Smart Agricultural Technology 10 (2025) 100802

Contents lists available at ScienceDirect

Smart Agricultural Technology

journal homepage: www.journals.elsevier.com/smart-agricultural-technology

Hyperspectral imaging for precision nitrogen management: A comparative exploration of two methodological approaches to estimate optimal nitrogen rate in processing tomato

Vito Aurelio Cerasola [a], Francesco Orsini [a], Giuseppina Pennisi [a,*], Gaia Moretti [a], Stefano Bona [b], Francesco Mirone [b], Jochem Verrelst [c], Katja Berger [d], Giorgio Gianquinto [a]

[a] Department of Agricultural and Food Sciences (DISTAL), Alma Mater Studiorum–University of Bologna, Viale Fanin 44, Bologna 40127, Italy
[b] Department of Agronomy, Food, Natural resources, Animals and Environment (DAFNAE), University of Padova, Viale dell'Università, 16, Legnaro, PD 35020, Italy
[c] Image Processing Laboratory (IPL), Universitat de València, València, Spain
[d] GFZ Helmholtz Centre for Geosciences, Telegrafenberg, Potsdam 14473, Germany

ARTICLE INFO

Keywords:
Precision nitrogen management
Hyperspectral imaging
Machine learning
Nitrogen nutrition index
Nitrogen rate

ABSTRACT

Hyperspectral imaging is widespread in crop nitrogen (N) monitoring for precision agriculture, although approaches that address the agronomical recommendation of the optimal N rate are still lacking. Here, two approaches are explored in defining the optimal N rate to be supplied in fertigated processing tomatoes through hyperspectral imaging. The first one, called the N uptake approach, focuses on the virtual reproduction of the critical N uptake curve through the estimation of both aboveground biomass and crop N uptake. The estimated biomass is used to derive the critical N uptake, and the optimal N rate is computed as the difference between the critical N uptake and the estimated actual N uptake. The second approach focuses on the monitoring of the Nitrogen Nutrition Index (NNI) and biomass. Again, the biomass is used to calculate the critical N uptake, which, when combined with the estimated NNI, resolves the equation to retrieve the actual crop N uptake. A modeling stage was included to estimate the N-related variables from crop canopy reflectance across the full spectrum (400–1000 nm). Canopy reflectance was measured by using an unmanned aerial vehicle at five growth stages of processing tomatoes grown under experimental plot conditions with different N rates. Three nonparametric algorithms were trained, i.e., Gaussian Process Regression (GPR), Support Vector Regression (SVR), and Partial Least Square Regression (PLSR). Multicollinearity of spectral bands was prevented with a principal component analysis, and models were 5-fold cross-validated. Considering the pivotal role of biomass in the selected N rate estimation approaches, two distinct biomass estimation methods were explored. The direct biomass retrieval from spectral data was compared with the indirect biomass retrieval from the remotely sensed LAI applying empirical regressions. PLSR outperformed the other algorithms in estimating N uptake (Relative Root Mean Square Error, RRMSE=21.8 %), while SVR better estimated NNI (RRMSE=10.2 %) and direct biomass (RRMSE=19.4 %). The indirect estimation of biomass outperformed the direct approach when GPR is used (RRMSE 18.2 % vs. 21.4 %), although the influence of soil background at early growth stages determines an unreliable biomass estimation for both methods. The NNI approach outperformed the N uptake approach in estimating the optimal N rate, especially when the biomass is directly retrieved from GPR. The promising estimation performances in N rate estimation (R^2=0.88 and RRMSE=36 %) revealed the effectiveness of hyperspectral imaging in entering the agronomical scheduling of precision N management.

1. Introduction

Nitrogen (N) is among the most important macronutrients for boosting vegetable crop yields, thus playing a pivotal role in ensuring food security for the growing world population [28]. However, the intensification of fertilization has led farmers to often use N fertilizers above the actual crop N requirements, resulting in negative environmental implications [44]. Therefore, precise management of N is

* Corresponding author.
E-mail address: giuseppina.pennisi@unibo.it (G. Pennisi).

https://doi.org/10.1016/j.atech.2025.100802
Received 27 November 2024; Received in revised form 22 January 2025; Accepted 23 January 2025
Available online 27 January 2025

required to minimize losses and optimize fertilization costs. Yet, during the early 1990s, the strategy of "dynamic optimization of N" was proposed [50]. This strategy aims to match the in-season N supply through fertilization with the actual crop N requirements, thus making rapid and non-destructive monitoring of the crop N status imperative. Dynamic N fertilization can reduce the N supply up to 50 % without any effect on cantaloupe yield, even increasing fruits' quality when hand-held chlorophyll meters (e.g., Hydro N-tester) are used as N monitoring devices [13]. Furthermore, SPAD threshold values were developed for guiding N fertilization in processing tomatoes by adopting a linear-plateau model built on relative yield [11]. Similarly, by relating vegetation indices with the Nitrogen Nutrition Index (NNI) and solving the equation for NNI=1, threshold values were also developed for greenhouse tomatoes [31]. However, these approaches only define the optimal timing for supplying N fertilizers, providing limited information on the optimal N rate. In this regard, other authors ventured into developing approaches for estimating the N rate. For instance, Scharf and Lory [37] developed empirical relationships to estimate the economically optimal N rate (EONR) of corn from NDVI measurements. Yet, EONR is affected by annual variability and requires in-season calibration with an N-rich plot [37]. An elegant and physiology-based approach to estimating the optimal N rate was proposed by Houles et al. [18], which is based on the critical N dilution curve. This model relates the aboveground biomass (AGB) to the minimum crop N concentration needed to maximize biomass production. From here, the critical N uptake curve along the AGB is easily derived, and the N rate is calculated as the difference between the critical N uptake and the actual N uptake. In this regard, Cilia et al. [7] computed the N rate following the estimation of NNI (i.e., the ratio between the actual and critical N concentration). Their methodology involved an empirical equation where vegetation indices are used to estimate AGB and crop N concentration, and from which the optimal N rate is subsequently calculated.

Noteworthy, the aforementioned works rely on the use of parametric regression with vegetation indices. Although this is by far the most popular approach since the birth of remote sensing because of their inherent simplicity, they are highly noise-sensitive and they often lack the generic capacity among environmental conditions [48]. As such, nonparametric approaches, mostly referred to as machine learning regression algorithms (MLRAs), emerged as a potentially superior modeling choice, as they can retrieve N-related variables using the full spectrum as input predictors [48]. The use of the full reflectance spectrum monitored by a hyperspectral sensor outperformed the parametric approach based on vegetation indices to monitor crop N content in sugarcane with excellent estimation performances [29]. Although partial least square regression (PLSR) is not strictly considered as a machine learning algorithm because of its linear transformation [48], it was identified as the most popular linear nonparametric algorithm in N retrieval from spectral imaging, as extensively reviewed by Berger et al. [1]. PLSR can provide improvements when compared to vegetation indices in estimating N concentration [16], but nonlinear nonparametric algorithms have gained attention in the last decades as they can capture nonlinear relationships among features [48]. Among these, the Gaussian Process Regression (GPR) is a nonparametric algorithm employing Bayesian principles, which stands out by providing not only an estimate of the variable but also a measure of uncertainty alongside the estimate [2,47]. This makes GPR particularly valuable for tasks demanding reliable predictions despite inherent data variability. In addition to N uptake retrieval [2], GPR is successfully used to retrieve AGB [52] and leaf area index (LAI, [49]). However, the nonparametric algorithms are known to suffer from the Hughes phenomenon, where the multicollinearity of contiguous spectral bands monitored by hyperspectral sensors can penalize the regression performances [48]. Such multicollinearity is typically handled with linear spectral transformations such as principal component analysis (PCA, [35]).

However, the estimation of the optimal N rate utilizing N-related variables like NNI, N uptake, and AGB retrieved from nonparametric algorithms remains unexplored, as extensively reviewed by Li et al. [23]. Therefore, in this study, we aim to explore two distinct workflows to estimate the in-season N rate in a processing tomato (*Solanum lycopersicum* L.) cultivation from unmanned aerial vehicle (UAV) hyperspectral imaging by retrieving the AGB, N uptake, LAI, and NNI through linear and nonlinear nonparametric regression. To avoid any misunderstandings, we refer to N uptake as the total N stored in AGB per unit area (kg N ha^{-1}). The approaches analyzed, called the N uptake approach (Nup-App, [18]) and the NNI approach (NNI-App, [7]), are based on the critical N uptake curve previously developed for processing tomatoes by Tei et al. [42]. Both approaches utilize the AGB to estimate the critical N uptake, which is functional for the N rate computation by difference with the actual N uptake. While in the Nup-App the actual N uptake is directly estimated from nonparametric algorithms, in the NNI-App the N uptake is derived from the NNI which is previously estimated. Furthermore, given the pivotal role of AGB in the N rate estimation, we also explore two additional methods to retrieve it from spectral reflectance: direct AGB retrieval from the reflectance profile, and indirect AGB retrieval, where empirical regression with the remotely sensed LAI is used. Indeed, LAI is among the most frequently retrieved variables from reflectance data. Physic-based retrieval approaches, as radiative transfer models (RTMs), are gaining huge popularity in remote crop monitoring, especially when hybrid workflows with MLRAs are adopted [2]. These approaches require a smaller dataset for validation, representing an excellent opportunity to accelerate the retrieval of crop traits from satellites. However, as AGB is not a state variable of RTMs, its direct retrieval is not possible. Wocher et al. [52] succeeded in the indirect retrieval of the AGB in wheat, barley, and corn by estimating the LAI, specific leaf area (SLA), and the leaf water thickness from the RTMs. However, their methodology did not account for the fruits in the model, while the development of empirical relationships that link the LAI with the AGB could address this issue. Linking the LAI to the AGB is crucial for precise N management when alternative retrieval approaches (e.g., RTMs) are adopted, but their impact on the estimation accuracy of the N rate was never assessed. Although in this work we will not rely on RTMs, we aim to clarify the impact of the indirect AGB retrieval using empirical regression with the LAI retrieved from MLRAs on the estimation of the N rate.

A field experiment with different N rates was carried out where hyperspectral reflectance was monitored from an unmanned aerial vehicle (UAV) to address the following specific objectives: i) comparing different methodologies to calculate the optimal N rates according to the critical N uptake curve (Nup-App vs NNI-App) aided by hyperspectral imaging and machine learning; ii) to investigate whether the indirect estimation of the AGB from the remotely retrieved LAI can find application in N fertilization schemes with sufficient accuracy; iii) to make a simulation of N fertilization with the approaches investigated using the crop parameters (AGB, LAI, N uptake, NNI) get from the validation dataset of MLRAs.

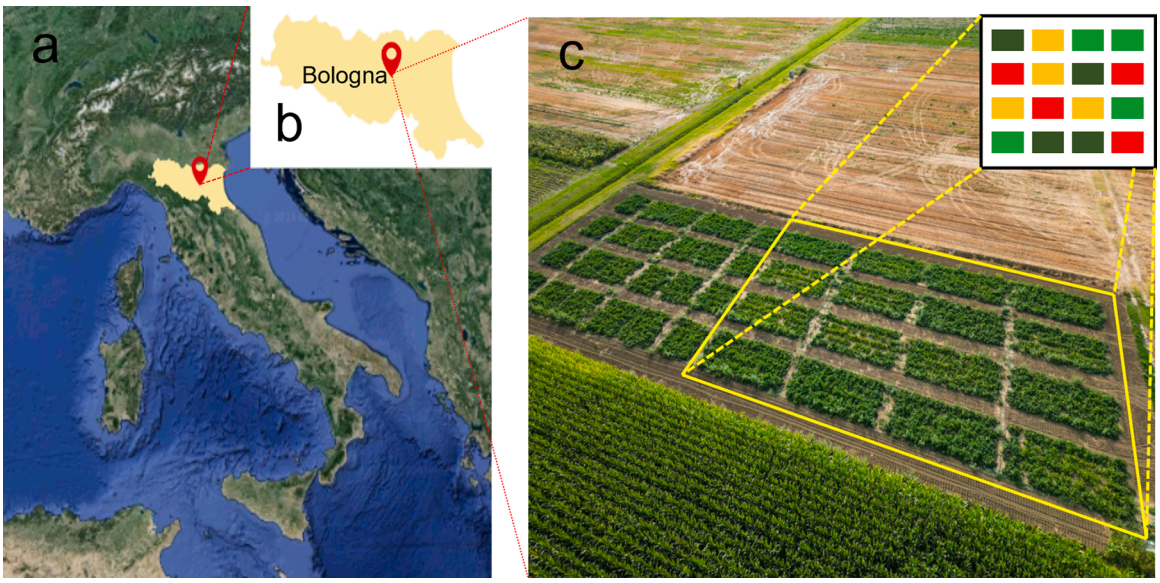

Fig. 1. Location of Bologna in Emilia Romagna region (b) in the Italian peninsula (a). The aerial view of the experimental area in Cadriano (Bologna, Italy) is presented too (c). The yellow area represents the 16 experimental plots that took part in the experiment, with a schematic representation of the full randomization of N treatments (red, yellow, green, and dark green plots represent N0, N20, N60, and N100 treatments, respectively).

Table 1
Sampling dates conducted in the experiment. Day After Transplant (DAT), Growing Degrees Days (GDDs), and BBCH phenological scale are reported.

Day	DAT	GDDs	BBCH
16 June	24	268	17/205/501
22 June	30	377	206/509/606
4 July	42	531	613/704
11 July	49	640	610/706
17 July	55	772	708

2. Materials and methods

2.1. Experimental design and crop management

A plot-size experiment was conducted from May to August 2023 at the experimental farm of the University of Bologna located in Cadriano in Northern Italy (44°55′48'' N°, 11°41′31'' E°, local climate Cfa according to the Köppen classification). Processing tomato plants of cultivar K2206 (ISI Sementi SpA, Fidenza, Italy) were manually transplanted in a single-row cropping system, with a distance between the rows of 1 m, and the distance between the plants in the same row of 0.36 m, resulting in a total of 27,778 plants ha^{-1}. Four nitrogen levels were supplied to the experimental plots sized 9 m x 5 m arranged in a four-times replicated complete randomized design, totaling to 16 experimental plots (Fig. 1). N fertilizer was supplied through fertigation, consisting of varying percentages of the total crop N requirement (0, 20, 60, and 100 %, henceforth N0, N20, N60, and N100). The total crop N requirement was calculated through the balance sheet method [12] which resulted in 180 kg N ha^{-1}. Ammonium nitrate (N = 34 %) was supplied six times through drip fertigation according to the daily crop N uptake rate calculated per three growth stages. Tei et al. [43] reported that in processing tomatoes 6 % of the total N uptake is absorbed by the crop from 0 days after transplant (DAT) to 28 DAT, 78 % from 29 to 77 DAT, and 16 % from 78 to 105 DAT. Therefore, for N100, the daily N uptake in the three growth stages was calculated as 0.39, 2.85, and 1.07 kg N ha^{-1} d^{-1}. Phosphorus was fully supplied at soil tillage at the rate of 80 kg P_2O_5 ha^{-1}, while a total of 150 kg of K_2O ha^{-1} was supplied through fertigation. The chemical-physical characteristics of soil were as follows: sand = 53.7 %, silt = 46.3 %, clay = 0 %; pH = 8.12; total $CaCO_3$ = 99.1 g kg^{-1}; bulk density = 1.34 g cm^{-3}; total organic matter = 1.95 %; total nitrogen = 1.41 g kg^{-1}; the wilting point and field capacity were 6.48 % v:v and 21.9 % v:v, respectively. Crop water requirements were satisfied with drip irrigation adopting the water balance sheet method that maintains soil moisture at the field capacity, as reported by Cerasola et al. [3].

2.2. UAV and spectral data acquisition

Five data collection campaigns were organized at five growth stages along the growing season, where both spectral imagery and ground data (LAI, AGB, NNI, and N uptake) were collected. BBCH growth stages are detailed in Table 1, also reporting the cumulative Growing Degree Days (GDDs) that were calculated using 10 °C and 30 °C as basal and maximum temperatures, respectively.

Spectral images were acquired using an Unmanned Aerial Vehicle (UAV) DJI Matrice 600 (DJI, Shenzhen, China) equipped with the Headwall Nano-Hyperspec® sensor (Headwall Photonics Inc., Bolton, MA, USA) as an airborne imaging sensor. Flights were performed between 11.00 and 15.00 at a height of 40 m above the ground. The sensor was mounted on the stabilized gimbal Ronin MX Gimbal system (DJI, Shenzhen, China) to minimize the movements caused by the UAV. The camera works as a push-broom sensor, mounting a 12 mm lens and detecting the incoming radiation in the visible and near-infrared range (VNIR) from 400 to 1000 nm in 273 bands, with a Full Width at Half Maximum (FWHM) of 6.6 nm and a sampling interval of 2.2 nm, and with a ground sampling distance (GSD) of about 4 cm. Dark and white references were acquired before each flight, acquiring the radiation with the cap on the lens (dark) and with a reference Spectralon® white panel (Labsphere, Inc., North Sutton, NH) placed in the field of view of the sensor. Reflectance calibration tarps (Group 8 Technology, Provo, UT, USA) were placed on the ground to allow the radiometric correction.

2.3. Ground data measurements

At each sampling date, three plants per experimental plot were randomly selected for the destructive monitoring of the AGB, LAI, and N uptake. First, plants were collected, excluding roots, and leaves were separated. Leaf area was measured with a leaf area meter (LI-3000; LI-COR, Lincoln, Nebraska, USA) and LAI (m^2 m^{-2}) was then calculated. Fresh plant biomass was, thereafter, measured with a precision scale, and then each plant was stored in an oven at 60 °C until it reached a constant weight for the measurement of dry plant biomass. Dry AGB was then calculated as t ha^{-1}. Entire dried plants were then crushed into small pieces, maintaining all the organs using a mill. Then, a sub-sample was selected and further ball milled into powder. Total N concentration (N %) was determined with a CHN elemental analyzer (Thermo Fisher, mod. EA 1110, Bremen, Germany) as further detailed in Marcolini et al. [27]. Then, aboveground N uptake (kg N ha^{-1}) was calculated by multiplying the N concentration per the dry mass per unit area (kg ha^{-1}) as obtained by the field campaign. At the end of the growing season which occurred in August, two plants per plot were harvested to determine the total fruit yield (t ha^{-1}).

2.4. Statistical analysis

2.4.1. Data preprocessing

Spectral images were preprocessed on Headwall's SpectralView software as described in the sensor manual. Before each flight, the dark and white references were measured to obtain the gain and offset. These have been used for radiometric correction, therefore, the digital numbers were converted into radiance data by the software. Then, the reflectance was calculated using the white calibration tarp placed on the soil. Finally, orthorectification was obtained, and the overlap with the area of interest was checked with QGIS software. Reflectance data were extracted from spectral images in Google Earth Engine, where the Region of Interest was a polygon that overlapped the experimental plot, excluding the external borders to mitigate the border effect. A total of 78 spectra profiles were extracted from vegetation, and 8 bare soil spectra were included to account for soil variability, resulting in a dataset of 86 observations. Spectra were smoothed through the Savitzky-Golay filter [36] to remove the high-frequency noise, implemented in R studio (version 4.2.0) with package "signal", adopting a window width of 21 nm with a quadratic polynomial order.

2.4.2. Dynamic N fertilization based on UAV spectral data: two approaches

An ideal Decision Support System (DSS) should recommend N fertilizer supply only when it is demanded by the crop, minimizing any errors related to the precise timing of the supply. Additionally, the DSS should provide dynamic recommendations for the optimal N rate, considering the current N status of the crop. Also, it should recommend a reliable N rate across the entire growing season, especially during the growth stages that are critical for N fertilization.

Two approaches for optimal N management based on spectral data are conceptualized here (Fig. 2). The first approach is the N uptake approach (Nup-App, [18]). It consists of the virtual reproduction of the critical N uptake curve by plotting AGB and crop N uptake. If the plotted observation is below the critical curve, then the crop is classified as "To be fertilized" and the N rate is easily calculated as reported in Eq. (1).

$$Optimal\ N\ rate\ \left[kg\ ha^{-1}\right] = N\ uptake_c - N\ uptake_{est} \quad (1)$$

Where the N $uptake_c$ is the critical N uptake calculated using the estimated AGB according to Tei et al. [42], and the N $uptake_{est}$ is the actual N uptake estimated by the regression algorithms that will be subsequently detailed. N $uptake_c$ curve for processing tomato has been developed by Tei et al. [42] in a four-year experiment in Italy, and the equation is reported in Eq. (2):

$$Critical\ N\ uptake\ \left[kg\ ha^{-1}\right] = \begin{cases} 45.3\ AGB, & AGB < 1.2 \\ 45.3\ AGB^{0.673}, & AGB \geq 1.2 \end{cases} \quad (2)$$

However, a threshold needs to be defined to trigger the N fertilization and define the optimal timing of the N fertilizer supply. NNI is commonly used as a rapid way to assess the crop N status, and it is calculated as the ratio of the actual crop N concentration above the critical N concentration, thus indicating an optimal N status around NNI=1.

$$NNI = \frac{N\%_{ac}}{N\%_c} = \frac{N\ uptake}{N\ uptake_c} \quad (3)$$

Using the N uptake instead of the N concentration for the computation of the NNI yields equivalent outcomes [23]. In the context of this study, NNI<0.9 was considered a rational threshold to trigger fertilization [7,17,56]. Therefore, in the framework of the Nup-App the crop is considered to be in N starvation when the plotted observation is below 90 % of the critical N uptake.

The second approach tested is the NNI approach (NNI-App, [7]). It is based on the monitoring of NNI and AGB. First, the NNI is used to assess whether the crop is in N starvation (if NNI<0.9). Then, the critical N uptake is calculated from the estimated AGB (Eq. (2)), and the actual N uptake is easily derived from the NNI estimated and from the critical N uptake (Eq. (3)). Finally, the optimal N rate is computed (Eq. (1)).

Therefore, the two approaches were compared in a simulation of the N fertilization using the validation dataset of the MLRAs (see Section 2.4.5 for details on the validation procedure). NNI-App and Nup-App are primarily used to classify the experimental plots as "To be fertilized" or "To not be fertilized". In this classification, the True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN), were then counted. Here, the experimental plots that are misclassified as FP and FN were identified, and the error (in percentage on the total) was calculated. Then, the N rate is calculated (Eq. (1)) only for the TP and FP cases, to better simulate the performances of the DSS under field conditions. The simulation was performed both for the full validation dataset (including all the N treatments), and also on a restricted validation dataset obtained by merging the N treatments that received more

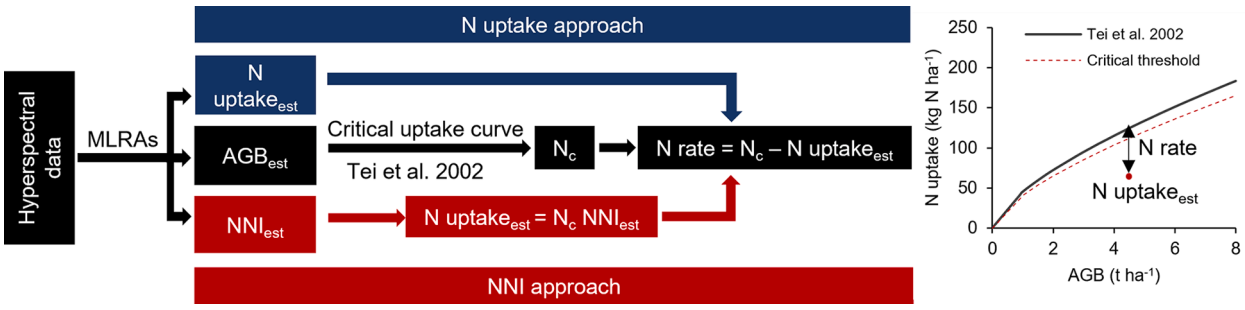

Fig. 2. Schematic representation of the NNI approach and the N uptake approach. The critical N uptake curve developed by Tei et al. [42] is represented as well. The red dashed line in the critical N curve represents the threshold to trigger the N fertilization, which is equal to the 90 % of the critical N uptake curve.

N fertilizers (N60 and N100). Indeed, as these N treatments received 108 and 180 kg N ha^{-1} respectively, they better represent the fertilization level typically supplied by local farmers in processing tomato cultivation in North Italy. Therefore, the data resulting from the merging of the N60 and N100 will be named as Higher N rate, as opposed to the merged N0 and N20 (Lower N rates).

2.4.3. Nonparametric regression algorithms

The extensive comparison of multiple MLRAs is behind the scope of this work, therefore we focus only on a limited number of well-established algorithms. Three nonparametric regression algorithms were selected to estimate N uptake, LAI, AGB, and NNI. For a detailed mathematical description of the algorithms, readers are directed to the following specific literature for an in-depth exploration. The selected algorithms are: (1) Partial Least Squares Regression (PLSR, [53]), (2) Gaussian Process Regression (GPR, [33]), and (3) Support Vector Regression (SVR, [39]).

Briefly, PLSR is a widely utilized linear multivariate method that tackles multicollinearity challenges among input variables (bands). Examining different MLRAs in spectral imaging for crop N retrieval, Berger et al. [1] found PLSR to be the most popular, which explains the choice to include this model in our evaluations. PLSR transforms all spectral bands into a reduced set of latent variables, maximizing the covariance with the dependent variable. Then, latent variables are used to build a multivariate linear model to estimate the dependent variable. However, GPR found a huge interest in the scientific community in the last decades, as it provides an estimate of the uncertainty alongside the retrieved variable [1,47]. This feature, and the adaptability of GPR to a small dataset (as that in the present study), make this algorithm among the most interesting candidates that will play a major role in vegetation monitoring with hyperspectral imaging [1], supporting our choice to include it in this work. GPR is a kernel-based and nonparametric Bayesian MLRA that excels in modeling complex relationships in spectral imaging retrieval issues [48]. GPR can capture intricate patterns between the input and output variables by utilizing a theoretically infinite number of parameters without any assumption on the specific functional form of such relationships. SVR is another kernel-based MLRA frequently adopted in spectral imaging retrieval tasks. Here, the kernel function is used to map the outputs in a continuous multidimensional space, and allow for the individuation of the hyperplane that best fits the output data in such a space. Strictly speaking, only GPR and SVR qualify as MLRAs, as they apply nonlinear transformations [48]. Conversely, as PLSR applies linear transformations, it does not strictly belong to the MLRA category. However, for simplicity and consistency in comparing all algorithms, we adopted the MLRA term to refer to all of them in the remainder of this work.

2.4.4. Software and modeling approach and statistical analysis

The AGB, LAI, N uptake, and NNI monitored in the field experiment were analyzed with the analysis of variance (ANOVA) carried out in R studio (version 4.4.2, packages *car* and *emmeans*), and the means were separated using the post hoc Tukey test. The assumption of homoscedasticity and normality were checked using Levene's test and Shapiro-Wilk's test, respectively. The development of MLRAs was implemented within the Automated Radiative Transfer Models Operator (ARTMO) software running within MATLAB (https://artmotoolbox.com/). ARTMO was born as an intuitive toolbox for facilitating the retrieval of biophysical variables through radiative transfer models (RTMs) inversion [45]. It was expanded with several retrieval toolboxes, such as the MLRA toolbox that can be run even in combination with different dimensionality reduction algorithms [34]. The performances of MLRAs are affected by the choice of various hyperparameters made before the models' training stage. Hyperparameters are algorithm-specific and may include the selection of the number of latent variables or the typology of the kernel function, to name a few. Hyperparameter optimization is crucial for minimizing prediction error and preventing overfitting, ensuring the model's accuracy and generalizability. Here, they are automatically optimized by the ARTMO toolbox. When numerous contiguous bands are entered as input data, however, collinearity hampers the development of fast and accurate MLRAs, and methods to reduce the dimensionality are imperative [35]. PCA is adopted in this study to reduce the huge number of wavelengths into a smaller number of uncorrelated principal components (PCs). PCs were then used as input predictors for the GPR and SVR algorithms. The rule adopted for selecting the number of PCs accounted for two criteria: i) including at least 99.95 % of the cumulative variance explained by the PCs [32], and ii) choosing the minimum number of components that minimize the prediction error.

Accordingly, we tested the variable estimation accuracy as a function of the total number of PCs from 0 (all bands are used) to 20. Considering the pivotal role of AGB in the computation of the optimal N rate, two methods for biomass estimation are explored, including a direct (D-AGB) and indirect (Ind-AGB) approach. The direct approach consists of estimating the AGB directly from the implementation of MLRAs, and the same approach is applied for N uptake, LAI, and NNI. Conversely, the indirect approach estimates the AGB from the LAI, which is previously retrieved from an MLRA. Empirical linear regression models between LAI and AGB were developed from experimental data obtained in the present study and from a similar experiment held in the same environmental condition in 2022, whose details are described in Cerasola et al. [4]. The Partial F test, implemented in Rstudio (version 4.4.2) using the aov() function, was employed to determine whether the effect of N treatments can be omitted in the regression approach (reduced model, AGB=f(LAI)), or if it should be considered, resulting in a more complex full model AGB=f(LAI+N treatment). LAI and AGB are linearly related, but the slope and intercept of these relationships are known to change across the growing season [18]. Therefore, the slope and intercept of the linear relationship between LAI and AGB are modeled as a function of GDDs as similarly reported in Cerasola et al. [5] and Houles et al. [18]. Subsequently, the estimated slope and intercept are used to infer the AGB from the LAI retrieved by an MLRA.

2.4.5. K-fold cross validation and performance assessment

The validation of the models was based on a k-fold cross-validation technique to prevent model overfitting [40]. Briefly, the dataset was randomly split into k different subset of equal size. Therefore, the k-1 subsets were selected as training datasets, and the remaining one was used as a validation set. This process is iteratively repeated k times until each subset is used both for training and for validation only once. In this study, $k = 5$ has been chosen. Goodness-of-fit validation statistics are averaged for the five validation subsets. In particular, model performances are evaluated by computing the following validation indices, namely coefficient of determination (R^2), root mean square error (RMSE), relative RMSE (RRMSE), and Lin's concordance correlation coefficient (p_c) [22], as reported in the following equations:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2}$$

$$RRMSE = \frac{RMSE}{\overline{X}}$$

$$p_c = \frac{2\sigma_{x,y}}{\sigma_x{}^2 + \sigma_y{}^2 + (\overline{X} - \overline{Y})^2}$$

Where x_i and y_i are respectively the observed and estimated variables, $\overline{X}$ and $\overline{Y}$ are respectively the average of the observed and estimated variables, $\sigma_{x,y}$ the covariance between the observed and estimated variables, and $\sigma_x{}^2$ and $\sigma_y{}^2$ are the variance of observed and estimated variables, respectively. The R^2 offers insights into the strength of the relationship between two sets of data, even if they are not perfectly aligned on the 1:1-line. To identify this misalignment, Lin's concordance coefficient

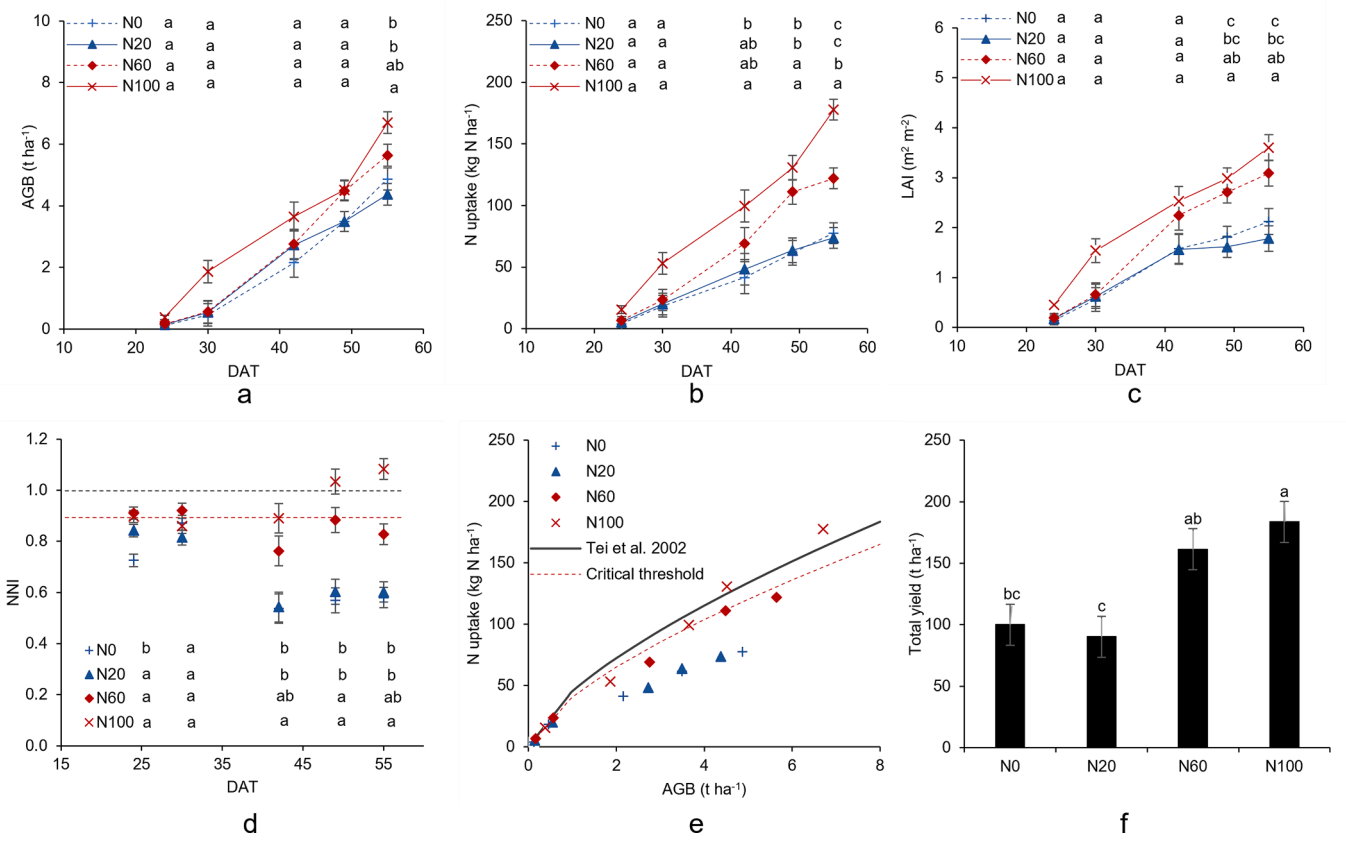

Fig. 3. Crop traits (AGB, N uptake, LAI, NNI, and final total yield) monitored during the experiment. Different letters indicate significant differences at ANOVA analysis for $p < 0.05$ (means separated with Tukey's HSD test), while vertical bars represent the standard error. The N treatments were plotted in the critical N uptake curve [42]. The dashed red lines in (d) and (e) represent the threshold value of NNI (d) and N uptake (e) to trigger the fertilization.

(p_c) specifically quantifies the displacement of the estimated value from the 1:1-line in a scatterplot. A higher p_c indicates greater concordance, reflecting the extent to which estimated values align with observed values.

3. Results and discussion

3.1. Crop traits analysis

The analysis of variance revealed a significant difference among N treatments for different crop traits along the growing season (Fig. 3). While differences in AGB were observed only at the fifth growth stage monitored (Fig. 3a), the LAI was significantly different among treatments also in the fourth growth stage, with higher values observed in N100 and N60 for both parameters (Fig. 3c). The N uptake was not significantly different in the first two sampling dates, but it increased according to the N rate supplied starting from the third sampling (Fig. 3b). The NNI positively reflected the N treatments, with a better N nutrition observed for N100 which remained always above, or close to, the critical threshold of the NNI set at 0.9 as defined in the methodology section (Fig. 3d, [7,56]). Conversely, N0 remained always well below the critical threshold of NNI (Fig. 3d). The treatments were plotted in the critical N uptake curve developed by Tei et al. [42], confirming the need for supplemental N fertilizer for N0 and N20, as they always stayed below the threshold curve, while N60 placed close to the critical threshold (Fig. 3e). This differences in the N status among treatments reflected also in a different total yield, with higher values monitored for N60 and N100 (Fig. 3f).

3.2. MLRAs selection

The cumulative variance reached 99.95 % yet at 6 components for all the estimated variables (Supplementary material, S1), and the number of PCs chosen for each MLRA is reported in Fig. 4 alongside the performance metrics. Although all the MLRA excelled in the estimation of the respective variables (Fig. 4), the best model was selected based on RRMSE, and p_c. These criteria led us to choose $SVR_{D\text{-}AGB}$, PLSR, GPR, and SVR as the best models for the estimation of AGB, N uptake, LAI, and NNI, respectively (Fig. 4). Similar findings were also reported by Verrelst et al. [49], Berger et al. [2], and Liu et al. [25].

Although GPR can overcome the saturation effect at higher LAI values, as remarked by Verrelst et al. [49], LAI retrieval is often affected by the saturation phenomenon at values >3 [55]. However, in this study, no saturation was experienced (Fig. 4), and LAI values ranged up to values below 4. A follow-up analysis would be required to assess the behavior of GPR in LAI estimation at late growth stage.

3.3. AGB retrieval with the direct and indirect approach

While the direct and indirect retrieval of AGB has been widely explored in the literature [9,25,41,52], fewer studies were conducted to assess the impact of the retrieval approach on the estimation performances [20]. Empirical relationships between LAI and AGB of tomatoes were developed at each growth stage in the present study (Supplementary material S2). The partial F test was used to investigate whether N treatments could affect such relationships at the different phenological stages. The full model, where the AGB was estimated by a multivariate linear model of LAI and N treatment (AGB=*f*(LAI+N treatment)), was compared with the less complex reduced model, where N treatment was not included as a predictive variable. N treatment exhibited a significant

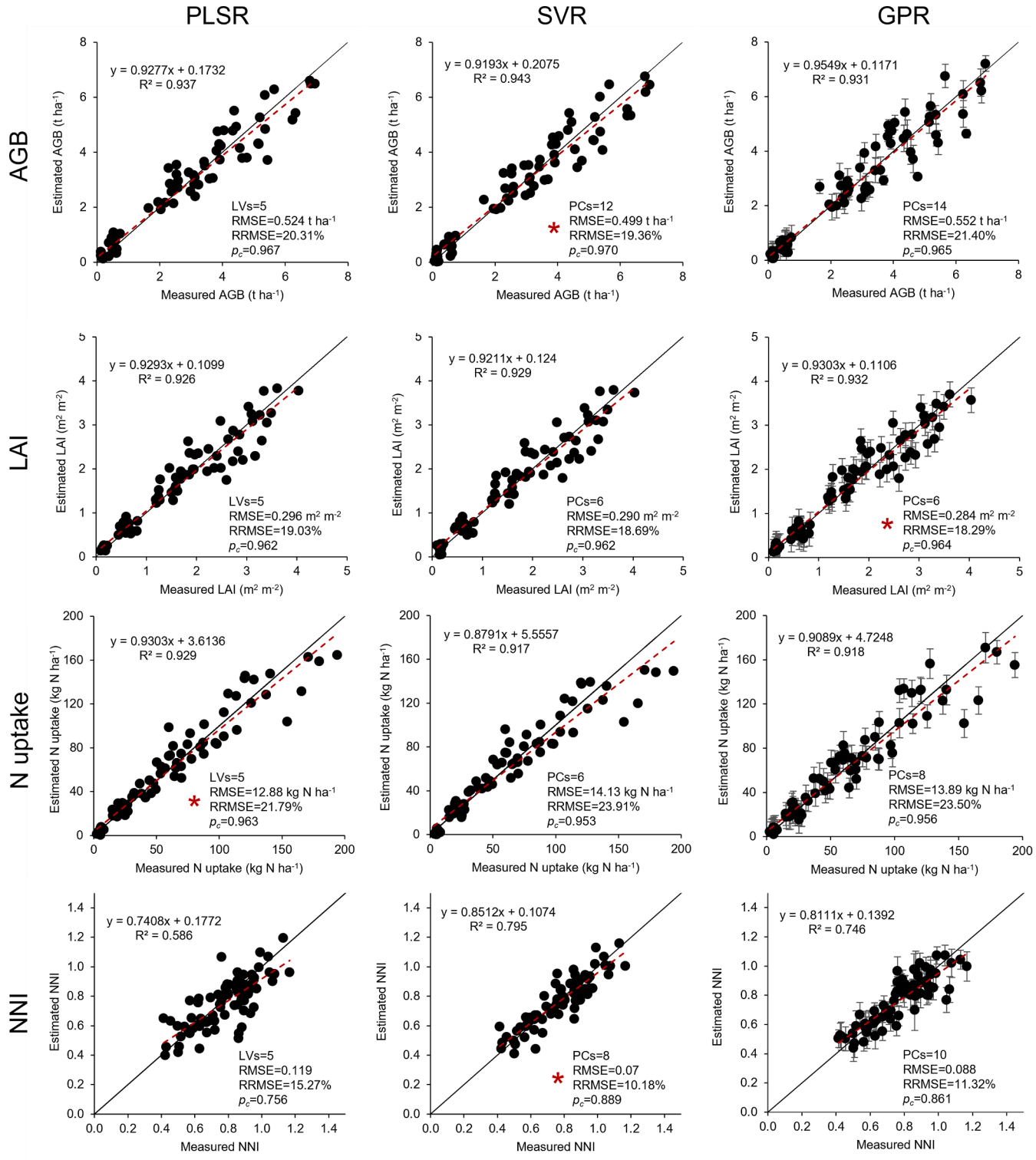

Fig. 4. Measured vs estimated scatterplot of the validation dataset along the 1:1-line (black line) for different crop parameters. Red asterisks indicate the best algorithm for a given crop variable, while the red dashed line represents the fitted line. PCs indicate the number of principal components included in the model, while LVs indicate the number of latent variables included in the PLSR model.

effect only at 531 and at 640 GDDs (Supplementary material, S3), meaning that the more complex full model should be preferred in these sampling dates. Therefore, the treatments that received a lower N rate (N0+N20) were pooled together, and the same was done for the higher N rates (N60+N100). According to the Partial F test, N60 and N100 can be merged in the higher N treatments at both growing stages, while N0 and N20 can be merged in lower N treatments only at 640 GDDs (data not shown). Therefore, at 640 GDDs the resulting combined model (AGB=*f*(LAI+ lower N, higher N)) was tested in comparison with the reduced model, which was chosen as it is less complex (see Supplementary material, S4).

Again, at 531 GDDs the lower N rate treatment cannot be merged

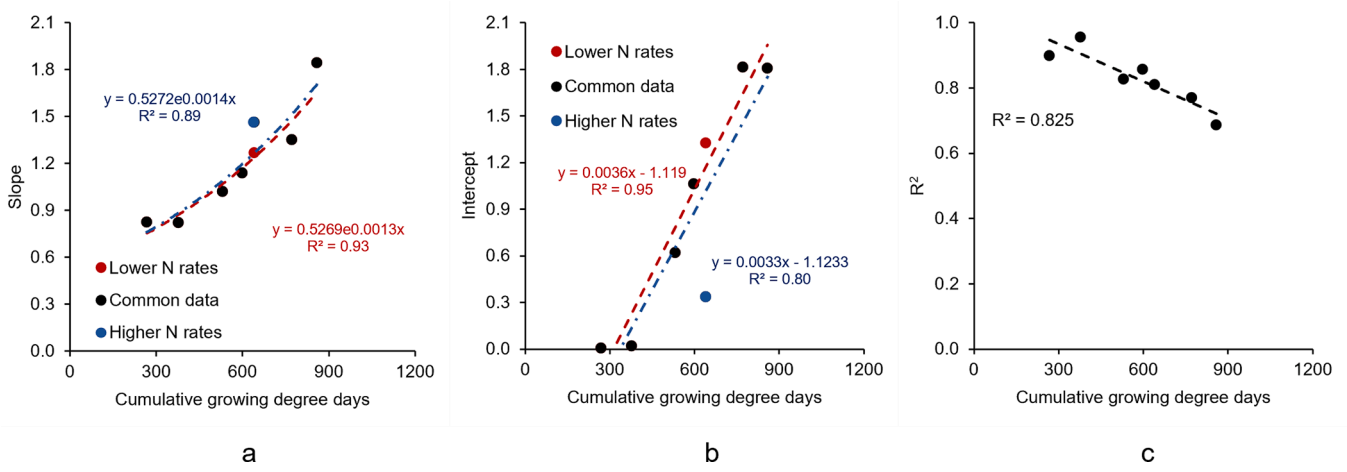

Fig. 5. Empirical models developed to estimate the slope (a) and intercept (b) of the linear relationship between LAI and AGB as a function of growing degree days. Lower N rates represent the combination of N0+N20 experimental treatments, while higher N rates is the combination of N60 and N100 treatments. The pattern of R^2 is reported as well (c).

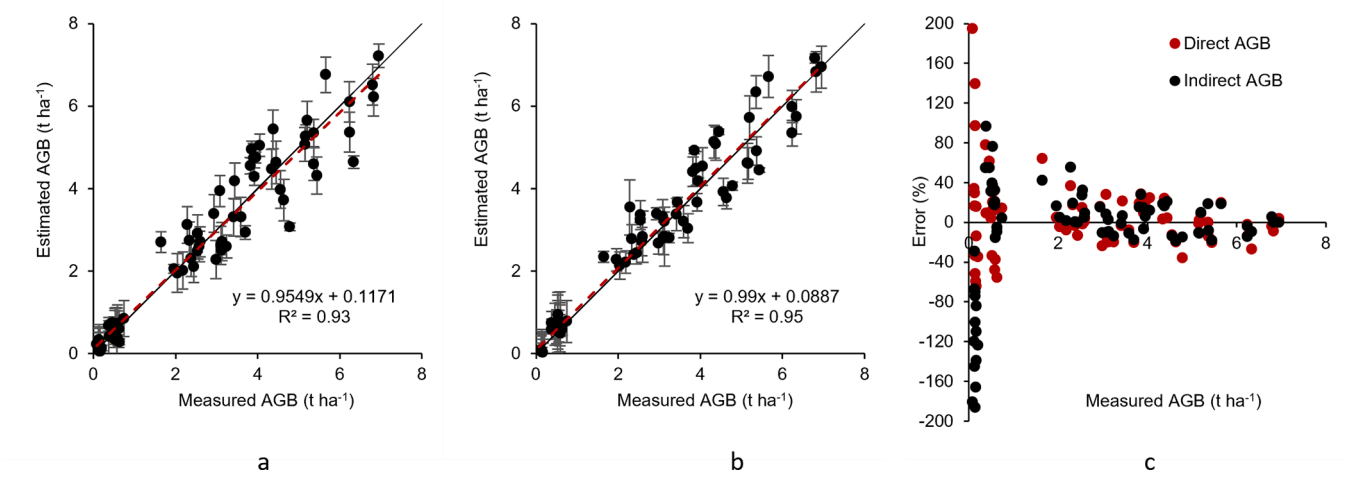

Fig. 6. Validation set of $GPR_{D\text{-}AGB}$ (a) and $GPR_{Ind\text{-}AGB}$ (b) retrieval models. The vertical bars represent the standard deviation estimated by the GPR model. The percentage error is reported as a function of the measured AGB for both retrieval approaches (c).

Table 2
Performances of Indirect AGB retrieval (Ind-AGB). The percentage difference as compared with the direct estimation approach (D-AGB) is reported.

Approach	Algorithm	RMSE	RRMSE	NRMSE	R^2	p_c
Indirect AGB	$PLSR_{Ind\text{-}AGB}$	0.498	19.3	7.24	0.946	0.972
	$SVR_{Ind\text{-}AGB}$	0.495	19.2	7.20	0.946	0.972
	$GPR_{Ind\text{-}AGB}$	**0.472**	**18.2**	**6.87**	**0.951**	**0.974**
Δ% in relation to direct retrieval	$PLSR_{Ind\text{-}AGB}$	−4.91	−4.92	−4.88	+0.96	+0.51
	$SVR_{Ind\text{-}AGB}$	−0.91	−1.07	−0.863	+0.31	+0.20
	$GPR_{Ind\text{-}AGB}$	−14.61	−14.52	−14.45	+2.25	+0.93

(data not shown), as a significant effect of the N0 and N20 treatments exists in the linear regression between AGB and LAI. Despite these results and our awareness of the limits of a model as such, we decided to adopt the reduced model at 531 GDDs to facilitate the subsequent modeling approach.

As processing tomato is a fertigated crop, N fertilizer can be supplied at any moment of the growing season. Thus, to allow for N recommendation along the entire growing season, the slopes and the intercepts of each LAI-AGB model were modeled with the cumulative GDDs. This procedure generated different models for lower N rates and higher N rates conditions. The analysis revealed that slopes and intercept varied along GDDs according to an exponential and linear function, respectively (Fig. 5a and 5b), confirming that the empirical relationships between LAI and AGB are growth-stage specific [18].

The slope models show a similar trend among N conditions (Fig. 5a), but the main differences are observed for the intercept model. In particular, a downward translation of the fitted intercept is observed for higher N treatment, indicating that the AGB is lower in comparison with plants grown under lower N rates at a given leaf area (Fig. 5b). Such a difference can be explained by a varying specific leaf area (SLA, leaf area per unit of mass) among different N treatments, which is an indicator of leaf thickness. Severely N-stressed plants may exhibit thicker leaves and, therefore, a higher mass per unit area, as already observed in tomatoes [38] and in pepper plants [8]. Furthermore, the coefficient of determination of relationships between LAI and AGB was reduced by the increase of GDDs (Fig. 5c), probably because they are gradually noised by

the fruit's development. The equations derived from Fig. 5a and 5b were used to infer the linear models that relate the LAI to AGB. In turn, they were used to indirectly estimate the AGB from the LAI retrieved by the GPR model. Although $SVR_{D\text{-}AGB}$ proved to be the best model for the direct AGB estimation, the indirect AGB retrieval adopted by the GPR model enabled a more accurate estimation of LAI (Fig. 4). For a fair comparison, the $GPR_{Ind\text{-}AGB}$ model was compared with the respective $GPR_{D\text{-}AGB}$ (Fig. 6). The indirect estimation of AGB outperformed the direct AGB estimation approach (Table 2). More specifically, the RRMSE in the $GPR_{Ind\text{-}AGB}$ reduced by about −14 % as opposed to $GPR_{D\text{-}AGB}$. (Table 2). Furthermore, the p_c increased in the indirect approach, indicating a greater agreement between the measured and estimated values [22]. However, it is noteworthy that a potential decline in accuracy may occur throughout the growing season (Fig. 5c).

The estimation of AGB is of utmost importance in crop monitoring and does not rely only on N management but also plays an important role in predicting crop yield [25]. However, the direct estimation of AGB utilizing MLRAs from spectral reflectance requires ground data for the model's training stage. Conversely, LAI is by far the most retrieved biophysical crop trait through RTM inversion, where fewer ground data are needed for the validation stage. Further details related to RTM-based retrieval methods are reported in Verrelst et al. [48]. The notable relationship between LAI and AGB [18] allowed for the indirect estimation of the biomass. Although the indirect estimation of AGB from the retrieved LAI is not a new concept [18], to the best of our knowledge, only Kayad et al. [20] compared the estimation performances of the direct versus indirect approach to AGB estimation in maize. The authors reported that the direct AGB estimation from spectral vegetation indices outperformed the indirect approach, where AGB was retrieved from an empirical equation with the LAI inverted from an RTM. However, such discrepancies with the current study could be due to the non-destructive monitoring of LAI in the maize field, which may have penalized the indirect AGB retrieval in their study [21]. Additionally, the reduced performance of LAI inversion from the RTM (R^2=0.50) was a contributing factor.

In our study, LAI was estimated with high accuracy with generally improved performances in comparison to AGB models (Fig. 4), probably due to the higher correlation of leaf pigments per area to the reflectance in the visible spectrum [19]. An elegant approach to indirect AGB estimation in corn, winter wheat, and barley, was proposed by Wocher et al. [52] where the AGB was retrieved through the simultaneous inversion of the equivalent water thickness (C_W), leaf dry matter content, and LAI, using a GPR model in a hybrid workflow. Despite the distinctive accuracies obtained in their study ($R^2 = 0.80$), it is noteworthy that the AGB did not account for fruit organs "*as radiation is limited in penetrating thick tissues of plant fruit organs*". A similar conclusion was reported by Berger et al. [2], who identified fruit organs as an obstacle to N uptake retrieval in maize and wheat. Our approach, relying on empirical regressions between the LAI and AGB, accounts for the fruits into the AGB estimate, overcoming the limitations experienced by Wocher et al. [52]. However, in our work spectral campaigns were performed during green fruit enlargement (BBCH 704, 706, and 708) that were, therefore, exposed to the field of view of the sensor, thus penalizing the direct AGB retrieval, as opposed to the indirect approach. Since green fruits contribute to photosynthesis [54], their reflectance could be partially confused with the one of leaves. Therefore, the degree to which the AGB retrieval accuracy would be affected during the fruit's ripening stage remains questionable, as tomato reflectance can be influenced by the presence of red fruits [14].

Despite the general improved performance of the indirect AGB retrieval, the percentage error of the estimate was computed and plotted against the measured biomass. It is noteworthy that for AGB values below 1 t ha^{-1} the percentage error of the estimate is significantly increased in absolute values for both AGB retrieval approaches (Fig. 6c), suggesting that soil background still disturbs crop reflectance, resulting in diminished estimate accuracy [15]. While the percentage error is centered around 0 for $GPR_{D\text{-}AGB}$, a significant underestimation is observed for $GPR_{Ind\text{-}AGB}$, which is associated with the estimated intercept needed to build the $GPR_{Ind\text{-}AGB}$ model (Fig. 5b). The estimated intercept, assuming negative values at the early stage of the growing season (Fig. 5b), leads to an estimation of the biomass that even falls below 0 (data not shown). Based on the presented results, there is the potential for incorporating indirect AGB estimation into the development of DSS for precision nitrogen management, although the estimation at early growth stages turns out to be unreliable. Caution is advised when creating empirical relationships between LAI and AGB. The inclusion of SLA is desirable and might provide further improvements to the indirect AGB estimation [24].

Table 3
Retrieval models tested under two different approaches for N rate calculation.

N uptake approach	NNI approach
$SVR_{D\text{-}AGB}$ – $PLSR_{Nupt}$	$SVR_{D\text{-}AGB}$ – SVR_{NNI}
$GPR_{D\text{-}AGB}$ – $PLSR_{Nupt}$	$GPR_{D\text{-}AGB}$ – SVR_{NNI}
$GPR_{Ind\text{-}AGB}$ – $PLSR_{Nupt}$	$GPR_{Ind\text{-}AGB}$ – SVR_{NNI}

3.4. Simulation of dynamic N fertilization based on UAV spectral data

Based on the selected MLRAs, we explored two approaches addressing the optimal management of N fertilizer Nup-App and NNI-App (Table 3). Although $SVR_{D\text{-}AGB}$ presents itself as the best model for direct AGB retrieval (Fig. 4), the $GPR_{D\text{-}AGB}$ is also explored to ensure a fair comparison with the $GPR_{Ind\text{-}AGB}$ and better clarify the effect of the AGB retrieval approach in the overall N management. We made a simulation of the N fertilization for each N management approach, and the assessment was carried out utilizing the validation set of the retrieval MLRAs in two steps: i) definition of the optimal time for N fertilizer supply, and ii) calculation of the optimal N rate.

3.4.1. Definition of the optimal moment for N fertilization

The classification of the crop N status ("To be fertilized" or "To not be fertilized") is reported in Table 4 for the different approaches.

Most of the False Positive and False Negative errors are associated with the higher N treatments, which was expected as these experimental treatments were close to the trigger threshold (critical N uptake curve, or NNI=0.9), therefore the probability of misclassification is increased. Although the sample size of the higher N rate set ($n = 38$) is not large enough to provide extensive arguments, some trends can be highlighted. The AGB estimation approach did not yield tangible differences in the classification of the N status. Furthermore, the NNI-App yielded the highest classification accuracy in comparison with the Nup-App (21 % vs. 33.3 % on average, respectively). This result makes sense, as in the Nup-App the classification error depends on two sources of errors related to the simultaneous estimation of the AGB and N uptake, resulting in a higher cumulative error compared to the NNI-App. In the latter approach, the classification error of the N status depends only on the NNI model and is approximately equal to the residual variance not explained by the SVR_{NNI} model (equal to 1-R^2). Further assessments are recommended with a larger dataset.

3.4.2. Calculation of the optimal N rate

The optimal N rate (Eq. (1)) has been calculated only for the detected positive cases (TP and FP), both for the full validation dataset and only for higher N treatments for a better simulation of the model performances under field conditions. The Nup-App underperformed when compared to the NNI-App for each pair of models under assessment (Fig. 7). Here, the presence of outliers penalizes the estimation performances of the optimal N rate, increasing the RRMSE, and reduces the agreements between the measured and estimated N rate as shown by the rather low p_c [22]. Such outliers are associated with the

Table 4
Classification of crop N status according to the N rate estimation approaches performed on the validation set. Data are classified as True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). The table also reports the percentage error (ε%) of misclassification (FP+FN) for higher N rates (N60+N100) and for lower N rates (N0+N20).

N rate		Higher N rates					Lower N rates				
approach	Algorithms	TP	TN	FP	FN	ε%	TP	TN	FP	FN	ε%
N uptake	$SVR_{D\text{-}AGB;}$ $PLSR_{Nupt}$	15	12	6	5	28.9	39	1	0	0	0
N uptake	$GPR_{D\text{-}AGB;}$ $PLSR_{Nupt}$	13	11	9	5	36.8	37	1	2	0	5
N uptake	$GPR_{Ind\text{-}AGB;}$ $PLSR_{Nupt}$	15	10	7	6	34.2	33	0	6	1	17.5
NNI	SVR_{NNI}	18	12	4	4	21.0	37	1	1	1	5

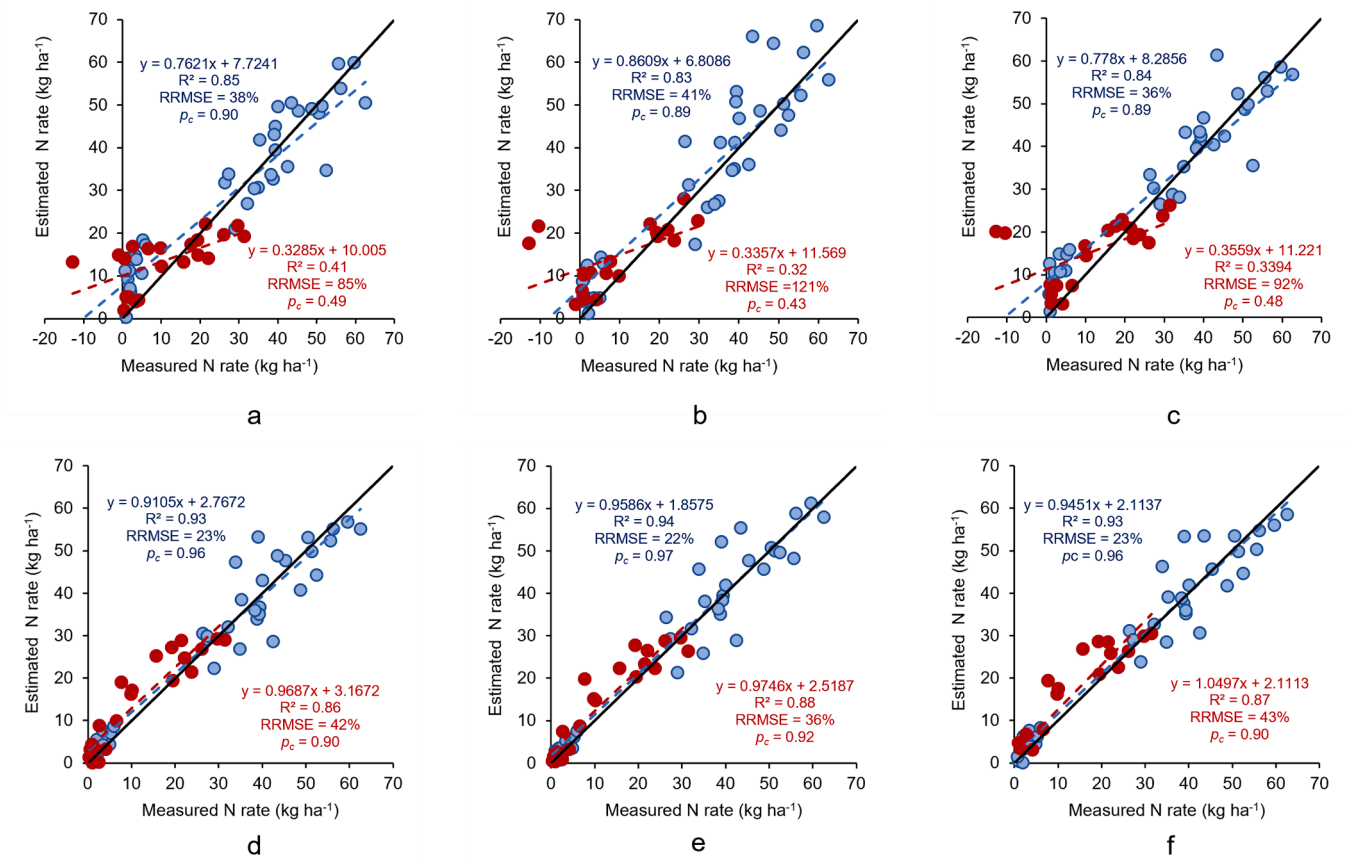

Fig. 7. Scatterplots of N rate estimated adopting the Nup-App (a, b, c) and NNI-App (d, e, f). Red equation refers to higher N rates treatments, while the blue equation refers to the full dataset. The different scatterplots refer to the following approaches: a $SVR_{D\text{-}AGB}$;$PLSR_{Nupt}$, b $GPR_{D\text{-}AGB}$;$PLSR_{Nupt}$, c $GPR_{Ind\text{-}AGB}$;$PLSR_{Nupt}$, d $SVR_{D\text{-}AGB}$;SVR_{NNI}, e $GPR_{D\text{-}AGB}$;SVR_{NNI}, f $GPR_{Ind\text{-}AGB}$;SVR_{NNI}.

misclassification of the crop N status which increased the frequency of false positive cases in the Nup-App (Table 4). According to the presented results, the NNI-App is more adequate than the Nup-App in estimating the optimal N rate from the remotely sensed spectral reflectance.

Similarly, Berger et al. [2] encountered challenges in estimating N uptake during the late growth stage due to the presence of fruit organs in corn, which aligns with our preference for the NNI-App. In this context, the direct estimation of AGB using the GPR model demonstrated a superior performance compared to other model combinations. However, indirect AGB retrieval is promising for its integration in N management strategies as an alternative to direct retrieval, especially when only limited AGB ground data are available, as indicated by the optimal p_c and R^2 values (Fig. 7f). Notably, LAI inversion through RTMs requires only a small dataset for validating the GPR algorithm [49]. Given that tomato growth follows an exponential trend from the development of the seventh leaves (approximately from 20 DAT), early growth stages play a crucial role in N management [10]. Nevertheless, when adopting the NNI-App, a high Relative Root Mean Square Error (RRMSE) of 36 % persists in N rate estimation (Fig. 7e). This can be attributed to the significant error associated with AGB estimation in the early stages up to 30 DAT (Fig. 6c), which also affects the accuracy of N rate estimation. Excluding observations from the first two early growth stages reduces the RRMSE to 28 % for $GPR_{D\text{-}AGB}$/SVR_{NNI} and 33 % for $GPR_{Ind\text{-}AGB}$/SVR_{NNI}, making it sufficiently accurate for N rate management (data not shown). Thus, optimal N management strategies during early growth stages should focus on alternative techniques that ensure critical N uptake until AGB≈1 t ha^{-1} (equivalent to 45.3 kg ha^{-1} of N). A starter fertilization between 20 and 40 kg N ha^{-1} after the transplant might be adequate to allow the crop to overcome the first critical stage (approximately up to 35 DAT). After that, the dynamic management of N guided by hyperspectral sensors can be adopted.

To the best of our knowledge, this is the first UAV-based study where the N rate is estimated by exploiting the full crop reflectance spectrum with nonparametric methods. While previous efforts by various authors have aimed to estimate optimal N rates from reflectance data, they typically relied on monitoring vegetation indices [7,26,37,51].

Accordingly, to the methodology employed in the current study, Cilia et al. [7] also determined the N rate following the estimation of NNI. Their methodology involved the remotely estimated AGB and estimated crop N concentration using empirical relationships with vegetation indices calculated from hyperspectral reflectance and derived NNI, N uptake, and N rate accordingly. No specific metrics were provided by Cilia et al. [7] to assess the error associated with the estimated N rate. However, the performance of AGB and NNI estimations in our study (with R^2 values of 0.94 and 0.80, respectively) exceeded those achieved in the study conducted by Cilia et al. [7], where corresponding R^2 values were 0.80 and 0.70, respectively. Relying on vegetation indices (as in [7]) rather than exploiting the full spectrum is limited in biophysical retrieval, as it may fail to include the hidden variability [6]. In this regard, in our estimation models, the number of PCs was chosen to include at least the 99.95 % of the cumulative variance, meaning that the models consider also the variance within the experimental treatments. This means that in real field conditions, our estimation models could estimate even the spatial variability of the field, finding application even for site-specific management. Verrelst et al. [46] demonstrated that, even when used as input data in a GPR algorithm, vegetation indices do not improve the prediction accuracy compared to using individual spectral bands. Rather, if only a few bands are available, up to a minimum of four, GPR trained by original bands still yields superior results as opposed to vegetation indices, which makes the algorithm ideal for multispectral monitoring [46]. In addition, the estimation accuracy of NNI is significantly enhanced by combining multispectral sensor data with soil, weather, and management practices, thereby improving the ability to generalize the N rate estimates to broader environmental and management conditions [26,51]. Nonetheless, besides the single-year location of the present study, the estimate of uncertainty provided by the GPR algorithm (standard deviation and coefficient of variation) can be used to trust the spatiotemporal portability of the model, as remarked by Verrelst et al. [47]. Future follow-up experiments are needed to assess the seasonal and location portability of the models, also including in the assessment different cultivars that are known to affect the reflectance profile [30], and an overfertilized N treatment to check the ability of the models to detect luxury N consumption. As different data source could improve the model accuracy [51], including in the machine learning algorithm the soil N content as input data is advisable. With the present study we demonstrated that AGB and NNI are crucial parameters for accurate N management, and we recommend their use with the NNI-App as presented in this paper for future DSSs development for N fertilization.

4. Conclusions

In this study, we demonstrated that the calculation approach (NNI-App or Nup-App) strongly affects the final accuracy of the optimal N rate, and should be considered when a new DSS is developed, with better results clearly shown by the NNI approach. However, some elements must be considered:

a) Using NNI had the best results when compared to the Nup-App, as the latter approach suffers from a double source of error associated with the estimation of N uptake and AGB. This made NNI-App the most effective also in estimating the optimal N rate: utilizing the GPR and SVR to estimate AGB and NNI by exploiting the full reflectance spectrum enables a sufficiently accurate estimate of the optimal N rate for processing tomatoes (R^2=0.88 and RRMSE=36 %), therefore hyperspectral imaging can support the dynamic N fertilization. As our work relied only on two workflows for the estimation of the N rate, alternative approaches can be explored (e.g., see [23] for more examples).
b) The high error experienced at AGB<1 t ha^{-1} suggests that hyperspectral reflectance cannot be adopted in precision N management schemes at early growth stages when soil background is prevalent. As the early growth stages are the most critical for N fertilization of processing tomatoes, further research is needed to improve AGB estimates when the soil coverage is low and the soil disturbances make the estimate unreliable, e.g., using LiDAR technology. Also, a high spatial resolution might enable the exclusion of soil pixels and improve the estimation performances but further assessments are required. Furthermore, it remains uncertain if spectral imaging can accurately support N rate estimation during the late growing season during fruit ripening in processing tomatoes.
c) We clearly show that the indirect approach to estimating the AGB utilizing empirical regression with the LAI is an alternative solution that performs similarly to the direct AGB estimation. This is particularly interesting when ground data are not sufficient for training an MLRA, and only a small ground dataset is available for validation (e.g., LAI retrieved from RTMs). Our approach, making use of empirical regression with LAI, allows to consider the fruits in the indirect estimation of the AGB, as opposed to alternative approaches [52]. However, we recommend optimizing empirical regressions that link LAI and AGB, for instance, including the specific leaf area (see Section 3.2). Also, as the R^2 between LAI and AGB reduces along the growing season, further assessment are warranted at late growth stages.
d) Lastly, the present dataset was generated in a single location and a single year, therefore, the model's portability to other farming conditions might be limited. Although the uncertainty estimate provided by GPR can be used to trust the spatiotemporal portability, validation under real field conditions is essential before assessing the efficacy of the NNI-App for scheduling in-season N rates in field trials. Data fusion approaches with agroclimatic data and soil N content are suggested, as they might ensure better spatiotemporal portability.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT in order to improve the readability of the manuscript. After using this tool/service, the authors reviewed and edited the content as needed and takes full responsibility for the content of the publication.

Ethics statement

Not applicable: This manuscript does not include human or animal research.

CRediT authorship contribution statement

Vito Aurelio Cerasola: Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Francesco Orsini:** Writing – review & editing, Validation, Project administration, Funding acquisition. **Giuseppina Pennisi:** Writing – review & editing, Visualization, Validation. **Gaia Moretti:** Visualization, Investigation. **Stefano Bona:** Writing – review & editing, Validation, Resources, Investigation, Funding acquisition, Formal analysis. **Francesco Mirone:** Visualization, Resources, Investigation. **Jochem Verrelst:** Writing – review & editing, Validation, Software, Resources, Methodology. **Katja Berger:** Writing – review & editing, Validation. **Giorgio Gianquinto:** Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Francesco Orsini reports financial support was provided by European Union. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have

appeared to influence the work reported in this paper.

Acknowledgements

The research leading to this publication has received funding from the European Union's PRIMA Program under grant agreement 2242 [Call 2022 Section 1 NEXUS WEFE IA], and from the European Union Next-GenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) – MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4 – D.D. 1032 17/06/2022, CN00000022) within the Agritech National Research Center. This manuscript reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them. Furthermore, we sincerely aknowledge Agata Morelli for the English proof reading of the manuscript, Massimiliano Beretta for sharing the plant materials used in the experimental set up, and Paola Gioacchini for performing the N concentration measurements.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.atech.2025.100802.

Data availability

Data will be made available on request.

References

[1] K. Berger, J. Verrelst, J.B. Féret, Z. Wang, M. Wocher, M. Strathmann, M. Danner, W. Mauser, T. Hank, Crop nitrogen monitoring: recent progress and principal developments in the context of imaging spectroscopy missions, Remote Sens. Environ. 242 (2020) 111758, https://doi.org/10.1016/j.rse.2020.111758.

[2] K. Berger, J. Verrelst, J.B. Féret, T. Hank, M. Wocher, W. Mauser, G. Camps-Valls, Retrieval of aboveground crop nitrogen content with a hybrid machine learning method, Int. J. Appl. Earth Obs. Geoinf. 92 (2020) 102174, https://doi.org/10.1016/j.jag.2020.102174.

[3] V.A. Cerasola, L. Perlotti, G. Pennisi, F. Orsini, G. Gianquinto, Potential use of superabsorbent polymer on drought-stressed processing tomato (Solanum lycopersicum L.) in a Mediterranean climate, Horticulturae 8 (8) (2022) 718, https://doi.org/10.3390/horticulturae8080718.

[4] V.A. Cerasola, G. Pennisi, F. Orsini, S. Bona, G. Gianquinto, Hybridization of vegetation index with agroclimatic data to improve biomass estimation in tomato for precision N management, in: Proceedings of the IEEE International Workshop on Metrology for Agriculture and Forestry (MetroAgriFor), IEEE, 2023, pp. 86–91, https://doi.org/10.1109/MetroAgriFor58484.2023.10424265.

[5] V.A. Cerasola, A. Di Marco, G. Pennisi, F. Orsini, S. Bona, G. Gianquinto, Modeling the canopy reflectance to predict tomato biomass for precision nitrogen management, in: Proceedings of the Precision Agriculture'23, Wageningen Academic, 2023, pp. 1049–1057, https://doi.org/10.3920/978-90-8686-947-3.

[6] A. Chlingaryan, S. Sukkarieh, B. Whelan, Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: a review, Comput. Electron. Agric. 151 (2018) 61–69, https://doi.org/10.1016/j.compag.2018.05.012.

[7] C. Cilia, C. Panigada, M. Rossini, M. Meroni, L. Busetto, S. Amaducci, M. Boschetti, P. Picchi, R. Colombo, Nitrogen status assessment for variable rate fertilization in maize through hyperspectral imagery, Remote Sens. 6 (7) (2014) 6549–6565, https://doi.org/10.3390/rs6076549.

[8] L. de Ávila Silva, R.P. Omena-Garcia, J.A. Condori-Apfata, P.M.D.A. Costa, N. M. Silva, F.M. DaMatta, A. Zsögön, W.L. Araùjo, E.A. de Toledo Picoli, R. Sulpice, A. Nunes-Nesi, Specific leaf area is modulated by nitrogen via changes in primary metabolism and parenchymal thickness in pepper, Planta 253 (2021) 1–13, https://doi.org/10.1007/s00425-020-03519-7.

[9] T. Dong, J. Liu, B. Qian, L. He, J. Liu, R. Wang, Q. Jing, C. Champagne, H. McNairn, J. Powers, Y. Shi, J.M. Chen, J. Shang, Estimating crop biomass using leaf area index derived from Landsat 8 and Sentinel-2 data, ISPRS J. Photogramm. Remote Sens. 168 (2020) 236–250, https://doi.org/10.1016/j.isprsjprs.2020.08.003.

[10] M. Farneselli, P. Benincasa, G. Tosti, E. Simonne, M. Guiducci, F. Tei, High fertigation frequency improves nitrogen uptake and crop performance in processing tomato grown with high nitrogen and water supply, Agric. Water Manag. 154 (2015) 52–58, https://doi.org/10.1016/j.agwat.2015.03.002.

[11] G. Gianquinto, P. Sambo, D. Borsato, Determination of SPAD threshold values for the optimisation of nitrogen supply in processing tomato, in: Proceedings of the International Symposium Towards Ecologically Sound Fertilisation Strategies for Field Vegetable Production 700, 2004, pp. 159–166, https://doi.org/10.17660/ActaHortic.2006.700.26.

[12] G. Gianquinto, P. Muñoz, A. Pardossi, S. Ramazzotti, D. Savvas, Soil fertility and plant nutrition, in: Proceedings of the Good Agricultural Practices for Greenhouse Vegetable Crops, 2013, p. 205.

[13] G. Gianquinto, M. Fecondini, M. Mezzetti, F. Orsini, Steering nitrogen fertilisation by means of portable chlorophyll meter reduces nitrogen input and improves quality of fertigated cantaloupe (Cucumis melo L. var. cantalupensis Naud.), J. Sci. Food Agric. 90 (3) (2010) 482–493, https://doi.org/10.1002/jsfa.3843.

[14] G. Gianquinto, F. Orsini, G. Pennisi, S. Bona, Sources of variation in assessing canopy reflectance of processing tomato by means of multispectral radiometry, Sensors 19 (21) (2019) 4730, https://doi.org/10.3390/s19214730.

[15] M.L. Gnyp, Y. Miao, F. Yuan, S.L. Ustin, K. Yu, Y. Yao, S. Huang, G. Bareth, Hyperspectral canopy sensing of paddy rice aboveground biomass at different growth stages, Field Crops Res. 155 (2014) 42–55, https://doi.org/10.1016/j.fcr.2013.09.023.

[16] P.M. Hansen, J.K. Schjoerring, Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression, Remote Sens. Environ. 86 (4) (2003) 542–553, https://doi.org/10.1016/S0034-4257(03)00131-7.

[17] M. Hoogmoed, A. Neuhaus, S. Noack, V.O. Sadras, Benchmarking wheat yield against crop nitrogen status, Field Crops Res. 222 (2018) 153–163, https://doi.org/10.1016/j.fcr.2018.03.013.

[18] V. Houles, M. Guerif, B. Mary, Elaboration of a nitrogen nutrition indicator for winter wheat based on leaf area index and chlorophyll content for making nitrogen recommendations, Eur. J. Agron. 27 (1) (2007) 1–11, https://doi.org/10.1016/j.eja.2006.10.001.

[19] T. Kattenborn, F. Schiefer, P. Zarco-Tejada, S. Schmidtlein, Advantages of retrieving pigment content [µg/cm2] versus concentration [%] from canopy reflectance, Remote Sens. Environ. 230 (2019) 111195, https://doi.org/10.1016/j.rse.2019.05.014.

[20] A. Kayad, F.A. Rodrigues Jr, S. Naranjo, M. Sozzi, F. Pirotti, F. Marinello, U. Schulthess, P. Defourny, B. Gerard, M. Weiss, Radiative transfer model inversion using high-resolution hyperspectral airborne imagery–Retrieving maize LAI to access biomass and grain yield, Field Crops Res. 282 (2022) 108449, https://doi.org/10.1016/j.fcr.2022.108449.

[21] A. Klingler, A. Schaumberger, F. Vuolo, L.B. Kalmár, E.M. Pötsch, Comparison of direct and indirect determination of leaf area index in permanent grassland, PFG–J. Photogramm. Remote Sens. Geoinf. Sci. 88 (5) (2020) 369–378, https://doi.org/10.1007/s41064-020-00119-8.

[22] L. Lin, A concordance correlation coefficient to evaluate reproducibility, Biometrics 45 (1) (1989) 255–268, https://doi.org/10.2307/2532051.

[23] X. Li, S.T. Ata-UI-Karim, Y. Li, F. Yuan, Y. Miao, K. Yoichiro, T. Cheng, L. Tang, X. Tian, X. Liu, Y. Tian, Y. Zhu, W. Cao, Q. Cao, Advances in the estimations and applications of critical nitrogen dilution curve and nitrogen nutrition index of major cereal crops. A review, Comput. Electron. Agric. 197 (2022) 106998, https://doi.org/10.1016/j.compag.2022.106998.

[24] Z. Li, C. Zhan, S. Hu, L. Ning, L. Wu, H. Guo, Implementation of a dynamic specific leaf area (SLA) into a land surface model (LSM) incorporated crop-growth model, Comput. Electron. Agric. 213 (2023) 108238, https://doi.org/10.1016/j.compag.2023.108238.

[25] Y. Liu, H. Feng, J. Yue, Y. Fan, X. Jin, X. Song, H. Yang, G. Yang, Estimation of potato above-ground biomass based on vegetation indices and green-edge parameters obtained from UAVs, Remote Sens. 14 (21) (2022) 5323, https://doi.org/10.3390/rs14215323.

[26] J. Lu, E. Dai, Y. Miao, K. Kusnierek, Improving active canopy sensor-based in-season rice nitrogen status diagnosis and recommendation using multi-source data fusion with machine learning, J. Clean. Prod. 380 (2022) 134926, https://doi.org/10.1016/j.jclepro.2022.134926.

[27] G. Marcolini, P. Gioacchini, M. Quartieri, E. Baldi, G. Sorrenti, M. Toselli, Evaluation of plant derivatives of Meliaceae family as a source of nitrogen for trees, J. Plant Nutr. 45 (11) (2022) 1642–1653, https://doi.org/10.1080/01904167.2021.2014868.

[28] C. Masso, F. Baijukya, P. Ebanyat, S. Bouaziz, J. Wendt, M. Bekunda, B. Vanlauwe, Dilemma of nitrogen management for future food security in sub-Saharan Africa–a review, Soil Res. 55 (6) (2017) 425–434.

[29] P. Miphokasap, K. Honda, C. Vaiphasa, M. Souris, M. Nagai, Estimating canopy nitrogen concentration in sugarcane using field imaging spectroscopy, Remote Sens. 4 (6) (2012) 1651–1670, https://doi.org/10.3390/rs4061651.

[30] T.J. Nigon, D.J. Mulla, C.J. Rosen, Y. Cohen, V. Alchanatis, J. Knight, R. Rud, Hyperspectral aerial imagery for detecting nitrogen stress in two potato cultivars, Comput. Electron. Agric. 112 (2015) 36–46, https://doi.org/10.1016/j.compag.2014.12.018.

[31] F.M. Padilla, M.T. Peña-Fleitas, M. Gallardo, R.B. Thompson, Threshold values of canopy reflectance indices and chlorophyll meter readings for optimal nitrogen nutrition of tomato, Ann. Appl. Biol. 166 (2) (2015) 271–285, https://doi.org/10.1111/aab.12181.

[32] A.B. Pascual-Venteo, E. Portalés, K. Berger, G. Tagliabue, J.L. Garcia, A. Pérez-Suay, J.P. Rivera-Caicedo, J. Verrelst, Prototyping crop traits retrieval models for CHIME: dimensionality reduction strategies applied to PRISMA data, Remote Sens. 14 (10) (2022) 2448, https://doi.org/10.3390/rs14102448.

[33] C.E. Rasmussen, C.K. Williams, Gaussian Processes for Machine Learning, 1, MIT Press, Cambridge, MA, 2006, p. 159.

[34] J.P. Rivera-Caicedo, J. Verrelst, J. Muñoz-Marí, J. Moreno, G. Camps-Valls, Toward a semiautomatic machine learning retrieval of biophysical parameters, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 7 (4) (2014) 1249–1259, https://doi.org/10.1109/JSTARS.2014.2298752.

[35] J.P. Rivera-Caicedo, J. Verrelst, J. Muñoz-Marí, G. Camps-Valls, J. Moreno, Hyperspectral dimensionality reduction for biophysical variable statistical retrieval, ISPRS J. Photogramm. Remote Sens. 132 (2017) 88–101, https://doi.org/10.1016/j.isprsjprs.2017.08.012.
[36] A. Savitzky, M.J. Golay, Smoothing and differentiation of data by simplified least squares procedures, Anal. Chem. 36 (8) (1964) 1627–1639.
[37] P.C. Scharf, J.A. Lory, Calibrating reflectance measurements to predict optimal sidedress nitrogen rate for corn, Agron. J. 101 (3) (2009) 615–625, https://doi.org/10.2134/agronj2008.0111.
[38] J. Scholberg, B.L. McNeal, K.J. Boote, J.W. Jones, S.J. Locascio, S.M. Olson, Nitrogen stress effects on growth and nitrogen accumulation by field-grown tomato, Agron. J. 92 (1) (2000) 159–167, https://doi.org/10.2134/agronj2000.921159x.
[39] A.J. Smola, B. Schölkopf, A tutorial on support vector regression, Stat. Comput. 14 (2004) 199–222, https://doi.org/10.1023/B:STCO.0000035301.49549.88.
[40] R.D. Snee, Validation of regression models: methods and examples, Technometrics (1977) 415–428.
[41] H. Tao, H. Feng, L. Xu, M. Miao, H. Long, J. Yue, Z. Li, G. Yang, X. Yang, L. Fan, Estimation of crop growth parameters using UAV-based hyperspectral remote sensing data, Sensors 20 (5) (2020) 1296, https://doi.org/10.3390/s20051296.
[42] F. Tei, P. Benincasa, M. Guiducci, Critical nitrogen concentration in processing tomato, Eur. J. Agron. 18 (1–2) (2002) 45–55.
[43] F. Tei, P. Benincasa, M. Farneselli, G. Tosti, M. Guiducci, Environmentally sustainable nitrogen nutrition management in processing tomato, Acta Hortic. (1081) (2015) 41–48, https://doi.org/10.1016/S1161-0301(02)00096-5.
[44] F. Tei, S. De Neve, J. de Haan, H.L. Kristensen, Nitrogen management of vegetable crops, Agric. Water Manag. 240 (2020) 106316, https://doi.org/10.1016/j.agwat.2020.106316.
[45] J. Verrelst, J. Rivera, L. Alonso, J. Moreno, ARTMO: an automated radiative transfer models operator toolbox for automated retrieval of biophysical parameters through model inversion, in: Proceedings of the EARSeL 7th SIG-Imaging Spectroscopy Workshop, Edinburgh, UK, 2011, pp. 11–13.
[46] J. Verrelst, L. Alonso, G. Camps-Valls, J. Delegido, J. Moreno, Retrieval of vegetation biophysical parameters using Gaussian process techniques, IEEE Trans. Geosci. Remote Sens. 50 (5) (2011) 1832–1843.
[47] J. Verrelst, J.P. Rivera, J. Moreno, G. Camps-Valls, Gaussian processes uncertainty estimates in experimental Sentinel-2 LAI and leaf chlorophyll content retrieval, ISPRS J. Photogramm. Remote Sens. 86 (2013) 157–167, https://doi.org/10.1016/j.isprsjprs.2013.09.012.
[48] J. Verrelst, G. Camps-Valls, J. Muñoz-Marí, J.P. Rivera, F. Veroustraete, J. G. Clevers, J. Moreno, Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties–a review, ISPRS J. Photogramm. Remote Sens. 108 (2015) 273–290, https://doi.org/10.1016/j.isprsjprs.2015.05.005.
[49] J. Verrelst, J.P. Rivera, F. Veroustraete, J. Muñoz-Marí, J.G. Clevers, G. Camps-Valls, J. Moreno, Experimental Sentinel-2 LAI estimation using parametric, non-parametric and physical retrieval methods–a comparison, ISPRS J. Photogramm. Remote Sens. 108 (2015) 260–272, https://doi.org/10.1016/j.isprsjprs.2015.04.013.
[50] J. Vos, P.C. Struik, Crop responses to nitrogen, in: J.L. Meulenbroek (Ed.), Agriculture and Environment in Eastern Europe and The Netherlands, Wageningen Agricultural University, Wageningen, 1992, pp. 195–205.
[51] X. Wang, Y. Miao, R. Dong, H. Zha, T. Xia, Z. Chen, K. Kusnierek, G. Mi, H. Sun, M. Li, Machine learning-based in-season nitrogen status diagnosis and side-dress nitrogen recommendation for corn, Eur. J. Agron. 123 (2021) 126193, https://doi.org/10.1016/j.eja.2020.126193.
[52] M. Wocher, K. Berger, J. Verrelst, T. Hank, Retrieval of carbon content and biomass from hyperspectral imagery over cultivated areas, ISPRS J. Photogramm. Remote Sens. 193 (2022) 104–114, https://doi.org/10.1016/j.isprsjprs.2022.09.003.
[53] S. Wold, M. Sjöström, L. Eriksson, PLS-regression: a basic tool of chemometrics, Chemom. Intell. Lab. Syst. 58 (2) (2001) 109–130, https://doi.org/10.1016/S0169-7439(01)00155-1.
[54] H.L. Xu, L. Gauthier, Y. Desjardins, A. Gosselin, Photosynthesis in leaves, fruits, stem and petioles of greenhouse-grown tomato plants, Photosynthetica 33 (1997) 113–123, https://doi.org/10.1023/A:1022135507700.
[55] J. Zhao, J. Li, Q. Liu, L. Yang, A preliminary study on mechanism of LAI inversion saturation, in: Proceedings of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 39, 2012, pp. 77–81, https://doi.org/10.5194/isprsarchives-XXXIX-B1-77-2012.
[56] N. Ziadi, G. Bélanger, F. Gastal, A. Claessens, G. Lemaire, N. Tremblay, Leaf nitrogen concentration as an indicator of corn nitrogen status, Agron. J. 101 (4) (2009) 947–957, https://doi.org/10.2134/agronj2008.0172x.