



## OPEN ACCESS

### EDITED BY

Sara Álvarez,  
Instituto Tecnológico Agrario de Castilla  
y León, Spain

### REVIEWED BY

Muhammad Ather Nadeem,  
University of Sargodha, Pakistan  
Sergio Vélez,  
Wageningen University and Research,  
Netherlands

### \*CORRESPONDENCE

Heiplanmi Rymbai  
✉ rymbaihort@gmail.com  
Anoop Kumar Srivastava  
✉ aksrivastava2007@gmail.com  
Vasileios Ziogas  
✉ ziogas@elgo.gr

RECEIVED 06 October 2025

REVISED 05 February 2026

ACCEPTED 12 February 2026

PUBLISHED 04 March 2026

### CITATION

Rymbai H, Parent L-E, Srivastava AK,  
Talang H, Verma VK, Ziogas V,  
Mawlein J, Ramesh T, Hazarika S and  
Baishya LK (2026) Machine learning  
tools-based diagnosis of soil  
nutrient constraints to increase the  
productivity of citrus orchards.  
*Front. Agron.* 8:1719690.  
doi: 10.3389/fagro.2026.1719690

### COPYRIGHT

© 2026 Rymbai, Parent, Srivastava, Talang,  
Verma, Ziogas, Mawlein, Ramesh, Hazarika  
and Baishya. This is an open-access article  
distributed under the terms of the  
[Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/)  
(CC BY). The use, distribution or  
reproduction in other forums is  
permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication  
in this journal is cited, in accordance  
with accepted academic practice. No  
use, distribution or reproduction is  
permitted which does not comply with  
these terms.

# Machine learning tools-based diagnosis of soil nutrient constraints to increase the productivity of citrus orchards

Heiplanmi Rymbai<sup>1\*</sup>, Leon-Etienne Parent<sup>2</sup>,  
Anoop Kumar Srivastava<sup>3\*</sup>, Hammylliende Talang<sup>1</sup>,  
Veerendra Kumar Verma<sup>1</sup>, Vasileios Ziogas<sup>4\*</sup>,  
Joiedevivreson Mawlein<sup>1</sup>, Thangavel Ramesh<sup>1</sup>,  
Samarendra Hazarika<sup>1</sup> and Lohit Kumar Baishya<sup>5</sup>

<sup>1</sup>ICAR Research Complex for North Eastern Hill Region, Meghalaya, India, <sup>2</sup>Department of Soils and Agrifood Engineering, Laval University, Quebec, QC, Canada, <sup>3</sup>Indian Council of Agricultural Research (ICAR)-Central Citrus Research Institute, Nagpur, Maharashtra, India, <sup>4</sup>Hellenic Agricultural Organization – DIMITRA (ELGO – DIMITRA), Institute of Olive Tree, Subtropical Crops and Viticulture, Chania, Crete, Greece, <sup>5</sup>Indian Council of Agricultural Research (ICAR)-Indian Agricultural Research Institute, Assam, India

**Introduction:** The multiple nutritional disorders producing the early decline of citrus productivity are commonly observed across the citrus belts of northeast India. This situation is further compounded by a mismatch between annual addition and consumption of fertilizers, in the backdrop of an erroneous diagnosis of nutrient imbalance. In this background, we attempted to diagnose nutrient balance in Khasi mandarin (*Citrus reticulata* Blanco) orchards using several diagnostic tools comprising machine learning (ML) tools.

**Methods:** A database of soil available nutrients (KMnO<sub>4</sub>-N, Brays-P, NH<sub>4</sub>OAc-K, DTPA-Fe, DTPA-Mn, DTPA-Cu, DTPA-Zn) and fruit yield documenting 180 Khasi mandarin orchards of seedling origin (10–30 years old with row-to-row distance of 4 m and trees 6 m apart) raised under rainfed conditions in the Meghalaya state of northeast India. Diagnosis methods were compared: the sufficiency level of available nutrients (SLAN), the basic cation saturation ratio (BCSR), the compositional nutrient diagnosis (CND), the diagnosis and recommendation integrated system (DRIS) and ML tools like random forest and xgboost.

**Results:** Soil test interpretation of a low-yielding and nutritionally imbalanced orchard differed among diagnostic methods. DRIS predicted deficient-to-low concentrations of Zn, Ca, P, N, and K; other nutrients like Fe, Cu, Mn, and Mg were at optimum-to-high concentrations. CND standards diagnosed Zn deficiency and Cu excess with potential agronomic manifestations for early decline in productivity. SLAN interpretation was highly skewed for Ca; moderately skewed for Mg, Cu and Zn, and unskewed for N, P, K, Fe and Mn. The accuracy of ML regression models relating nutrient expressions to fruit yield was invariably high, followed by SLAN. The ML xgboost regression model exhibited the highest accuracy in predicting fruit yield from soil test. Conversely, the BCSR, which considers only three cationic dual ratios, was inaccurate. There were distortions when relating concentration values to DRIS indices to determine ‘optimum’ concentration ranges. The ML classification models showed that concentration

values were also less accurate than *clr* to classify data as true negative (TN) or true positive (TP). The *xgboost* classification model showed optimum ranges for N, P, Ca, Mg, Mn, and Cu; whereas K, Fe, and Zn fell below the lower limits.

**Conclusion:** ML, hence, as a classification approach, aided in discarding cases of poor yields and high yields showing luxury nutrient consumption or suboptimal nutrient levels. The soil test standards and site-by-site comparisons can further support site-specific nutrient management and precision fertilization. Considering Zn as the most deficient nutrient, Zn biofortification interventions are recommended to increase fruit yield and quality in Khasi mandarin.

#### KEYWORDS

BCSR, *Citrus reticulata*, DRIS norms, Khasi mandarin, ML tools, nutrient management, SLAN, sustainable production

## 1 Introduction

Citrus is one of the important fruit crops grown worldwide. Countries like China, Brazil, India, Mexico, and the USA account for nearly 60% of global production. Mandarins (*Citrus reticulata* Blanco) dominate over sweet oranges and acid limes/lemons, totaling an annual production of 167.82 million metric Mg from an area of 10.62 million ha with a current productivity of 15.80 Mg ha<sup>-1</sup> (Megagrams = Mg) (USDA, 2024). In India, citrus production is estimated at 14.8 million Mg occupying an area of 1.09 million ha (Agricultural Statistics, 2024). 'Khasi' mandarin is grown as seedlings primarily on coarse textured acid soils under rainfed conditions, representing a diversity of landscapes represented predominantly by hillside slopes, frequently erosion and loss of fertile topsoil (Rymbai et al., 2022). The average productivity of 4.9 Mg ha<sup>-1</sup>, far below 13.6 Mg ha<sup>-1</sup> as the national average (Horticultural Statistics at a Glance, 2024), is the major cause for concern, calling for an in-depth scientific investigation. Low citrus yield is therefore a serious issue in the Indo-Myanmar region due to multiple nutrient deficiencies, the world's primary center of citrus diversity (Rymbai et al., 2024).

Current fertilization programs heavily rely on general guidelines. However, such programs are still not customized in relation to orchard age and crop phenology, resulting in premature decline in orchard productivity (Srivastava and Singh, 2004). Despite large scale deficiencies of nutrients like N, K, P, Zn, Ca, Mg, Cu, and Mn (Barthakur, 2018; Srivastava and Singh, 2004), the response of mandarins to fertilization, however, continues to be inconsistent. Regional standards should be customized to enhance diagnostic accuracy (Ferrández-Cámara et al., 2021) and relieve nutrient constraints.

Soil tests are still not common in northeastern India (Srivastava and Singh, 2005) compared to other regions (Obreza and Morgan, 2008). The diagnostic tools like SLAN and BCSR are commonly used to interpret soil test results (Dahnke and Olson, 1990). The SLAN diagnoses nutrients separately, whereas BCSR considers relationships among the basic cations (Ca, Mg, and K). On the other hand, SLAN is much simpler, but often less precise than DRIS or CND at detecting imbalances and predicting yield when applied across a variety of perennial woody crops (Morais et al., 2019). The

BCSR rarely gives consistent yield benefits over more conventional liming/fertility guidance and is poorly validated for nutrient-limitations in many crops, including tree crops. BCSR used for soil physical/long-term management is not as precise as tissue-based yield diagnostics (Chaganti et al., 2021). There have been some disagreements regarding the effective interpretation of soil test results using SLAN (Dahnke and Olson, 1990) and BCSR (Culman et al., 2021). On the other hand, DRIS combines nutrient indices and all available dual ratios from soil test reports, including soil cationic dual ratios (Beaufils and Sumner, 1976). DRIS and CND are suitable to diagnose nutrient balance after treatment in citrus and illustrate the practical sensitivity of DRIS and CND relative to univariate thresholds (Huang et al., 2012). DRIS and CND are reported to be equivalent in linking tissue data to orchard performance (Labaieda et al., 2018).

Log ratios have been developed to address the back-looping of added proportions upon closure and control data distribution (Aitchison, 1986). Centered log ratios combine dual ratios into multi-nutrient variables to address nutrient interactions as CND diagnosis (Parent and Dafir, 1992). The non-parametric ML models are known to predict the yield and quality of fruit crops from several explanatory variables, including soil and tissue test results. Interactions can be handled effectively during modelling (Chlingaryan et al., 2018; Huynh-Thu and Geurts, 2019; Parent, 2025a). ML tools and CND methods offer an alternative to SLAN, BCSR, and DRIS. ML models and CND can customize fertilization for precision farming based on site-specific conditions documented in a database (Parent, 2025b). ML regressors (PLS-R, Random Forest, SVM, XGBoost, CNNs) also achieve moderate to high R<sup>2</sup> for several nutrients (often R<sup>2</sup> > 0.60 for P, K, B; lower for others) and classification accuracies for deficiency detection frequently >80% in controlled datasets (Acosta et al., 2023). ML is promising (good R<sup>2</sup>, classification scores) in citrus trials (Costa et al., 2022). The ML could thus improve diagnostic coverage and enables precision interventions.

It is hypothesized that (1) ML modelling can provide higher diagnostic accuracy compared to classical soil fertility evaluation systems, and (2) nutrient diagnoses derived from SLAN, BCSR, DRIS, and CND differ in Khasi mandarin orchards of northeast India. Our eventual objective is to identify the underlying causes of

nutrient imbalances in Khasi mandarin orchards using diagnostic tools supporting site-specific fertilization decisions.

## 2 Materials and methods

### 2.1 Citrus sites and experimental orchards

A total of 180 10- to 30-year-old orchards (250–1200 trees by orchard on average) of ‘Khasi’ mandarin were documented during 2016 to 2023 period in multiple years in the Himalayan foothill tracts of Meghalaya state, India (25.03333 to 26.11667 N latitude; 89.8000’ to 92.83333 E longitude, and altitude 100 to 1400 m). The other contextual variables of identified orchards are highlighted in [Supplementary Table 1](#). The terrain consisted of numerous penepains, hilly relief, plateaus, rocky, uneven hills, and several river and stream valleys with sporadically distributed human settlements. The uninterrupted range of hills is characterized by steep slopes in the south, and by a highly undulating and irregular terrain in the west and the north. Tree spacings were 4 m row-wise and 6 m as inter-plant distance. The region is characterized by a humid subtropical monsoon climate, highly suitable for growing mandarins. At lower altitudes, particularly on the southern and western sides, the climate tends to be warmer than at higher altitudes. Annual precipitations ranged from 4000 to 11,436 mm (mean annual rainfall of 3135 mm), temperatures varying between 5°C and 32°C (mean temperature in summer and winter are 31°C and 4°C, respectively, with average annual temperature of 20°C). Relative humidity ranged from 76% to 92%.

The fruit yield of the test trees was determined by visual observation on the performance of the majority of trees and later identifying 4–5% trees within this group in an orchard. The number of fruits in the selected trees was counted during harvesting, and the fruit weight (50 fruits per tree) was measured. The average fruit weight was multiplied by the number of fruits to give yield per tree expressed as metric Mg per hectare ( $\text{Mg ha}^{-1}$ ). These orchards were maintained under uniform management practices.

### 2.2 Collection and analysis of soil samples

Soil samples were collected from the 0–20 cm layer within the skirt belt zone trees (near the canopy drip line, having maximum concentration of feeder roots). The composite sampling consisted of 10–25 subsamples collected from the uniformly performing block of trees in each orchard. The soil samples were collected at a time matching the maximum expansion at growth stage of leaves of the current season flush (July–September, 4–6 months-old growth of new leaves). These samples emerged from the same season of multiple years.

The soil samples were mixed thoroughly, all the clods broken, debris removed, air-dried, ground and sieved through 2-mm sieve, then subjected to various analyses. The analysis of soil available nitrogen (available N) was performed using the alkaline potassium permanganate ( $\text{KMnO}_4$ ) steam distillation method, as described in [Subbiah and Asija \(1956\)](#). The available phosphorous was determined by the Bray-1 extraction method, followed by

molybdenum blue colorimetric determination using a spectrophotometer at wavelength of 660 nm (Model: Spectronic 21D, USA) ([Bray and Kurtz, 1945](#)). Available potassium was extracted using neutral 1 N ammonium acetate ( $\text{NH}_4\text{OAc}$ ) solution at a soil-to-extractant ratio of 1:2 and quantified by flame emission photometry (Model: Elico 24C, India) ([Chapman and Pratt, 1978](#)). Soil micronutrients (Zn, Cu, Mn, and Fe) were extracted using the DTPA extractant (0.005 M DTPA, 0.01 M  $\text{CaCl}_2$ , and 0.1 M triethanolamine) adjusted to pH 7.3 ([Lindsay and Norvell, 1978](#)) and quantified by atomic absorption spectrophotometry (Model: GBC 908, Australia).

### 2.3 Data distribution

Data distribution was classified as symmetrical, where skewness is observed between -0.5 and +0.5; while skewness is considered high if values are observed as less than -1 or greater than +1 ([Hatemi et al., 2022](#)).

### 2.4 Computations through SLAN and BCSR

The SLAN is a philosophy of fertilizing the crop aimed to maximize crop yield with minimum nutrient use, excluding any concern about building up the nutrient stock of soil. Nutrient availability is addressed individually based on established critical levels ([Black, 1993](#)) for low/very low, medium, high/very high categories, depending on potential crop response. On the other hand, BCSR calculates proportions of the cation exchange capacity of soil occupied by basic cations  $\text{Ca}^{2+}$ ,  $\text{Mg}^{2+}$ ,  $\text{Na}^+$  and  $\text{K}^+$ , and computes cationic ratios. Values thought to be ‘ideal’ are as follows ([McLean, 1984](#)): 65–75% Ca saturation of the cation exchange capacity (CEC), 2.5–5% Mg saturation, 2.5–5% K saturation, 2–10% exchangeable acidity, 7:1 as Ca/Mg molar ratio, 15:1 Ca/K molar ratio, and 3:1 Mg/K molar ratio ([Dahnke and Olson, 1990](#)).

### 2.5 Computations through DRIS

For the establishment of DRIS ‘norms’, the sample population was carefully separated into high-productivity (reference population) and low-productivity (non-reference population) orchards using the critical value approach ( $R^2$ ) as per the statistical procedure of [Cate and Nelson \(1971\)](#). The yield value was initially listed in descending order. The adjusted sum of squares of the two populations was estimated from the low yield data, progressing to each successive yield value known as  $R^2$  ([Figure 1](#)). The sample population with  $R^2$  value greater than the maximum yield data was regarded as the “high yielding” or “reference population”; whereas the remaining samples were regarded as the “low productivity” population.

Dual ratios between nutrients *A* and *B* were computed as  $A/B$  or  $B/A$ . The variance ratio of  $A/B$  or  $B/A$  ratios between both subpopulations were computed as per F-test as follows ([Walworth and Sumner, 1987](#)):

$$F_{A/B} = \frac{S^2(\frac{A}{B})_{\text{non-reference population}}}{S^2(\frac{A}{B})_{\text{reference population}}}$$

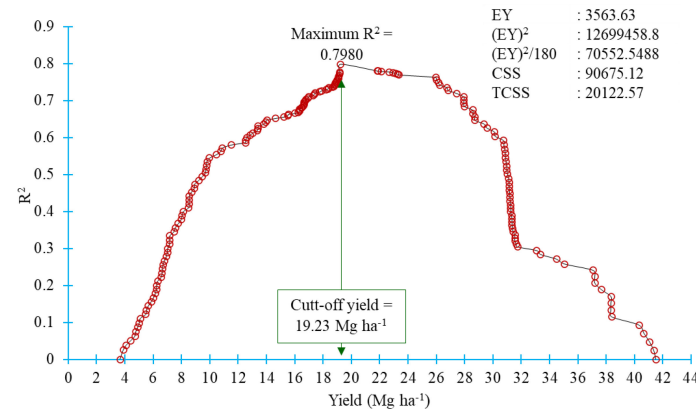


FIGURE 1

Determining the cut-off value to separate the high - and low - yield orchards (180 orchards) using the critical value approach ( $R^2$ ) as per the statistical procedure of Cate and Nelson (1971). EY – expected yield, CSS – corrected sum of squares for 180 orchards, TCSS – total corrected sum of squares for 180 orchards, Max  $R^2$  – maximum  $R^2$ .

$$F_{B/A} = \left[ \frac{S^2\left(\frac{B}{A}\right)_{\text{non-reference population}}}{S^2\left(\frac{B}{A}\right)_{\text{reference population}}} \right]$$

The variance showing the highest F-value across iterations was selected to set a statistically significant yield cutoff and compute DRIS norms and functions ( $f$ ) for the dual ratio, as follows:

$$\text{For } A/B < a/b: f\left(\frac{A}{B}\right) = \left[ 1 - \frac{a/b}{A/B} \right] \frac{100\kappa}{CV\%}$$

$$\text{If } A/B = a/b: f\left(\frac{A}{B}\right) = 0$$

$$\text{If } A/B > a/b: f\left(\frac{A}{B}\right) = \left[ \frac{A/B}{a/b} - 1 \right] \frac{100\kappa}{CV\%}$$

where  $A/B$  is the dual ratio of the diagnosed specimen,  $a/b$  is the corresponding mean of the high-yielding specimens,  $CV$  (%) is the coefficient of variation of the dual ratio of the high-yielding specimens, and  $\kappa$  (= 10) is a constant to control the number of decimals.

DRIS index for nutrient A was computed as the sum of DRIS functions involving nutrient A, where DRIS functions was multiplied by +1 if A is at numerator and by -1 if A is at denominator. The sum of DRIS indices is zero as a result of the symmetry inherent to the balance concept. A critical soil test value ( $n_i$ ) where crop yield exceeded 90% of the maximum yield was computed from the ratio ( $Q$ ) between yield ( $Y$ ) and ( $n_i$ ) as follows (Maia et al., 2001):

$$Q = \frac{Y}{n_i}$$

$Q$  was determined similarly, using equations

$$CLi = \frac{(1.281552 \cdot s1) + m1}{(1.281552 \cdot s2) + m2}$$

$m1$  and  $s1$  are the arithmetic mean (AM) and standard deviation (SD) of  $Y$ , and  $m2$  and  $s2$  are the AM and SD of  $Q$ , respectively. The tabulated value 1.282 was used as the standard distribution.

The DRIS nutritional balance index (NBI) was computed across DRIS indices irrespective of sign, as follows across nutrients A to N (Beaufils, 1973):

$$NBI = |I \text{ DRIS } A| + |I \text{ DRIS } B| + |I \text{ DRIS } C| + \dots + |I \text{ DRIS } N|$$

The NBI was averaged as NBIA as follows across  $n$  DRIS indices (Wadt 1996):

$$NBIA = \frac{NBI}{n}$$

As many as five interpretation classes were used for establishing potential fertilization response (PFR) by comparing each nutrient index ( $I_N$ ) to NBIA (Supplementary Table 2). Classes were further grouped as follows (Silva et al., 2005): (LL) limiting (high insufficiency and insufficiency), (NL) non-limiting (balance), and (LE) limiting by excess (excess and high excess). The statistical test (chi-square, 5%) compared PFR classes using observed (OF) and expected (EF) frequencies computed as follows (Urano et al., 2006):

The statistical test (chi-square, 5%) compared PFR classes using observed (OF) and expected (EF) frequencies computed as follows (Urano et al., 2006):  $EF\% = \left[ \frac{\text{total number of plots evaluated} / \text{total number of nutrients}}{\text{total number of plots evaluated}} \right] \times 100$ ; and  $OF\% = \left[ \frac{\text{number of plots in which the nutrient was } p}{\text{total number of plots evaluated}} \right] \times 100$ .

$$EF\% = \left[ \frac{\frac{\text{Total number of plots evaluated}}{\text{Total number of nutrients}}}{\text{Total number of plots evaluated}} \right] \times 100$$

$$OF\% = \left[ \frac{\text{Number of plots in which the nutrient was } p}{\text{Total number of plots evaluated}} \right] \times 100$$

Raw concentration values are more familiar to stakeholders than DRIS indices. Raw concentrations were related to DRIS indices to

allow computing concentration means and standard deviations ( $s$ ) for each nutrient across the entire population, as  $mean \pm \frac{4}{3}s$  and  $mean \pm \frac{8}{3}s$ . Nutrients showing values less than  $mean - \frac{8}{3}s$  were deemed deficient. Values between  $mean \pm \frac{4}{3}s$  and  $mean \pm \frac{8}{3}s$  were diagnosed in the high range, while those above  $mean + \frac{8}{3}s$  were considered excessive. Sufficiency ranges were set as lower and upper bounds of  $mean \pm \frac{4}{3}s$  (Srivastava and Singh, 2008).

## 2.6 Computations through CND

Soil data are compositional data, i.e. strictly positive and intrinsically multivariate data (Aitchison, 1986). Multivariate relationships as dual ratios were transformed into dual log ratios to project them into the real space ( $\pm\infty$ ) appropriate to conduct statistical analyses (Steel and Torrie, 1980). Dual log ratios are additive as centered log ratios ( $clr$ ) as follows for the component  $x_i$  in a  $D$ -parts composition (Greenacre et al., 2023):

$$clr_{x_i} = \ln\left(\frac{x_i}{g(x)}\right) = \frac{1}{D} \left( \ln\left(\frac{x_i}{x_1}\right) + \ln\left(\frac{x_i}{x_2}\right) + \dots + \ln\left(\frac{x_i}{x_D}\right) \right)$$

where  $g(x)$  is the geometric mean across components. To close the simplex and allow back-transforming  $clr$  values into familiar concentration values, a filling value is computed by the difference between the measurement unit and the sum of quantified components. The  $clr$  is the average of dual log ratios (Greenacre et al., 2023), as follows for nutrient  $N$ :

$$\begin{aligned} & \ln\left(\frac{N}{G}\right) \\ &= \ln\left(\frac{N}{N} \times \frac{N}{P} \times \frac{N}{K} \times \frac{N}{Ca} \times \frac{N}{Mg} \times \frac{N}{S} \times \frac{N}{B} \times \frac{N}{Cu} \times \frac{N}{Zn} \times \frac{N}{Mn} \times \frac{N}{Fe} \times \dots \times \frac{N}{x_D}\right)^{\frac{1}{D}} \\ &= \frac{1}{D} \left[ \ln\left(\frac{N}{N}\right) + \ln\left(\frac{N}{P}\right) + \dots + \ln\left(\frac{N}{x_D}\right) \right] \end{aligned}$$

The back-transformation is computed as follows:

$$x_i = \exp(clr_{x_i}) \times \frac{unit}{\kappa}$$

Where  $unit$  is a measurement unit and  $\kappa$  is the sum of  $\exp(clr_{x_i})$  values.

The  $clr$ 's have Euclidean geometry, as follows (Aitchison, 1986):

$$\varepsilon = \sqrt{\sum_{i=1}^D (clr_{x_i} - clr_{x_i}^*)^2}$$

where,  $(clr_{x_i} - clr_{x_i}^*)$  is the difference between two equal-length compositions, allowing to compare a performing plant ( $clr_{x_i}^*$ ) to a defective one ( $clr_{x_i}$ ) in portions of the field or in comparable fields (Munson and Nelson, 1990). A small database can thus suffice for diagnostic purposes at local scale allowing us to meet the objective of precision farming with minimum effort for data acquisition. The  $(clr_{x_i} - clr_{x_i}^*)$  distance between two equal-length soil compositions, i.e., those for abnormal and normal performing plants, can be displayed in a histogram to illustrate relative nutrient shortage, sufficiency or excess. Soil test for the abnormal plant can be rebalanced by adding the most limiting nutrients in such amounts to reach and maintain soil test for the normal plant.

Where a database is large enough as is the case here,  $clr$  indices can be computed using nutrient standards as follows for  $clr_{x_i}$  (Parent and Dafir, 1992):

$$Index \ clr_{x_i} = \frac{(clr_{x_i} - clr_{x_i}^*)}{s_{clr_{x_i}}^*}$$

where the  $clr_i$  refers to  $t$  diagnosed specimen, and the asterisk indicates the corresponding means and standard deviations for the high-yielding and nutritionally balanced subpopulation. The workflow of steps involved with computation of DRIS norms is further presented (Figures 2A, B).

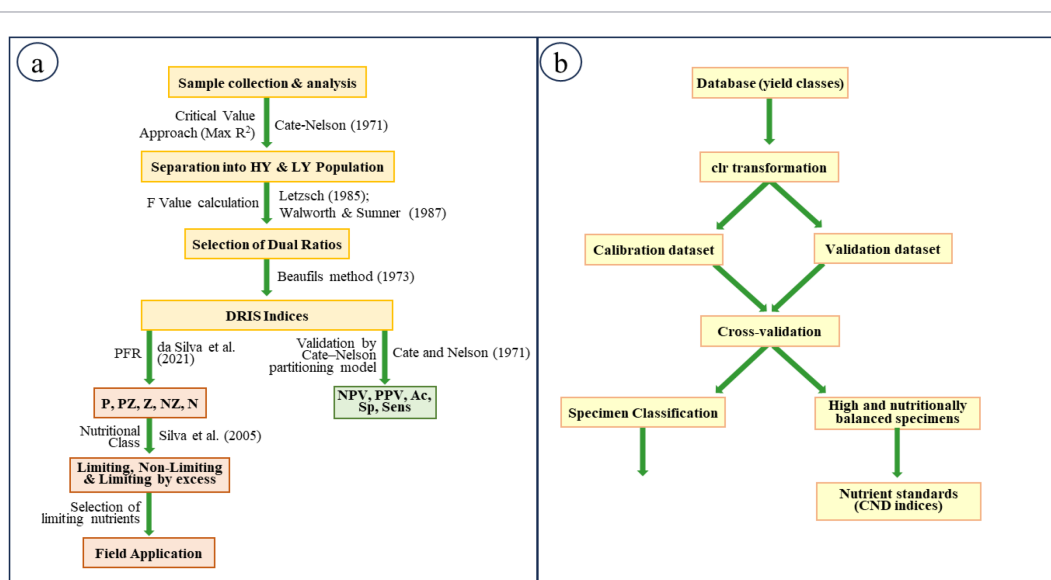


FIGURE 2

Work flowchart for computation of (A) DRIS indices and (B) Machine learning. HY – high yielding; LY– low yielding; PFR, potential fertilization response; P, positive, with high probability; PZ, positive or null, with low probability; Z, Null; NZ, Negative, with low probability; N, Negative, with high probability; NPV, negative predictive value; PPV, positive predictive value; Ac, Accuracy; Sp, Specificity; Sens, Sensitivity; clr, centered log ratios.

## 2.7 Procedural steps for ML models

Nonparametric ML decision-tree models could handle interactions in high-dimensional databases (Chlingaryan et al., 2018; Huynh-Thu and Geurts, 2019). Classification ML models were run using the Orange Data Mining freeware vs. 3.37.0 (University of Ljubljana, Slovenia). Nonparametric ML classification models set apart nutritionally balanced and imbalanced subpopulations in a confusion matrix. Random Forest is the classical ML decision tree model. Gradient boosting assembles weak learners and minimizes the loss function to make a stronger learner.

There was no missing data. Concentration and log ratio-transformed data were analyzed for skewness (Hatem et al., 2022). Crop yield was the target variable. Concentration and log ratio-transformed data were the explanatory variables. There were no categorical variables, hence no need for stratification. The entire procedure involved following important steps:

### 2.7.1 Model hyperparameters

Classification ML models were run using the Orange Data Mining freeware (University of Ljubljana, Slovenia). Nonparametric ML classification models set apart nutritionally balanced and imbalanced subpopulations in a confusion matrix. Random forest is the classical ML decision tree model. Gradient boosting assembles weak learners and minimizes the loss function to make a stronger learner. Hyperparameters were those suggested by the Orange Data Mining software. For Random Forest, there were ten trees and at least five subsets per split. For XGBoost, there were 100 trees, a learning rate of 0.300, replicable training, six individual trees as limit depth, 1.00 as fraction of training instances, features for each tree, features for each level, and features for each split.

### 2.7.2 Variable importance ranking

Different nutrients such as Ca (clr Ca 90.078) ranked as first discriminating variable between the high- versus low-yielding subpopulations, followed by Cu (clr Cu 0.054), Zn (clr Zn 0.049), Mg (clr Mg 0.043), N (clr N 0.019), Fe (clr Fe 0.019), K (clr K 0.015), Mn (clr Mn 0.011), P (clr P 0.011) in increasingly decreasing order.

### 2.7.3 Method to control overfitting

There were no categorical variables, hence no need for stratification to avoid overfitting. We conducted calibration/validation tests (70:30). If the area under curve (AUC) exceeded 0.7, model performance was acceptable (Swets, 1988) and cross-validation was run across the whole database. Five, ten, and 20 splits of mutually exclusive datasets were tested in cross-validation. Ten splits returned to the most performing cross-validation models.

### 2.7.4 Addressing the spatial autocorrelation

Data were collected across a large region, not within the well-defined limits of an orchard.

## 2.7.5 Rationale for choosing random forest and XGBoost

We tested seven models provided by the software, but Random Forest and XGBoost were the most performing. The workflow of ML tools is further highlighted (Figures 2A, B).

Random Forest and XGboost ML classification models were run using the Orange Data Mining freeware vs. 3.37 (University of Ljubljana, Slovenia). Each quadrant represented a class of probable response to inputs as follows:

- True Negative (TN, Quadrant I): adequate nutritional status, as the high-yield subpopulation was correctly diagnosed as balanced. No response expected.
- False Positive (FP, Quadrant II): nutrient luxury consumption by the plant or slightly below-optimum concentration. Reduce nutrients found in relative excess.
- False Negative (FN, Quadrant III): low yield caused by another limiting factor, as the low-yield subpopulation is diagnosed as nutritionally balanced. Identify the limiting factors.
- True Positive (TP, Quadrant IV): at least one nutrient limiting crop performance. Strong response expected.

## 3 Results

### 3.1 Crop yield and soil tests

Fruit yield varied widely from less than 4 to 66 Mg ha<sup>-1</sup> (Table 1). Soil test values showed coefficients of variation varying from 15.77 to 35.72%. Soil test data used for SLAN interpretation were highly skewed for Ca; moderately skewed for yield, Mg, Cu and Zn; and symmetrical for N, P, K, Fe and Mn. All *clr* data were symmetrical.

### 3.2 Soil -based DRIS norms

DRIS norms for soils of 'Khasi' mandarin orchards (Table 2) showed the coefficients of variation ranging from 12.88 to 31.99%.

††Cut-off yield was worked out to create a reference high yielding subpopulation representing nutrient status leading to high yield using critical value approach, although it has very little effect on the final outcome of DRIS norms. The mean nutrient ratios and their coefficients of variation from high yielding subpopulation produced the DRIS norms. The variability (variance ratio) between the low and high-yielding subpopulations for specific nutrient ratios was used to select discriminatory expression of nutrients, and thus suitable for inclusion in the calculation of DRIS model.

### 3.3 DRIS- based optimal range of nutrients

DRIS indices were converted into concentration values. Applying confidence intervals as  $mean \pm \frac{4}{3}s$  and  $mean \pm \frac{8}{3}s$  for each nutrient, soil test results were interpreted (Table 3).

TABLE 1 Statistical variation in crop yield and soil test results (xD = filling value).

| Variable | Mean                      | Standard deviation | Minimum | Maximum | Coefficient of variation | Skewness |
|----------|---------------------------|--------------------|---------|---------|--------------------------|----------|
|          | <b>Mg ha<sup>-1</sup></b> |                    |         |         | <b>%</b>                 |          |
| Yield    | 25.22                     | 16.64              | 3.66    | 66.07   | 66.00                    | 0.57     |
|          | <b>mg kg<sup>-1</sup></b> |                    |         |         |                          |          |
| N        | 253.92                    | 40.05              | 150.70  | 367.82  | 15.77                    | 0.22     |
| P        | 5.51                      | 1.03               | 4.10    | 7.40    | 18.71                    | 0.46     |
| K        | 173.67                    | 38.61              | 98.33   | 247.22  | 22.23                    | -0.02    |
| Ca       | 225.66                    | 45.26              | 153.30  | 361.00  | 20.06                    | 1.24     |
| Mg       | 50.10                     | 15.00              | 24.98   | 91.48   | 29.94                    | 0.85     |
| Fe       | 29.89                     | 9.22               | 13.00   | 52.27   | 30.85                    | 0.04     |
| Mn       | 10.25                     | 2.23               | 6.01    | 16.38   | 21.72                    | -0.20    |
| Cu       | 0.99                      | 0.21               | 0.61    | 1.56    | 21.42                    | 0.59     |
| Zn       | 1.08                      | 0.39               | 0.49    | 1.97    | 35.72                    | 0.51     |
|          | <b>clr</b>                |                    |         |         |                          |          |
| N        | 1.419                     | 0.112              | 1.060   | 1.888   | –                        | 0.11     |
| P        | -2.416                    | 0.162              | -2.714  | -2.057  | –                        | 0.21     |
| K        | 1.025                     | 0.202              | 0.600   | 1.368   | –                        | -0.26    |
| Ca       | 1.295                     | 0.139              | 1.022   | 1.616   | –                        | 0.39     |
| Mg       | -0.234                    | 0.207              | -0.830  | 0.219   | –                        | -0.38    |
| Fe       | -0.761                    | 0.298              | -1.490  | -0.250  | –                        | -0.47    |
| Mn       | -1.804                    | 0.205              | -2.303  | -1.228  | –                        | -0.08    |
| Cu       | -4.139                    | 0.237              | -4.757  | -3.718  | –                        | -0.13    |
| Zn       | -4.094                    | 0.285              | -4.680  | -3.489  | –                        | -0.12    |
| xD       | 9.709                     | 0.121              | 9.376   | 9.928   | –                        | -0.44    |

TABLE 2 Soil-based DRIS norms for 'Khasi' mandarins at a yield cutoff of 19.23 Mg ha<sup>-1</sup>.

| Ratio | Mean    | CV†    | Ratio | Mean   | CV†    |
|-------|---------|--------|-------|--------|--------|
| N/Ca  | 1.268   | 12.884 | Fe/Ca | 0.153  | 27.656 |
| N/Zn  | 275.617 | 21.225 | Fe/Mg | 0.578  | 24.814 |
| P/N   | 0.021   | 13.293 | Fe/Mn | 2.902  | 22.673 |
| P/Ca  | 0.026   | 13.441 | Fe/Cu | 32.724 | 25.601 |
| P/Zn  | 5.768   | 22.458 | Fe/Zn | 33.240 | 31.987 |
| K/N   | 0.634   | 14.783 | Mn/N  | 0.042  | 13.861 |
| K/P   | 30.435  | 15.970 | Mn/P  | 2.018  | 18.157 |
| K/Ca  | 0.797   | 14.979 | Mn/K  | 0.067  | 16.999 |
| K/Zn  | 173.266 | 22.519 | Mn/Ca | 0.053  | 17.775 |
| Ca/Zn | 219.333 | 21.515 | Mn/Mg | 0.201  | 16.316 |
| Mg/N  | 0.212   | 16.386 | Mn/Zn | 11.536 | 26.185 |
| Mg/P  | 10.173  | 17.537 | Cu/N  | 0.004  | 20.225 |
| Mg/K  | 0.338   | 16.863 | Cu/P  | 0.182  | 21.349 |
| Mg/Ca | 0.267   | 17.845 | Cu/K  | 0.001  | 22.107 |
| Mg/Zn | 57.802  | 22.549 | Cu/Ca | 0.005  | 21.472 |
| Fe/N  | 0.121   | 24.891 | Cu/Mg | 0.018  | 21.941 |
| Fe/P  | 5.796   | 24.221 | Cu/Mn | 0.091  | 19.332 |
| Fe/K  | 0.193   | 26.716 | Cu/Zn | 1.037  | 28.521 |

†Coefficient of variation (%).

Estimated critical levels (90% maximum yield) were obtained within 'optimum' nutrient ranges assessed statistically.

### 3.4 Nutrient constraints and NBI

NBI described the general status of nutrient balance in citrus. NBI was negatively related to crop yield ( $r = 0.83$ ,  $p < 0.01$ ). High NBI values were associated with lower yields, and vice-versa (Supplementary Figure 1). The soil-based DRIS indices ranging from -8.09 to -2.12 diagnosed low levels of Zn, Ca, P, N and K. Nutrients like Mg, Cu, Mn, and Fe showed increasingly positive indices (ranging from 2.48 to 11.90), indicating relatively high to excessive levels (Figure 3).

### 3.5 Potential nutrient response

The statistically constructed crop response classes are presented in Table 4. It provided an insight about the high possibility (P) of improving fruit yield as a result of fertilizer application and established the crop response -based nutrient deficiencies. In contrast, some nutrients showed a low probability (N) of nutrient response upon fertilization. In the high-yielding subpopulation, the relative frequencies of possible nutrient response to fertilization of Zn were observed 10.67-12.00% higher over P, Ca and Mg fertilization (P). Soils being acidic in nature with lime application, Zn response was observed still higher by 17.33% and 18.67%,

TABLE 3 Soil test interpretation for 'Khasi' mandarin orchards in this study and comparisons with other studies.

| Nutrient            | Deficient            | Low           | Optimum       | High          | Excess  | Critical level | Optimum              |                       |
|---------------------|----------------------|---------------|---------------|---------------|---------|----------------|----------------------|-----------------------|
| mg kg <sup>-1</sup> |                      |               |               |               |         |                |                      |                       |
|                     | This study ('Khasi') |               |               |               |         |                | 'Khasi' <sup>†</sup> | 'Nagpur' <sup>‡</sup> |
| N                   | <178.32              | 178.32-213.04 | 213.04-282.46 | 282.46-317.18 | >317.18 | 262.12         | 142-222              | 94.8–154.8            |
| P                   | <3.69                | 3.69-4.71     | 4.71-6.74     | 6.74-7.76     | >7.76   | 5.84           | 5.9-8.3              | 6.6–15.9              |
| K                   | <152.03              | 152.03-181.95 | 181.95-241.81 | 241.81-271.73 | >271.73 | 209.30         | 110-270              | 146.6–311.9           |
| Ca                  | <151.96              | 151.96-185.7  | 185.70-253.20 | 253.20-286.94 | >286.94 | 217.87         | 152-281              | 401.0–601.6           |
| Mg                  | <36.48               | 36.48-44.96   | 44.96-61.90   | 61.90-70.37   | >70.37  | 53.90          | 28-79                | 85.2–369.6            |
| Fe                  | <28.03               | 28.03-35.01   | 35.01-48.99   | 48.99-55.97   | >55.97  | 43.19          | 32-49                | 10.9-25.2             |
| Mn                  | <7.25                | 7.25-8.90     | 8.90-12.18    | 12.18-13.82   | >13.82  | 10.58          | 26-41                | 7.5–23.2              |
| Cu                  | <0.58                | 0.58-0.75     | 0.75-1.10     | 1.10-1.28     | >1.28   | 0.88           | 0.89-1.4             | 2.5-5.1               |
| Zn                  | <1.52                | 1.52-1.82     | 1.82-2.44     | 2.44-2.74     | >2.74   | 1.94           | 1.2-2.4              | 0.59-1.26             |

<sup>†</sup>Manipur, Indo-Myanmar region, Eastern India (Sarangthem et al., 2014).

<sup>‡</sup>Maharashtra and Madhya Pradesh, Central India (Srivastava and Singh, 2008).

respectively, for possible response of N and Mn fertilization (P), besides Cu and Ca/K showing highly negative response to fertilization (N). In low-yielding orchards, these nutrient response dynamics displayed Zn fertilization as having the highest probability of response, followed by Ca and P in decreasing order (P), with Fe and Mn producing highly negative nutrient response (N).

Correspondingly, the yield limited by nutrient shortage or excess in the high-yielding subpopulation of orchards versus low-yielding orchards showed a significantly uniform pattern of expected crop response on account of fertilization (Table 5). In high-yielding orchards, the most frequent cases of nutrient shortage were observed for Zn (37.33%), followed by P (28.00%). Yield was limited by excess of Fe (41.34%) and Cu and Ca (29.34%). In low-yielding subpopulation orchards, the most frequent cases of apparent shortage were Zn (65.71%), Ca (61.90%), P (49.52%), N (44.76%) and K (38.10%). There was an apparent excess of Fe, Mn, Cu, and Mg in 78.10%, 41.90%, and 33.33% of the cases, respectively. Chi-square tests were significant across these nutrient response classes (Supplementary Table 3).

### 3.6 Validation of soil-based DRIS norms

In the high-performance subpopulation, 82.7% of all orchards showed an apparent nutrient shortage correctly as true negative (TN). The remainder was classified as false positive (FP), an imbalanced nutritional status (Figure 4). In the low-performance population, 89.5% of the orchards were TP and the remainder 10.5% were FN. The PPV was 88%, the NPV, 85% and the accuracy, 87%. The probability of correct classification of orchards as balanced was 83%.

### 3.7 ML models

There were nine nutrient concentrations to assess the performance of raw soil test values representing SLAN, three dual ratios (Ca/Mg, Ca/K, K/Mg) representing BCSR, and ten *clr*

variables representing CND. The xgboost regression model showed the highest accuracy using *clr* as nutrient expression (Table 6). SLAN, based on nutrient concentrations, was the second most accurate model. The BCSR that addressed only three cationic dual ratios was inaccurate.

The ML classification models showed that concentration values were less accurate than *clr* to classify data as TN or TP (Table 7).

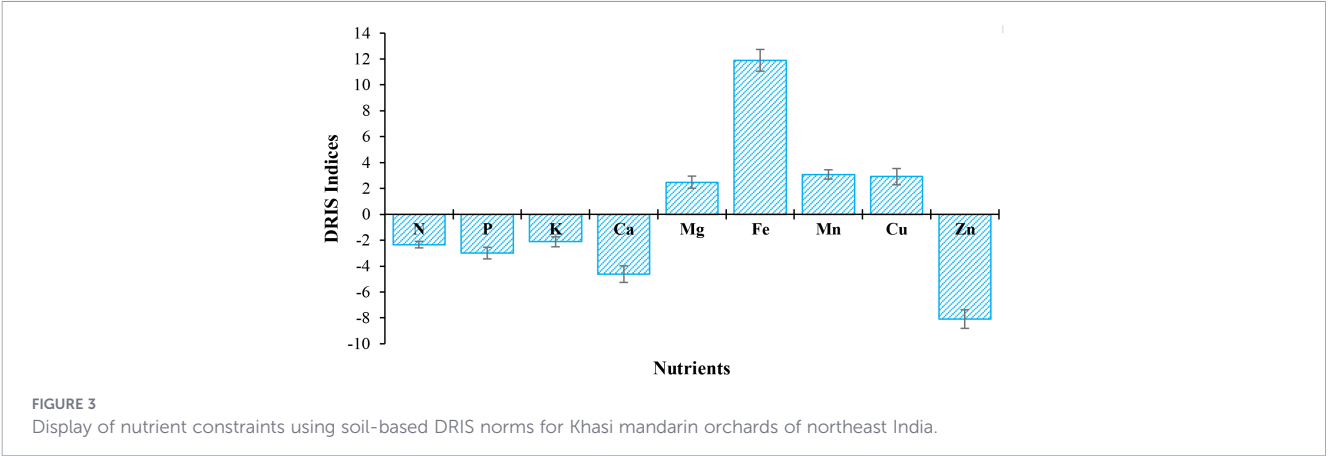
The *clr* nutrient standards were computed for the 92 nutritionally balanced and high-yielding subpopulation set apart by the xgboost classification model (Table 8). The back-transformed *clr* means into the more familiar concentration values were compared to the 'optimum' ranges. N, P, Ca, Mg, Mn, and Cu centroids fell within the concentration range. K, Fe, and Zn centroids were below the lower bounds, indicating distortion at the step of relating concentration values to DRIS indices to derive 'optimum' concentration ranges.

The orchard showing the lowest fruit yield was diagnosed against the closest high-yielding specimen showing 32.17 Mg ha<sup>-1</sup> and Euclidean distance of 0.42. Zn was found to be in relative shortage and Cu in relative excess (Supplementary Figure 1). The Zn CND index was significant at the 0.02 level and Cu excess at the 0.08 level (Supplementary Figure 2), indicating Zn apparent limitation and intensive use of Cu-based fungicides.

## 4 Discussion

### 4.1 Yield cutoff

The citrus orchards were primarily located in the hilly terrains of the eastern Himalayas, presenting a unique agroecological setup requiring crop-specific management practices leading to differential yield potentials (Rymbai et al., 2024). There are several selection methods to set apart a reference subpopulation (Letzsch, 1985; Walworth and Sumner, 1987). In the present study, the critical value approach (R<sup>2</sup>) was adopted as per the statistical procedural



steps of [Cate and Nelson \(1971\)](#) to define a yield cut-off value. Yield cut-off ( $19.23 \text{ Mg ha}^{-1}$ ) was lower compared to  $38 \text{ Mg ha}^{-1}$  for mandarin budded on sour orange [*Citrus aurantium* (L.)] and grown on alkaline calcareous clay loam soils in Tunisia ([Labaieda et al., 2018](#)) and to  $20 \text{ Mgha}^{-1}$  for Nagpur mandarin grafted on rough lemon (*Citrus jambhiri* Lush) and grown on alkaline black clay soils in central India ([Srivastava and Singh, 2008](#)). Mandarin was raised from exclusively seedlings on the Himalayan highly acidic sandy loam-loam soils under rainfed conditions. In other studies, mandarin plants were specific rootstock-scion combinations grown under irrigated conditions. Moreover,

ecotypes, climate, geographical landscape and management practices differed.

## 4.2 Soil test interpretation

### 4.2.1 SLAN

The interpretation of soil test results is commonly conducted using concepts developed circa 80- years ago as SLAN and BCSR ([McLean, 1984](#)). SLAN diagnoses nutrient concentrations in the soil, disregarding distorting effects of nutrient interactions and non-

TABLE 4 Frequency of distribution % expected crop response to fertilization in 'Khasi' mandarin orchards.

| Nutrient                 | Positive (P) | Positive/Null (PZ) | Null (Z) | Negative/Null (NZ) | Negative (N) | Total |
|--------------------------|--------------|--------------------|----------|--------------------|--------------|-------|
| %                        |              |                    |          |                    |              |       |
| High-yield subpopulation |              |                    |          |                    |              |       |
| N                        | 5.33         | 16.00              | 53.33    | 17.33              | 8.00         | 100   |
| P                        | 13.33        | 14.67              | 49.33    | 13.33              | 9.33         | 100   |
| K                        | 9.33         | 16.00              | 49.33    | 10.67              | 14.67        | 100   |
| Ca                       | 12.00        | 13.33              | 45.33    | 14.67              | 14.67        | 100   |
| Mg                       | 12.00        | 9.33               | 56.00    | 17.33              | 5.33         | 100   |
| Fe                       | 4.00         | 17.33              | 37.33    | 30.67              | 10.67        | 100   |
| Mn                       | 6.67         | 12.00              | 58.67    | 13.33              | 9.33         | 100   |
| Cu                       | 10.67        | 10.67              | 49.33    | 10.67              | 18.67        | 100   |
| Zn                       | 24.00        | 13.33              | 38.67    | 17.33              | 6.67         | 100   |
| Low yield subpopulation  |              |                    |          |                    |              |       |
| N                        | 10.48        | 34.29              | 41.90    | 10.48              | 2.86         | 100   |
| P                        | 13.33        | 36.19              | 39.05    | 9.52               | 1.90         | 100   |
| K                        | 5.71         | 32.38              | 46.67    | 14.29              | 0.95         | 100   |
| Ca                       | 23.81        | 38.10              | 30.48    | 5.71               | 1.90         | 100   |
| Mg                       | 1.90         | 11.43              | 53.33    | 20.95              | 12.38        | 100   |
| Fe                       | 1.90         | 9.52               | 10.48    | 30.48              | 47.62        | 100   |
| Mn                       | 2.86         | 12.38              | 42.86    | 20.00              | 21.90        | 100   |
| Cu                       | 3.81         | 15.24              | 39.05    | 32.38              | 9.52         | 100   |
| Zn                       | 36.19        | 29.52              | 23.81    | 9.52               | 0.95         | 100   |

†A clear-cut difference in expected crop response to fertilization is visible, while comparing low-yielding versus high-yielding subpopulation via variation in P-values (4.00-24.00 versus 1.90-36.19) and N-values (5.33-18.67 versus 0.95-47.62).

TABLE 5 Frequency distribution of 'Khasi' mandarin orchards in each nutritional class.

| Soil nutrients              | Relative shortage | No limitation | Relative excess | Total |
|-----------------------------|-------------------|---------------|-----------------|-------|
| %                           |                   |               |                 |       |
| High-yielding subpopulation |                   |               |                 |       |
| N                           | 21.33             | 53.33         | 25.33           | 100   |
| P                           | 28.00             | 49.33         | 22.66           | 100   |
| K                           | 25.33             | 49.33         | 25.34           | 100   |
| Ca                          | 25.33             | 45.33         | 29.34           | 100   |
| Mg                          | 21.33             | 56.00         | 22.66           | 100   |
| Fe                          | 21.33             | 37.33         | 41.34           | 100   |
| Mn                          | 18.67             | 58.67         | 22.66           | 100   |
| Cu                          | 21.34             | 49.33         | 29.34           | 100   |
| Zn                          | 37.33             | 38.67         | 24.00           | 100   |
| Low-yielding subpopulation  |                   |               |                 |       |
| N                           | 44.76             | 41.90         | 13.33           | 100   |
| P                           | 49.52             | 39.05         | 11.43           | 100   |
| K                           | 38.10             | 46.67         | 15.24           | 100   |
| Ca                          | 61.90             | 30.48         | 7.62            | 100   |
| Mg                          | 13.33             | 53.33         | 33.33           | 100   |
| Fe                          | 11.43             | 10.48         | 78.10           | 100   |
| Mn                          | 15.24             | 42.86         | 41.90           | 100   |
| Cu                          | 19.05             | 39.05         | 41.90           | 100   |
| Zn                          | 65.71             | 23.81         | 10.48           | 100   |

normal data distribution. SLAN is generally supported by fertilizer trials. General guidelines are proposed in terms of soil test classes using relative yields as response variable. Compared to absolute yields, relative yields are more closely correlated to soil tests (Pearce et al., 2022). Crop response curves provide yield response patterns to make representative fertilizer recommendations in each soil test class. Alternatively, SLAN can take advantage of closure and higher accuracy of *clr* expressions compared to raw concentrations.

#### 4.2.2 BCSR

BCSR addresses relationships between soil cations only, setting *a priori* ranges for 'ideal' soils and disregarding other interactions and data distribution. BCSR computes dual cationic ratios and the saturation of the cation exchange capacity (CEC) by each cation. In the present study, the Ca/Mg, Ca/K and Mg/K molar ratios were  $2.9 \pm 0.7$ ,  $2.8 \pm 0.7$  and  $1.1 \pm 0.7$ , respectively, much lower than in the

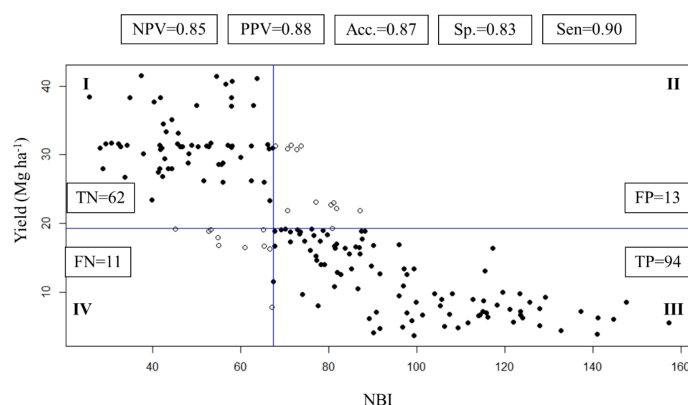


FIGURE 4

Cate-Nelson partitioning into quadrants of the surveyed *Khasi* mandarin orchards. I: adequate nutritional status. II: luxury consumption or suboptimal level. III: nutritional imbalance caused by at least one nutrient. IV: another limiting factors impacting crop performance. NPV, Negative Predictive Value; PPV, Positive Predictive Value; TN, True Negative; FP, False Positive; TP, True Positive; FN, False Negative; Acc., Accuracy; Sp., Specificity; and Sens., Sensitivity.

TABLE 6 Performance of the regression model at yield cutoff of 19.23 metric Mg ha<sup>-1</sup>.

| Nutrient expression | Random forest |                | xgboost |                |
|---------------------|---------------|----------------|---------|----------------|
|                     | RMSE          | R <sup>2</sup> | RMSE    | R <sup>2</sup> |
| Concentrations      | 8.074         | 0.763          | 7.378   | 0.802          |
| Cationic ratios     | 14.333        | 0.254          | 13.495  | 0.339          |
| <i>clr</i>          | 7.025         | 0.821          | 6.913   | 0.827          |

suggested ‘ideal’ soil (Dahnke and Olson, 1990). There is a widespread agreement among soil fertility scientists that BCSR is not a legitimate as a ‘soil balancing’ general approach (Culman et al., 2021). Indeed, crop performance is not impacted across extremely wide variations in dual cationic ratios (Dahnke and Olson, 1990). The K fertilization also depends on clay properties (Breker et al., 2019). The ML model did not support the suggested dual cationic ratios commonly reported in the literature (Dahnke and Olson, 1990). BCSR where CEC is computed as the sum of exchangeable cations and exchangeable acidity (Quaggio et al., 1985) proved to be useful for making lime and gypsum recommendations in tropical soils (Natale et al., 2012).

The lower accuracy of BCSR for soil fertility management is due to the assumption of a single ideal cation ratio across soils and crops. However, there could be the specific roles of Mg, K and Ca on crop yield, soil pH, aggregation, health and biological activity, over-emphasis on ratios over sufficiency level of nutrients, neglecting the actual concentration of all nutrients and undue risk of over-fertilization (Chaganti et al., 2021). Comparing BCSR and SLAN is a false dichotomy when considering the management of soil acidity (Culman et al., 2021).

### 4.3 Buildup and maintenance concept

The rapid build-up and maintenance of soil test levels is a common practice to adjust fertilization to crop performance (Dahnke and Olson, 1990). The P and K fertilizer recommendations should maintain soil test P and K the medium soil fertility classes to avoid soil mining. However, basing nutrient requirements on nutrient balance sheets and nutrient-use efficiency coefficients may lead to overfertilization in tropical soils (Nowaki et al., 2017). The right fertilizer dosage is speculative in the absence of fertilizer trials. Alternatively, soil test results could be integrated into ML models to make more accurate predictions of absolute yields and fertilizer requirements for comparable growing conditions.

### 4.4 DRIS and CND

Concentration means and standard deviation to generate soil fertility classes may return distorted results. Indeed, means may not be additive to 100% or measurement unit after conducting statistical analyses across interacting components (Diaz-Zorita et al., 2002; Weltje, 2006). CND uses centered log ratios to correct distortions caused by interactions and data distribution in raw concentration values. DRIS and CND can integrate all dual ratios computable from *D*-parts compositions. Nutrient ranges have been estimated from relationships between concentration values and DRIS indices as suggested by de Lima Neto et al. (2022).

The log ratio methodology can be useful to test the apparent representativity of current concentration ranges. In the present study, the centroid of the optimum concentration was located below the ‘optimum’ ranges of soil K, Fe and Zn derived statistically from the relationship between DRIS indices and concentration values. This is especially critical for Zn because Zn is recognized globally as the most common nutrient limiting the sustainable production of citrus (Srivastava and Singh, 2004). Indeed, Zn is the third most growth-limiting nutrients for crop production in India (Arunachalam et al., 2013). High-revenue fruit crops like citrus, considered highly susceptible to Zn- deficiency, could be biofortified with Zn to support human health (Noulas et al., 2018). Too high an ‘optimum’ range for Zn may lead to overfertilization in orchards showing low Zn levels. The link between soil testing and biofortification however still remains an unattended issue.

Comparing SLAN, DRIS, and CND showed different diagnoses on nutrient constraints. SLAN diagnosed soil nutrients separately. Nutrient interactions distort concentration-based diagnoses (Bates, 1971). DRIS and CND integrated several dual ratios that reflect nutrient interactions (Kenworthy, 1967). As the number of nutrient ratios considered simultaneously increase, the yield-limiting nutrients become easier to identify reliably (Sumner, 1978).

The DRIS concept experienced several formulations and modifications since its inception (Walworth and Sumner, 1987). We used the original formulation (Beaufils, 1973) and statistically assessed crop response classes and concentration ranges to facilitate interpretation (Wadt, 1996; Maia et al., 2001; Silva et al., 2005; Urano et al., 2006). While several nutrients may be deficient or excessive, soil DRIS indices can focus on the most limiting nutrients (Beaufils and Sumner, 1976). The Nutritional Balance Index average (NBIa) is closely related to crop yield (Souza et al., 2015) and has been reported to have a significant relationship with productivity. The close relationship between NBI and yield in our study was also reported for sweet orange ‘Valencia’ (Mourão Filho and

TABLE 7 Performance of the classification model at yield cutoff of 19.23 metric t Mg ha<sup>-1</sup>.

| Nutrient expression | Random forest    |          | xgboost          |          |
|---------------------|------------------|----------|------------------|----------|
|                     | Area under curve | Accuracy | Area under curve | Accuracy |
| Concentrations      | 0.963            | 0.883    | 0.962            | 0.878    |
| <i>clr</i>          | 0.958            | 0.922    | 0.970            | 0.928    |

TABLE 8 Nutrient standards as *clr* values for the 92 nutritionally balanced and high-yielding subpopulation ( $x_D$  = filing value) set apart by the xgboost classification model.

| <i>clr</i> | Mean   | Standard deviation<br>unitless | Back-transformed <i>clr</i><br>mg kg <sup>-1</sup> | 'Optimum' range† |
|------------|--------|--------------------------------|----------------------------------------------------|------------------|
| N          | 1.427  | 0.120                          | 266.35                                             | 213.04-282.46    |
| P          | -2.402 | 0.184                          | 5.79                                               | 4.71-6.74        |
| K          | 0.978  | 0.210                          | 169.94                                             | 181.95-241.81    |
| Ca         | 1.355  | 0.124                          | 247.76                                             | 185.70-253.20    |
| Mg         | -0.171 | 0.198                          | 53.87                                              | 44.96-61.90      |
| Fe         | -0.805 | 0.308                          | 28.58                                              | 35.01-48.99      |
| Mn         | -1.837 | 0.168                          | 10.18                                              | 8.90-12.18       |
| Cu         | -4.216 | 0.211                          | 0.94                                               | 0.75-1.10        |
| Zn         | -3.986 | 0.237                          | 1.19                                               | 1.82-2.44        |

†Table 3.

Azevedo, 2003). As NBI approaches zero, high yields are expected (Angeles et al., 1990; Silveira et al., 2005). Nevertheless, low-yielding crops may still show adequate NBI levels (Teixeira et al., 2021), indicating that other factors were limiting crop yield. A strong correlation between NBI and yield can be achieved only where DRIS standards are site-specific, taking into consideration the important effects of the edaphic, climatic and stock-scion relationships as in the case for coffee (Nick, 1998) and sweet orange (Rodriguez et al., 1997). The theoretical basis behind the DRIS concept was not previously documented until DRIS ratios were reconfigured as log ratios (Parent and Dafir, 1992) following the theory of Compositional Data Analysis (CoDa) (Aitchison, 1986). In contrast to DRIS variables, log ratios are additive and have Euclidean geometry, allowing two equal-length compositions (defective vs. benchmark) to be compared under otherwise similar conditions.

The *clr* variables are appropriate to conduct statistical analyses on compositions by removing the noise inherent to closed systems caused by closure (Aitchison, 1986). Sources of noise in compositional data are nutrient dilution, interactions and signaling, and non-normal distribution. Using log ratios, CND is a unique means to eliminate the noise. CND has several advantages over DRIS as follows:

- CND is easier to compute
- The filing value as component of the whole instead of dry matter (that is not a component but a basis) as separating 'component' between nutrient shortage and excess
- Log ratios are additive as centered log ratios with Euclidean geometry (Greenacre et al., 2023) while dual ratios in DRIS are projected conceptually onto bisectors, assuming additivity among DRIS functions (Traspadini et al., 2018)
- Non-linear relationships between nutrients (Kenworthy, 1967) distort the projection onto bisectors (the slope of the relationship changes at any one point)
- While, compared to dual ratios, logarithms produce symmetrical ratios, i.e.  $\ln(A/B) = -\ln(B/A)$  allowing to use a single equation to compute DRIS functions (Beverly, 1987), log-transformed dual ratios are additive to log ratios

- Log ratios are close to normal distribution
- Variance ratios used by DRIS to set apart the high-yielding subpopulation do not consider false positive specimens that inject noise into nutrient standards (luxury consumption, suboptimal nutrient level), while CND focuses on true negative specimens set apart by ML classification methods and thus exclude FP specimens to generate nutrient standards
- CND may consider the importance of each dual log ratio (Parent, 2025a; Srivastava et al., 2025), while the *clr* expression set equal them (Greenacre et al., 2023)
- Centered log ratios have Euclidean geometry and are thus additive in the Euclidean space, allowing to compare two equal-length compositions under the *ceteris paribus* assumption.

The Mahalanobis distance computed across centered, amalgamated, isometric or pivot log ratios have chi-square distribution, allowing to test the statistical significance of the difference between a diagnosed composition and the centroid of the high-yielding and nutritionally balanced subpopulation weighted by the variance-covariance matrix; the CND index is the Mahalanobis distance for a single variable, that can be tested statistically as a *t* value (de Lima Neto et al., 2025).

The cutoff value to generate CND standards from ML classification models may be selected statistically as shown in this paper or by stakeholders to fit realistic grower's objectives.

## 4.5 Display of nutrient limitations

There are several ways to avoid over-optimistic multivariate models driven by overfitting (Ezenarro and Schorn-Garcia, 2025). Because the database was relatively small, the stratified cross validation (CV) model was used as an exploratory model after combining the calibration and validation datasets following AUC acceptability. The software averaged model accuracy across ten folds, as commonly reported in the literature, to provide maximum reliability. The software does not provide AUC and accuracy for each fold. In any event, external data from other sources, including experimental data, was still not available to us for the most rigorous validation step. This is an initial step toward the "golden" model for future use in orchard

management. Beyond the simple chemometrics of nutrient relationships in plant tissues (and soils as well) solved in part by ML under the heroic *ceteris paribus* assumption, there is much more knowledge to acquire to control variability across agroecosystems and increase the reliability of general models, such as including climatic variables at proper time scale, detailed soil and landscape properties, and orchard management features like pruning intensity and frequency, cover crops, fruit quality, nutrient dosage and pest control.

Nutrient limitations are easy to display using the DRIS or CND methodologies. In the example used in the present study, CND diagnosed Zn shortage significant at 0.02 level and Cu excess significant at 0.08 level (Supplementary Figure 2). Zn deficiency and Cu excess have different agronomical manifestations. Zn deficiency produces the rosetting of small leaves, frequently coupled with interveinal chlorosis that reduces the photosynthesis, leading to decline in productivity (Srivastava and Singh, 2003a). Cu excess accumulated in the root zone reduces the root-mediated rhizosphere acidification and an excessive generation of reactive oxygen species causing lipid peroxidation that increases cell membrane permeability, leading to reduced water and nutrient uptake and plant biomass development (Chen et al., 2022).

Nutrient deficiencies, whether single or in combination, were reported in citrus orchards worldwide (Srivastava and Singh, 2003b). In 'Nagpur' soils of Central India, concentrations of Zn, P, Fe, N, and K were found to be very low to low while those of Mg, Ca, Cu, and Mn were excessive (Srivastava and Singh, 2008). There were limitations of Ca, Mg, P, Zn, N, and K, and high to very high levels of Cu, Mn, and Fe in 'Khasi' soils of Manipur in the Indo-Myanmar region (Sarangthem et al., 2014). Nevertheless, comparisons between nutrient concentration optima of citrus cultivars are due to rootstock-scion combination, the nature and properties of soil, management practices, agroclimatic conditions, irrigation, the presence or absence of Huanglongbing (citrus greening disease caused by phloem limited pathogen, *Candidatus Liberibacter asiaticus* carried by Asian citrus psyllids as vector), orchard topography, and inter-plant distance amongst others (Srivastava and Singh, 2007, 2008; Sarangthem et al., 2014; Rymbai et al., 2024). Orchards of the Indo-Myanmar regions are highly acidic and slopy in geographical location in most of the cases. Acidic soils show relatively low base saturation and are prone to Ca deficiency (McLaughlin and Wimmer, 1999). In the present study, Ca was not the most limiting nutrient. The statistical constructions relating soil test classes to Zn requirements should be validated agronomically by fertilizer trials to support the results of the diagnosis. Tissue tests can complement soil tests to verify whether tissue Zn is properly balanced to achieve high mandarin yields.

DRIS standards and ranges, concentration ranges and CND standards were assessed at population scale to address site specific limitations under the *ceteris paribus* assumption. The *clr* transformation allowed computing the Euclidean distance between two equal-length compositions, i.e. the diagnosed composition and a compositionally close high-yielding and nutritionally balanced specimen, facilitating its interpretation by growers. At subpopulation scale, statistics across nutrient indices can be computed as a chi-square test to support the validity of conducting individual nutrient index.

## 5 Conclusion

Nutrient diagnostic tools are compared to determine nutrient limitations in 'Khasi' orchards of the Indo-Myanmar region of Himalaya. Interpretation tools were SLAN, BCSR, DRIS, CND, and ML. ML models, particularly *xgboost*, using *clr*-transformed nutrient data, were highly accurate and effectively identified limiting nutrients. Ca ranked the first as discriminating variable between high- versus low- yielding subpopulation followed by Cu, Zn, and Mn.

The Euclidean geometry of *clr* allowed diagnosing site-specific problems by comparing two equal-length compositions (abnormal plant vs. compositionally close normal plants), an important issue where the database is small. However, diagnostic tools alone should be supported by field trials and complemented by fertilizer recommendations. This paper screened several diagnostic tools and suggested using log ratio concepts to conduct nutrient diagnosis and display nutrient limitations to crop performance. Stakeholders generally validate the diagnosis with site-specific recommendations.

Nutrient imbalance in 'Khasi' mandarin orchards warrants field validation through progressive nutrient response studies related to soil test results and site-by-site comparisons to support site-specific actions like precision fertilization and biofortification, useful to stakeholders. However, the lack of external validation as universality test and the absence of an independent dataset still limits the strength of these claims.

## Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

## Author contributions

HR: Conceptualization, Validation, Investigation, Formal analysis, Methodology, Writing – original draft, Data curation, Software. L-EP: Writing – review & editing, Software, Formal analysis. AKS: Visualization, Writing – review & editing. HT: Writing – review & editing. VV: Writing – review & editing. VZ: Writing – review & editing. JM: Data curation, Investigation, Writing – review & editing. TR: Writing – review & editing. SH: Writing – review & editing, Supervision. LB: Writing – review & editing.

## Funding

The author(s) declared that financial support was not received for this work and/or its publication.

## Acknowledgments

The authors would sincerely express gratitude to the Director of the ICAR Research Complex for NEH Region, Umiam, Meghalaya, India, for assistance and support in conducting the research work under the Institute Project.

## Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author AKS declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

## Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

## References

- Acosta, M., Quiñones, A., Munera, S., de Paz, J. M., and Blasco, J. (2023). Rapid prediction of nutrient concentration in citrus leaves using vis-NIR spectroscopy. *Sens.* 23, 6530. doi: 10.3390/s23146530
- Agricultural Statistics (2024). *Government of India, ministry of agriculture & Farmers welfare, department of agriculture* (New Delhi, India: Cooperation & Farmers Welfare, Directorate of Economics & Statistics).
- Aitchison, J. (1986). *The statistical analysis of compositional data* (London, UK: Chapman and Hall).
- Angeles, D. E., Sumner, M. E., and Barbour, N. W. (1990). Preliminary nitrogen, phosphorus, and potassium DRIS norms for pineapple. *HortSci.* 25, 652–655. doi: 10.5897/AJArx12.015
- Arunachalam, P., Pandian, K., Gnanasekaran, P., and Govindarajan, M. (2013). Zinc deficiency in Indian soils with special focus to enrich zinc in peanuts. *Afr. J. Agric. Res.* 8, 6681–6688. doi: 10.5897/AJArx12.015
- Barthakur, I. K. (2018). Soil pH as a phenotype determinant in humans: proposing a scientific hypothesis. *Open J. Soil Sci.* 8, 36–46. doi: 10.4236/ojss.2018.81003
- Bates, T. E. (1971). Factors affecting critical nutrient concentrations in plants and their environment: a review. *Soil Sci.* 112, 116–130. doi: 10.1097/00010694-197108000-00005
- Beaufils, E. R. (1973). Diagnosis and recommendation integrated system (DRIS). *Soil Sci. Bull.* 1, 1–132.
- Beaufils, E. R., and Sumner, M. E. (1976). Application of the DRIS approach for calibrating soil and plant factors in their effects on yield of sugarcane. *Proc. S. Afr. Sugar Technol. Assoc.* 51, 118–124.
- Beverly, R. B. (1987). Modified DRIS method for simplified nutrient diagnosis of 'Valencia' oranges. *J. Plant Nutr.* 10, 1401–1418. doi: 10.1080/01904168709363672
- Black, C. A. (1993). *Soil fertility evaluation and control* (Boca Raton, FL: CRC Press, USA), 114–162.
- Bray, R. H., and Kurtz, L. T. (1945). Determination of total, organic, and available forms of phosphorus in soils. *Soil Sci.* 59, 39–45. doi: 10.1097/00010694-194501000-00006
- Breker, J. S., DeSutter, T., Rakkar, M. K., Chatterjee, A., Sharma, L., and Franzen, D. W. (2019). Potassium requirements for corn in North Dakota: influence of clay mineralogy. *Soil Sci. Soc. Am. J.* 83, 429–436. doi: 10.2136/sssaj2018.10.0376
- Cate, R. B. Jr., and Nelson, L. A. (1971). A simple statistical procedure for partitioning soil test correlation data into two classes. *Soil Sci. Soc. Am. J.* 35, 658–660. doi: 10.2136/sssaj1971.03615995003500040048x
- Chaganti, V. N., Culman, S. W., Herms, C., Sprunger, C. D., Brock, C., Leiva Soto, A., et al. (2021). Base cation saturation ratios, soil health, and yield in organic field crops. *Agron. J.* 113, 41.90–44200. doi: 10.1002/agj2.20785
- Chapman, H. D., and Pratt, P. F. (1978). *Methods of analysis of soils, plants, and waters* (Oakland, CA, USA: University of California, Division of Agriculture Sciences), 82–96.
- Chen, H. H., Chen, X. F., Huang, W. Z., Guo, J., Yang, L. T., and Chen, L. S. (2022). Characterization of copper-induced release of exudates by *Citrus sinensis* roots and their possible roles in copper tolerance. *Chemosphere.* 308, 1363–1348. doi: 10.1016/j.chemosphere.2022.136348
- Chlingaryan, A., Sukkarieh, S., and 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. doi: 10.1016/j.compag.2018.05.012
- Costa, L., Kunwar, S., Ampatzidis, Y., and Albrecht, U. (2022). Determining leaf nutrient concentrations in citrus trees using UAV imagery and machine learning. *Precis. Agric.* 23, 854–875. doi: 10.1007/s11119-021-09864-1
- Culman, S. W., Brock, C., Doohan, D., Jackson-Smith, D., Herms, C., Chaganti, V. N., et al. (2021). Base cation saturation ratios vs. sufficiency level of nutrients: a false dichotomy in practice. *Agron. J.* 113, 5623–5634. doi: 10.1002/agj2.20787
- Dahnke, W. C., and Olson, R. A. (1990). "Soil test correlation, calibration, and recommendation," in *the soil testing and plant analysis, 3rd ed.* Ed. R. L. Westerman (Madison, WI: Soil Science Society of America, Madison), 45–71. doi: 10.2136/sssabookser3.3ed.c4
- de Lima Neto, A. J., de Deus, J. A. L., Rozane, D. E., de Corrêa, M. C. M., Natale, W., Parent, E., et al. (2025). Customized nutrient standards to diagnose nutrient imbalance in fertigated 'Nanica' banana groves. *Horticulturae* 11, 1327. doi: 10.3390/horticulturae11111327
- de Lima Neto, A. J., Natale, W., Rozane, D. E., de Deus, J. A. L., and Rodrigues Filho, V. A. (2022). Establishment of DRIS and CND standards for fertigated 'Prata' banana in the northeast, Brazil. *J. Soil Sci. Plant Nutr.* 22, 765–777. doi: 10.1007/s42729-021-00687-7
- Diaz-Zorita, M., Perfect, E., and Grove, J. H. (2002). Disruptive methods for assessing soil structure. *Soil Till. Res.* 64, 3–22. doi: 10.1016/S0167-1987(01)00254-9
- Ezenarro, J., and Schorn-Garcia, D. (2025). How are chemometric models validated? A systematic review of linear regression models for NIRS data in food analysis. *J. Chemometrics* 39, e70036. doi: 10.1002/cem.70036

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fagro.2026.1719690/full#supplementary-material>

- Ferrández-Cámara, M., Martínez-Nicolás, J. J., Alfosea-Simón, M., Cámara-Zapata, J. M., Melgarejo-Moreno, P., and García-Sánchez, F. (2021). Estimation of diagnosis and recommendation integrated system (DRIS), compositional nutrient diagnosis (CND) and range of normality (RN) norms for mineral diagnosis of almonds trees in Spain. *Horticulturae* 7, 481. doi: 10.3390/horticulturae7110481
- Greenacre, M., Gunsky, E., Bacon-Shone, J., Erb, I., and Quinn, T. (2023). Aitchison's compositional data analysis 40 years on: a reappraisal. *Stat. Sci.* 38, 386–410. doi: 10.48550/arXiv.2201.05197
- Hatem, G., Zeidan, J., Goossens, M., and Moreira, C. (2022). Normality testing methods and the importance of skewness and kurtosis in statistical analysis. *BAU J. Sci. Technol.* 3, 7. doi: 10.54729/ktpe9512
- Horticultural Statistics at a Glance (2024). *Horticulture statistics division, department of agriculture* (New Delhi, India: Cooperation & Farmers Welfare, Ministry of Agriculture & Farmers Welfare, Government of India).
- Huang, H., Hu, C. X., Tan, Q., Hu, X., Sun, X., and Bi, L. (2012). Effects of Fe-EDDHA application on iron chlorosis of citrus trees and comparison of evaluations on nutrient balance with three approaches. *Sci. Hortic.* 146, 137–142. doi: 10.1016/j.scienta.2012.08.015
- Huynh-Thu, V. A., and Geurts, P. (2019). Unsupervised gene network inference with decision trees and random forests. *Methods Mol. Biol.* 1883, 195–215. doi: 10.1007/978-1-4939-8882-2\_8
- Kenworthy, A. L. (1967). "Interpreting the balance of nutrient elements in leaves of fruit trees," in *Plant analysis and fertilizer problems*. Ed. W. Reuther. (Washington, DC: American Institute of Biological Sciences), 28–43.
- Labaieda, M. B., Serra, A. P., and Ben Mimoun, M. (2018). Establishment of nutrients optimal range for nutritional diagnosis of mandarins based on DRIS and CND methods. *Commun. Soil Sci. Pl. Anal.* 49, 2557–2570. doi: 10.1080/00103624.2018.1526944
- Letzsch, W. S. (1985). Computer program for selection of norms for use in the Diagnosis and Recommendation Integrated System (DRIS). *Commun. Soil Sci. Pl. Anal.* 16, 339–347. doi: 10.1080/00103628509367609
- Lindsay, W. L., and Norvell, W. (1978). Development of a DTPA soil test for zinc, iron, manganese, and copper. *Soil Sci. Soc. Am. J.* 42, 421–428. doi: 10.2136/sssaj1978.03615995004200030009x
- Maia, C. E., de Moraes, E. R., and Oliveira, M. D. (2001). Critical level through the reduced normal distribution approach: a new proposal for interpretation of foliar analysis. *Rev. Bras. Eng. Agric. Ambient* 5, 235–238. doi: 10.1590/S1415-43662001000200010
- McLaughlin, S. B., and Wimmer, R. (1999). Calcium physiology and terrestrial ecosystem processes. *New Phytol.* 142, 373–417. doi: 10.1046/j.1469-8137.1999.00420.x
- McLean, E. O. (1984). "Contrasting concepts in soil test interpretation: sufficiency levels of available nutrients versus basic cation saturation ratios," in *Soil testing: correlating and interpreting the analytical results*. Ed. M. Stelly (American Society of Agronomy Special Publication 29, Madison, Wisconsin, USA), 39–54.
- Morais, T., Prado, R., Traspadini, E. I. F., Wadt, P. G. S., Paula, R., and Rocha, A. M. S. (2019). Efficiency of the CL, DRIS and CND methods in assessing the nutritional status of Eucalyptus spp. Rooted cuttings. *Forests* 9, 786. doi: 10.3390/f10090786
- Mourão Filho, F. A. A., and Azevedo, J. C. (2003). DRIS norms for 'Valencia' sweet orange on three rootstocks. *Pesqui. Agropecu. Bras.* 38, 85–93. doi: 10.1590/S0100-204X2003000100012
- Munson, R. D., and Nelson, W. L. (1990). "Principles and practices in plant analysis," in *Soil testing and plant analysis*, 3rd ed. Ed. R. L. Westerman. (Madison, WI: Soil Science Society of America), 359–387.
- Natale, W., Rozane, D. E., Parent, L. E., and Parent, S.É. (2012). "Soil acidity and liming of tropical fruit orchards," in *Soil fertility*. Ed. R. N. Issaka (InTech Open, New York), 173–192. doi: 10.5772/53345
- Nick, J. A. (1998). *DRIS for pruned coffee trees* (Piracicaba, Brazil: USP/ESALQ).
- Noulas, C., Tziouvalakas, M., and Karyotis, T. (2018). Zinc in soils, water and food crops. *J. Trace Elem. Med. Biol.* 49, 252–260. doi: 10.1016/j.jtemb.2018.02.009
- Nowaki, R. H. D., Parent, S.É., Cecílio Filho, A. B., Rozane, D. E., Meneses, N. B., Santos da Silva, J. A., et al. (2017). Phosphorus over-fertilization and nutrient misbalance of irrigated tomato crops in Brazil. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.00825
- Obreza, T., and Morgan, K. (2008). *Nutrition of Florida citrus trees*. 2nd ed (Gainesville, FL: University of Florida, Institute of Food and Agricultural Sciences, Florida Cooperative Extension Service, Soil and Water Science Department).
- Parent, L. E. (2025a). Networking 13 berry minerals to sustain a high yield of firm cranberry fruits. *Horticulturae* 11, 705. doi: 10.3390/horticulturae11060705
- Parent, L. E. (2025b). Compositional and machine learning tools to model plant nutrition: overview and perspectives. *Horticulturae* 11, 161. doi: 10.3390/horticulturae11020161
- Parent, L. E., and Dafir, M. (1992). A theoretical concept of compositional nutrient diagnosis. *J. Am. Soc. Hortic. Sci.* 117, 239–242. doi: 10.21273/JASHS.117.2.239
- Pearce, A. W., Slaton, N. A., Lyons, S. E., Bolster, C. H., Bruulsema, T. W., Grove, J. H., et al. (2022). Defining relative yield for soil test correlation and calibration trials in the fertilizer recommendation support tool. *Soil Sci. Soc. Am. J.* 86, 1338–1353. doi: 10.1002/saj2.20450
- Quaggio, J. A., van Raij, B., and Malavolta, E. (1985). Alternative use of the SMP buffer solution to determine lime requirement of soils. *Commun. Soil Sci. Plant Anal.* 16, 245–260. doi: 10.1080/00103628509367600
- Rodriguez, O., Rojas, E., and Sumner, M. (1997). Valencia orange DRIS norms for Venezuela. *Commun. Soil Sci. Plant Anal.* 28, 1461–1468. doi: 10.1080/00103629709369887
- Rymbai, H., Deshmukh, N. A., Ramesh, T., Verma, V. K., Baiswar, P., Mawlein, J., et al. (2024). DRIS standards for identifying yield-limiting nutrients in Khasi mandarin and their relationship with fruit yield on acidic soil of the eastern Himalayas, India. *J. Plant Nutr.* 47, 1331–1346. doi: 10.1080/01904167.2024.2308193
- Rymbai, H., Talang, H. D., Rymbai, D., Mawlein, J., Devi, M. B., and Assumi, S. R. (2022). Good orchard management practices in Khasi mandarin (Citrus reticulata Blanco). *Agric. J. World* 2, 1–7.
- Sarangthem, I., Sharma, L. D., and Srivastava, A. K. (2014). Development of nutrient diagnostic technique for Khasi mandarin (Citrus reticulata Blanco) grown in Manipur. *J. Indian Soc. Soil Sci.* 62, 118–125.
- Silva, G. D., Neves, J., Alvarez, V., and Leite, F. (2005). Assessment of the universality of DRIS, M-DRIS and CND. *Rev. Bras. Ciênc. Solo* 29, 755–761. doi: 10.1590/S0100-06832005000500011
- Silveira, C. P., Nachtigall, G. R., and Monteiro, F. A. (2005). Testing and validation of methods for the diagnosis and recommendation integrated system for signal grass. *Sci. Agric.* 62, 520–527. doi: 10.1590/S0103-90162005000600001
- Souza, H. A., Rozane, D. E., Amorim, D. A., Dias, M. J. T., Modesto, V. C., and Natale, W. (2015). Assessment of nutritional status of guava seedlings using preliminary DRIS norms and sufficiency ranges. *J. Plant Nutr.* 38, 1611–1618. doi: 10.1080/01904167.2015.1017050
- Srivastava, A. K., and Singh, S. (2003a). Plant and soil diagnostic norms for optimum productivity of Nagpur mandarin (Citrus reticulata Blanco) grown in Central India. *Fertil. News* 48, 47–54.
- Srivastava, A. K., and Singh, S. (2003b). Zinc nutrition, a global concern for sustainable citrus production. *J. Sustain. Agric.* 25, 5–42. doi: 10.1300/J064v25n03\_03
- Srivastava, A. K., and Singh, S. (2004). Soil and plant nutritional constraints contributing to citrus decline in Marathwada region, India. *Commun. Soil Sci. Plant Anal.* 35, 2537–2550. doi: 10.1081/LCSS-200030359
- Srivastava, A. K., and Singh, S. (2005). Diagnosis of nutrient constraints in citrus orchards of humid tropical India. *J. Plant Nutr.* 29, 1061–1076. doi: 10.1080/01904160600689183
- Srivastava, A. K., and Singh, S. (2007). DRIS-based nutrient norms for Nagpur mandarin (Citrus reticulata). *Indian J. Agric. Sci.* 77, 363–365.
- Srivastava, A. K., and Singh, S. (2008). DRIS norms and their field validation in Nagpur mandarin. *J. Plant Nutr.* 31, 1091–1107. doi: 10.1080/01904160802115359
- Srivastava, A. K., Huchche, A. D., Parent, L.-E., Malhotra, S. K., Ziogas, V., and Baishya, L. K. (2025). Nutrient balance of Citrus across the Mandarin Belts of India. *Horticulturae* 11, 254. doi: 10.3390/horticulturae11030254
- Subbiah, B. V., and Asija, G. L. (1956). A rapid procedure for determination of available nitrogen in soils. *Curr. Sci.* 25, 259–260.
- Sumner, M. E. (1978). Interpretation of nutrient ratios in plant tissue. *Commun. Soil Sci. Plant Anal.* 9, 335–345. doi: 10.1080/00103627809366812
- Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science* 246, 1285–1293. doi: 10.1126/science.3287615
- Teixeira, M. B., Donato, S. L. R., Silva, J. A. D., and Donato, P. E. R. (2021). Reference values for nutritional diagnosis of 'Gigante' cactus pear by DRIS and BIK methods. *Cienc. Agrotec.* 45, e010721. doi: 10.1590/1413-7054202145010721
- Traspadini, E. I. F., Prado, R. M., Vaz, G. J., Silva, F. C., Manoni, A. L., Silva, G. P., et al. (2018). *Practical guide for applying the nutritional composition diagnosis (CND) method: example of use in sugarcane cultivation* (Campinas, SP: Embrapa Informática Agropecuária, Ministry of Agriculture, Livestock and Supply).
- Urano, E. O. M., Kurihara, C. H., Maeda, S., Vitorino, A. C. T., Gonçalves, M. C., and Marchetti, M. E. (2006). Soybean nutritional status evaluation. *Pesqui. Agropecu. Bras.* 41, 1421–1428. doi: 10.1590/S1678-3921.pab2006.v41.7298
- USDA (2024). Outlook for global citrus production (U.S. Department of Agriculture, Foreign Agricultural Service). Available online at: <https://citrusindustry.net/2022/02/07/outlook-for-global-citrus-production/>.
- Wadt, P. G. S. (1996). *The methods of mathematical chance and the integrated diagnosis and recommendation system (DRIS) in the nutritional assessment of eucalyptus plantations* (Viçosa, MG: Federal University of Viçosa), 123.
- Walworth, J. L., and Sumner, M. E. (1987). The diagnosis and recommendation integrated system (DRIS). *Adv. Soil Sci.* 6, 149–188. doi: 10.1007/978-1-4612-4682-4\_4
- Weltje, G. J. (2006). "Ternary sandstone composition and provenance: an evaluation of the 'Dickinson model'," in *Compositional data analysis in the Geosciences: from theory to practice*. Eds. A. Buccianti, G. Mateu-Figueras and V. Pawłowsky-Glahn (Geological Society, Special Publication 264, London, UK), 79–99.