019). Predicting canopy nitrogen content in citrus-trees using random forest algorithm associated to spectral vegetation indices from UAV-Imagery. *Remote Sensing*, 11(24), 2925. https://doi.org/10.3390/rs11242925

Osco, L. P., Ramos, A. P. M., Faita Pinheiro, M. M., Moriya, Ã. A. S., Imai, N. N., Estrabis, N., et al. (2020). A machine learning framework to predict nutrient content in Valencia-orange leaf hyperspectral measurements. *Remote Sensing, 12*(6), 906. https://doi.org/10.3390/rs12060906.

Partel, V., Kakarla, S. C., & Ampatzidis, Y. (2019a). Development and evaluation of a low-cost and smart technology for precision weed management utilizing artificial intelligence. *Computers and Electronics in Agriculture*, 157, 339–350. https://doi.org/10.1016/j.compag.2018.12.048

Partel, V., Nunes, L., Stansly, P., & Ampatzidis, Y. (2019b). Automated vision-based system for monitoring Asian citrus psyllid in orchards utilizing artificial intelligence. *Computers and Electronics in Agriculture*, 162, 328–336. https://doi.org/10.1016/j.compag.2019.04.022

Qamar-uz-Zaman, & Schumann, A. (2006). Nutrient management zones for citrus based on variation in soil properties and tree performance. *Precision Agriculture*, 7, 45–63. https://doi.org/10.1007/s11119-005-6789-z

Ranganathan, P., Pramesh, C. S., & Aggarwal, R. (2017). Common pitfalls in statistical analysis: Measures of agreement. *Perspectives in Clinical Research*, 8(4), 187

Reynolds, D., Baret, F., Welcker, C., Bostrom, A., Ball, J., Cellini, F., et al. (2019). What is cost-efficient phenotyping? Optimizing costs for different scenarios. *Plant Science*, 282, 14–22. https://doi.org/10.1016/j.plantsci.2018.06.015

Scagel, C., Bi, G., Fuchigami, L., & Regan, R. (2007). Seasonal variation in growth, nitrogen uptake and allocation by container-grown evergreen and deciduous rhododendron cultivars. *HortScience*, 42(6), 1440–1449. https://doi.org/10.21273/HORTSCI.42.6.1440

Shaw, B., Thomas, T. H., & Cooke, D. T. (2002). Responses of sugar beet (*Beta vulgaris* L.) to drought and nutrient deficiency stress. *Plant Growth Regulation, 37*(1), 77–83.

Stammer, A., & Mallarino, A. (2018). Plant tissue analysis to assess phosphorus and potassium nutritional status of corn and soybean. *Soil Science Society of America Journal*, 82(1), 260–270. https://doi.org/10.2136/sssaj2017.06.0179

Sweeny, R. (1989). Generic combustion method for determination of crude protein in feeds: Collaborative study. *Journal of Association of Official Analytical Chemists*, 72(5), 770–774. https://doi.org/10.1093/jaoac/72.5.770

Toplu, C., Ugyur, V., Kaplankiran, M., Demirkeser, T., & Yildiz, E. (2011). Effect of citrus rootstocks on leaf mineral composition of 'Okitsu', 'Clausellina', and 'Silverhill' mandarin cultivars. *Journal of Plant Nutrition*, 35(9), 1329–1340. https://doi.org/10.1080/01904167.2012.684125

Uygur, V., & Yetisir, H. (2009). Effects of rootstocks on some growth parameters, phosphorous and nitrogen uptake watermelon under salt stress. *Journal of Plant Nutrition, 32*(4), 629–643. https://doi.org/10.1080/01904160802715448.

Vijayakumar, V., Costa, L., & Ampatzidis, Y. (2021). Prediction of citrus yield with AI using ground-based fruit detection and UAV imagery. Paper number: 2100493, St Joseph, MI, USA: ASABE. https://doi.org/10.13031/aim.202100493

Vincent, C., Vashisth, T., Zekri, M., & Albrecht, U. (2021). 2021–2022 Florida citrus production guide: Grove planning and establishment. UF/IFAS EDIS. Retrieved October 26, 2021, from https://edis.ifas.ufl.edu/publication/hs1302.

Watson, P., & Petrie, A. (2010). Method agreement analysis: A review of correct methodology. *Theriogenology, 73*(9), 1167–1179. https://doi.org/10.1016/j.theriogenology.2010.01.003.

Xie, Q., Huang, W., Liang, D., Chen, P., Wu, C., Yang, G., et al. (2014). Leaf area index estimation using vegetation indices derived from airborne hyperspectral images in winter wheat. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(8), 3586–3594. https://doi.org/10.1109/JSTARS.2014.2342291

Ye, X., Sakai, K., Sasao, A., & Asada, S. (2007). Estimation of citrus yield from canopy spectral features determined by airborne hyperspectral imagery. *International Journal of Remote Sensing*, 30(18), 4621–4642. https://doi.org/10.1080/01431160802632231

Yilmaz, B., Cimen, B., Incesu, M., Uysal, K., & Yesiloglu, T. (2018). Rootstock influences on seasonal changes in leaf physiology and fruit quality of rio red grapefruit variety. *Applied Ecology and Environmental Research*, 16(4), 4065–4080. https://doi.org/10.1080/01904167.2012.684125

Zhang, Y., & Haghani, A. (2015). A gradient boosting method to improve travel time prediction. *Transportation Research Part C: Emerging Technologies*, 58, 308–324. https://doi.org/10.1016/j.trc.2015.02.019