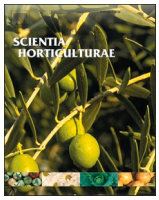
Scientia Horticulturae 347 (2025) 114144

Contents lists available at ScienceDirect

Scientia Horticulturae

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



Research Paper

Non-destructive sensing of nitrogen and phosphorus contents in plants via hyperspectral imaging of *Hydrangea macrophylla* at the seeding stage

Jun Yang [a], Guochun Shen [b,*], Jun Qin [a,*]

[a] *Urban Horticulture Research and Extension Center, Shanghai Chenshan Botanical Garden, Shanghai 201602, PR China*
[b] *Zhejiang Tiantong Forest Ecosystem National Observation and Research Station, School of Ecological and Environmental Sciences, East China Normal University, Shanghai 200241, PR China*

ARTICLE INFO

Keywords:
Hyperspectral reconstruction
LNC
LPC
Machine learning

ABSTRACT

Hyperspectral imaging is a powerful, non-destructive tool that has shown promise in estimating plant nutrition but remains underexplored for *Hydrangea macrophylla*, an important ornamental shrub. In this study, we developed regression and discrimination models for assessing nitrogen (N) and phosphorus (P) statuses in hydrangea at the seeding stage using spectral data combined with machine learning techniques. Spectral reflectance data, captured across various nutritional states following different N and P fertilisation treatments, were paired with laboratory-measured leaf nitrogen content (LNC) and leaf phosphorus content (LPC) data. Spectral reflectance data captured under various N and P treatments were processed with first derivative (FD) and continuous wavelet transform (CWT) to improve the data quality. The results of the correlation analysis revealed that the CWT was more strongly correlated with N (r = –0.90) and P (r = –0.87) than with FD. The machine learning models, which were provided with full-band and wavelet features, showed that the partial least squares regression (PLSR) model, which was integrated with the CWT, could accurately predict LNC (R^2 = 0.947, RRMSE = 9.6 %) and LPC (R^2 = 0.827, RRMSE = 10.6 %). Additionally, the PNN model classified nutrient statuses with over 95 % accuracy. Notably, N predictions outperformed those for P, probably because of weaker spectral correlations with P. These findings highlight the potential of hyperspectral imaging and machine learning for precise nutrient management in hydrangea cultivation and contribute to sustainable agricultural practices globally.

1. Introduction

Hydrangea macrophylla, commonly known as bigleaf or French hydrangea, is cultivated worldwide for ornamental and medicinal purposes (Li et al., 2018). It is grown as deciduous woody shrubs for landscaping or used as potted plants and cut flowers (Anderson et al., 2009). According to the 2019 Census of Horticultural Specialties, the sale value of hydrangeas as nursery stock in the United States was $106,862,000, ranking hydrangeas as the third most productive cultivated deciduous shrub (USDA, 2019). Although it is economically important, hydrangea cultivation faces challenges related to high nutrient demands, particularly for nitrogen (N) and phosphorus (P), which are required for many key physiological and biochemical processes (Guo et al., 2018). Nutrient deficiencies or excesses can limit productivity and lead to environmental issues such as pollution from fertiliser runoff (Shreckhise et al., 2019). Therefore, the rational application of N and P fertilisers is highly important for the production of hydrangea nurseries and for reducing global environmental pollution.

Recent advancements in agricultural technologies have revealed ways to address environmental and economic challenges while improving human life and promoting sustainable development (Amiri et al., 2025; Hanafi et al., 2024; Imani et al., 2025; Zinatloo-Ajabshir et al., 2025; Zonarsaghar et al., 2022). These innovative techniques highlight the importance of integrating sustainable practices into horticulture. Traditional nutrition diagnosis methods (such as the Kjeldahl digestion method) are time-consuming, expensive, and labour-intensive. Hyperspectral imaging is a valuable tool for non-destructive nutrient diagnosis. It is accurate, sustainable, and can process high-dimensional data (Ball et al., 2022). This technology has been applied to estimate leaf nitrogen content (LNC) and leaf phosphorus content (LPC) in various crop species, including rice, wheat, and cotton (Li et al., 2021; Sun et al., 2019; Wang et al., 2021). However, its application to hydrangea, a

* Corresponding authors.
E-mail addresses: gcshen@des.ecnu.edu.cn (G. Shen), cszwyqinjun@126.com (J. Qin).

https://doi.org/10.1016/j.scienta.2025.114144
Received 30 April 2024; Received in revised form 2 April 2025; Accepted 14 April 2025
Available online 27 May 2025

popular and economically important ornamental plant, remains underexplored.

Recently, machine learning, in conjunction with hyperspectral imaging, has offered new possibilities for improving the accuracy and efficiency of nutrient diagnosis. Models such as partial least squares regression (PLSR), random forest (RF), and convolutional neural networks (CNNs) have shown promise in estimating nutrient content from spectral data (Guo et al., 2018; Pourdarbani et al., 2021; Wang et al., 2023). The predictive accuracy of the N models at the leaf scale shows considerable variation, with R^2 values ranging from 0.37 to 0.99. In contrast, P models present R^2 values ranging from 0.32 to 0.95, indicating moderate-to-high precision (Watt et al., 2020). This variance highlights the disparate efficacy of different models in estimating N and P levels, highlighting the importance of choosing appropriate input variables for machine learning algorithms. A promising approach to further refine modelling accuracy and reduce dimensionality involves integrating wavelet analysis with machine learning.

The continuous wavelet transformation (CWT), a signal processing technique, excels in localising information in both the frequency and time domains and deconstructing signals across various scales and positions (Cheng et al., 2011). Its application has improved model precision for assessing LNC and LPC, outperforming traditional spectral transformations such as the first derivative (FD) (Gu et al., 2022). Some studies have revealed the efficacy of combining CWT with PLSR for extracting plant water status, phosphorus (P), or biomass from hyperspectral reflectance data (Gu et al., 2022; Zhang et al., 2023; Zhuang et al., 2023). Additionally, timely guidance and early detection of the nutrition status of hydrangea are highly important to farmers. The probabilistic neural network (PNN) is a robust method for rapidly detecting nutrient status and is crucial for timely agricultural management (Azlah et al., 2019). It has been applied in fields such as plant nutrition diagnosis (Lou et al., 2022), foliar biotic damage discrimination (Liu et al., 2018), and leaf identification (Abdulazeez et al., 2021).

Despite the advances in hyperspectral imaging and machine learning for plant nutrition, challenges remain in generalising these models owing to variations in spectral responses under different nutritional states, particularly during crucial growth phases. This study aimed to (1) elucidate the spectral responses of hydrangea leaves to different N and P fertilisation via hyperspectral imaging technology; (2) assess the efficacy of machine learning models, comparing different spectral transformations (FD and CWT), to identify an optimal model for estimating the N and P contents of hydrangea leaves; (3) validate the precision of discriminants for different N and P statuses using the PNN model. These findings can contribute to improving precise nutrient management practices in hydrangea cultivation and promote global efforts in sustainable agricultural development.

2. Materials and methods

2.1. Experimental design and crop growing

A pot experiment was conducted at Chenshan Botanical Garden in Songjiang District, Shanghai (31.08 N, 121.77 E). Root cuttings of *H. macrophylla* 'Hanatemari' (obtained from Hangzhou Landscaping Inc. in August 2021) were planted in pots (13 cm external diameter, 9 cm bottom diameter, and 11 cm height) filled with a substrate mixture of peat moss, perlite, and coconut husk (3:1:1 volume ratio). The chemical characteristics of the substrate were as follows: pH, 6.45; EC, 1.27 $mS{\cdot}cm^{-1}$; total N, 7.96 $mg{\cdot}g^{-1}$; total P, 1.09; total K, 3.33 $mg{\cdot}g^{-1}$; and organic matter, 865.18 $g{\cdot}kg^{-1}$. No basal fertiliser was added to the substrate before the treatment was applied. The plants were cultivated for eight weeks in a climatic room with no sunlight under a 16-h photoperiod with a PPFD of 101.93 μmol $m^{-2}s^{-1}$ supplied by LED lights. The temperature of the chamber was maintained at 25 ±0.5°C, and the relative humidity (RH) was 60 %. The plants were watered with unfertilised tap water during this period.

The plants had 4–5 leaf pairs before the treatments began. From November 2021 to January 2022, the plants were subjected to four N treatments (0, 0.4, 0.8, and 1.2 g per pot, corresponding to N_0, N_1, N_2, and N_3, respectively) and four P treatments (0, 0.15, 0.3, and 0.45 g per pot, corresponding to P_0, P_1, P_2, and P_3, respectively). The plants without N and P fertilisers were treated as the control (CK). A fixed K dosage (1.48 g K_2SO_4 per pot) was applied across all treatments. Based on our preliminary fertiliser experimental results, N_2, P_2, and K were determined using the appropriate application rates (Wang et al., 2022). A completely randomised block design with 30 replications was used to allocate each treatment (Table 1).

Urea (containing 46.6 % N), calcium superphosphate (containing 14.5 % P_2O_5), and K_2SO_4 (containing 54.1 % K_2O) were selected as the N, P, and K fertilisers, respectively. Seven parts of urea and K_2SO_4 were applied every seven days. Calcium superphosphate was applied in three portions every fifteen days. Images and plant tissue samples were collected at four development stages, namely, G0, G1, G2, and G3, which corresponded to 0, 20, 40, and 60 days following the initial fertilisation, respectively. Specifically, G0, G1, G2, and G3 were on November 23, 2021, December 13, 2021, January 2, 2022, and January 22, 2022.

2.2. Image acquisition and processing

A GaiaSky-mini2-VN hyperspectral imaging system (Dualix Spectral Imaging Technology Co., Ltd., Sichuan, China) was used to acquire hyperspectral images of the leaves. The system offered a spectral range of 400 to 1000 nm and a spectral resolution of 3 ±0.5 nm. The hyperspectral imager, equipped with a 16 mm lens, was mounted at a height of about 30 cm within a field of view of 36.25°. The images were captured at a spatial resolution of 0.2 mm, with dimensions of 960 × 1040 pixels. The push-broom method was used, which involved the camera moving across the scene to construct a two-dimensional image. Full-spectrum halogen lamps served as the light source for the system. To decrease the effect of ambient light, the entire setup was enclosed in a black box (Fig. 1).

The leaf sample was placed in the centre of the black background plate to ensure that it was spread as flat as possible without significant curling or folding. Hyperspectral images were captured using the data acquisition software SpecView. Reflectance values from the raw data were obtained via white and dark calibrations, following the approach described in another study (Baranowski et al., 2015). The calibrated hyperspectral images were imported into R-4.2.1, where the edges of the leaf were identified using the raster package, and the average spectral curve for all pixels within the leaf in each image was calculated. The second pair of fully expanded leaves, not shaded by other leaves, was picked to acquire hyperspectral images from the G1 to G3 growth periods, with 10–20 replicates for each treatment. In total, 405 leaf samples were collected.

Table 1
Fertiliser treatments of *H. macrophylla* 'Hanatemari'.

NO.	Treatment	Each nutrient consumption ($g{\cdot}pot^{-1}$)			Fertilising amount ($g{\cdot}pot^{-1}$)		
		N	P_2O_5	K_2O	Urea	Calcium superphosphate	K_2SO_4
T1	N_0	0	0.3	0.8	0	2.07	1.48
T2	N_1	0.4	0.3	0.8	0.86	2.07	1.48
T3	N_2&P_2	0.8	0.3	0.8	1.72	2.07	1.48
T4	N_3	1.2	0.3	0.8	2.58	2.07	1.48
T5	P_0	0.8	0	0.8	1.72	0	1.48
T6	P_1	0.8	0.15	0.8	1.72	1.03	1.48
T7	P_3	0.8	0.45	0.8	1.72	3.10	1.48
T8	CK	0	0	0.8	0	0	1.48

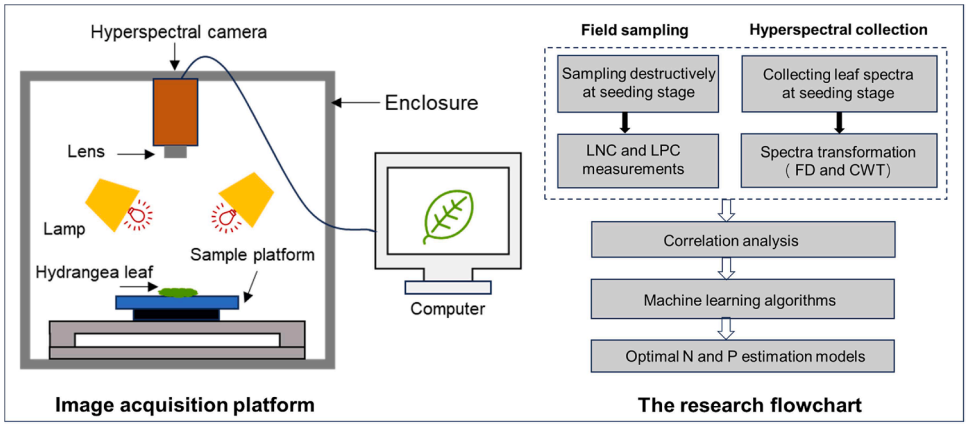


Fig. 1. Image acquisition platform and flowchart of the study.

2.3. Analytical reference measurements

After the spectral data were collected, the leaf samples were harvested to measure LNC and LPC. The samples were first oven-dried at 105°C for 30 min and then dried at 70°C until they reached a constant dry mass (Wang et al., 2022). The leaf samples were finely ground and analysed for their N and P contents using the Kjeldahl method (Li et al., 2021) and the molybdate blue colorimetric method (Zhang et al., 2023), respectively.

To assess the aboveground biomass dry matter of hydrangea, three plants per treatment, labelled G1, G2, and G3, were harvested during the seedling stage. Like the leaf samples, all fresh shoot samples were oven-dried at 105°C for 30 min and then at 70°C until a constant weight was reached to measure the shoot dry weight.

2.4. Spectral transformation

Before constructing the model, spectral reflectance data were pre-processed through standard normal variate (SNV) transformation and smoothed with a Savitzky-Golay (SG) filter, which considered the original spectra. To reduce noise and accurately reveal spectral characteristics, two transformation methods were used: FD and CWT. Sensitive spectral bands were identified by correlating the transformed spectra with nutrient content. The Pearson correlation coefficient (r) was calculated between individual spectral bands and nutrient contents. Using CWT, the analysis of the correlation between the original spectral curves and Gaussian functions across various positions and scales yielded continuous wavelet energy coefficients. The Mexican hat wavelet (Mexh) served as the selected mother wavelet basis function, with decomposition scales set from 1 to 10, namely, 2^1, 2^2..., 2^{10} (Wang et al., 2020). Wavelet features were then extracted based on coefficient values correlating wavelet energy coefficients with nutrient content (Cheng et al., 2024).

2.5. Model building and result validation

The data were processed and analysed using the R statistical environment. Three models, including PLSR, RF, and CNN, were developed to predict the LNC and LPC of hydrangea. For model development, three algorithms were executed using the pls, random forest, and torch packages in the R language (release 4.2.1) by modifying the methods described in another study (Guo et al., 2017). The CNN architecture was a basic one-dimensional convolutional neural network. It consisted of three convolutional layers with 50, 20, and 10 filters. Each filter had a kernel size of 3. A ReLU activation function followed every convolutional layer. The feature maps from the third convolutional layer were flattened into a single vector and passed to a dense layer with a single node for output. The mean squared error was used as the loss function to train the model parameters. A simple random sampling method was used to allocate 70 % of the total samples to the training dataset, with the remaining 30 % of the samples reserved for the test dataset. To prevent overfitting problems, 10-fold cross-validation was used in our model training process. The model performance was evaluated using the coefficient of determination (R^2) and the relative root-mean-square error (RRMSE) on the test dataset. The RRMSE was calculated by dividing the root mean square error (RMSE) by the average of the observed values (Ghorbani et al., 2017; Velumani et al., 2021). The model accuracy was rated as excellent when the RRMSE was < 10 %, good when 10 % $<$ RRMSE $<$ 20 %, fair when 20 % $<$ RRMSE $<$ 30 %, and poor if the RRMSE was $\geq$ 30 % (Li et al., 2013).

To classify hydrangea leaves based on different nutrient statuses, the hyperspectral data were processed using a supervised classification model (PNN). Similarly, 70 % of the total samples were selected as the training dataset, and the remaining 30 % were selected as the test dataset. The accuracy was evaluated based on overall accuracy (OA)= (TN + TP)/ (TN + TP + FN + FP), precision= TP/ (TP + FP), and recall= TP/ (TP + FN). However, accuracy alone is not sufficient for evaluating model performance, especially for imbalanced data or multiclass data (Guo et al., 2017). Therefore, Cohen's kappa coefficient (kappa) was also used to evaluate the model performance and was calculated as follows: kappa= $(p_0 - p_e)/(1-p_e)$, where p_0 was the observed agreement and p_e was the expected agreement. All analyses were performed in MATLAB 2023a and R-4.2.1. A flowchart of the study is shown in Fig. 1.

3. Results

3.1. Biomass, N, and P concentrations

The biomass, LNC, and LPC were significantly different under the N and P treatments (Fig. 2A). Biomass responded positively to optimal N addition (N_1–N_2) but was negatively affected by low and high N addition rates (N_0 and N_3). At the G3 stage, all shoots died under N_3 due to the high N level, resulting in no biomass recorded, while the biomass was greatest under N_1 (1.14 g·plant^{-1}). This reflected the N statuses of the deficient (N_0), medium (N_1–N_2), and surplus (N_3) groups. LNC decreased with time from G_1 to G_3 at the seedling stage. N treatment significantly affected LNC, which increased sharply from N_0 to N_2 but slightly decreased in the N_3 group at all stages.

Similarly, biomass was also significantly affected by P fertilisation, which was greatest under P_2 (0.72 g·plant^{-1}) and lowest under P_3 (0.19 g·plant^{-1}). Additionally, the shoots with an optimum P status presented a constant increase in biomass (P_1–P_2) but were limited by low (P_0) and surplus fertiliser addition (P_3) (Fig. 2A). The LNC exhibited greater variation (15.68 to 49.84 mg·g^{-1}) than the LPC (ranging from 2.63 to

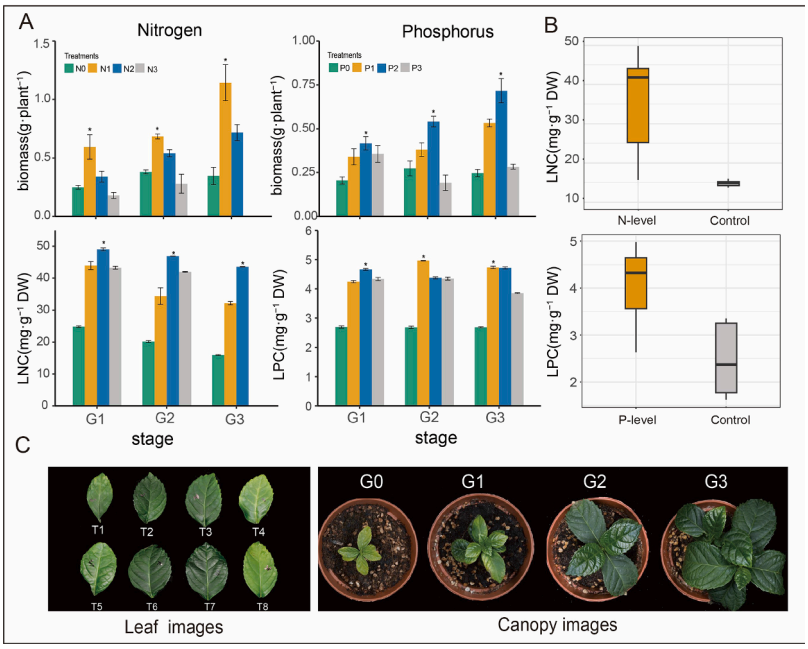


Fig. 2. Phenotypes of hydrangea seedlings grown under different N and P conditions. N was applied at four rates: N_0 (0 g·pot^{-1}), N_1 (0.4 g·pot^{-1}), N_2 (0.8 g·pot^{-1}), and N_3 (1.2 g·pot^{-1}). The P level was applied at four rates: P_0 (0 g·pot^{-1}), P_1 (0.15 g·pot^{-1}), P_2 (0.3 g·pot^{-1}), and P_3 (0.45 g·pot^{-1}). The control was treated without N or P fertilisers. G1, G2, and G3 correspond to the development phases at 20, 40, and 60 days following the initial fertilisation, respectively. (A) The biomass, LNC, and LPC under different treatments at the seedling stage are presented. At the G3 stage, all shoots died in the N_3 group due to high N levels, resulting in no recorded biomass. (B) Variation in LNC and LPC. (C) Leaves and canopies of seedlings under different treatments are shown. An asterisk (*) indicates a significant difference at $P < 0.05$.

4.98 mg·g^{-1}) (Fig. 2B). The phenotypic variations in hydrangea and their responses to N and P addition at the seedling stage were denoted as G1, G2, and G3, respectively (Fig. 2C).

3.2. Spectral features of leaves

As shown in Fig. 3A, the original spectra of all leaf samples had similar spectral characteristics, with prominent reflection peaks and absorption valleys across 400–1000 nm. A reflection peak was found around the green region (550 nm), whereas a prominent absorption valley occurred at about 670 nm in the red region. After applying FD transformation, baseline offsets and random noise were removed from the original spectra, leaving subtle yet significant wavelength information. The red edge position, located at about 710 nm, corresponded to the wavelength at which the FD value reached its maximum value (Fig. 3B).

The original spectra and the spectra after FD transformation under different N and P treatments are shown in Fig.4. These treatment effects were detected in either the visible (400–760 nm) or near-infrared (NIR, 760–900 nm) spectral region. At the N level, hydrangea leaves under N_0 presented the highest reflectance in the visible region (Fig. 4A). In contrast, at the P level, hydrangea leaves under P_3 presented greater reflectance in the visible region than those under the other treatments (Fig. 4C).

The "red-edge" region was involved in the transition from visible to NIR wavelengths. The FD values of the different N treatments were $N_0 > N_3 > N_2 > N_1$ at the peak position near 710 nm, and the red edge positions corresponding to N_0, N_1, N_2, and N_3 were 705, 713, 709, and 709 nm, respectively (Fig. 4B). In the P treatments, the values were $P_3 > P_2 > P_1 > P_0$, and the red edge positions corresponding to P_0, P_1, P_2, and P_3 were 728, 728,728, and 713 nm, respectively (Fig. 4D). When N deficient (N_0) and P surplus (P_3) leaves were present, the feature of the spectra characteristically shifted to lower wavelengths, a phenomenon observed in the range of the NIR region by FD. In general, the P treatments presented a lower variation in reflectance than the N treatments at all wavebands.

After a 10-dimensional CWT, the spectral curves were relatively flat, lacking identifiable features at the lower scales [1, 2, 3]. However, from scales [4, 5, 6], the spectra became wave-like before transitioning to parabolic forms at scales [7, 8] and finally appeared as nearly straight lines at the highest scales [9, 10]. The spectral features under the N and P treatments (Fig. 5A, B) were significantly more pronounced than those

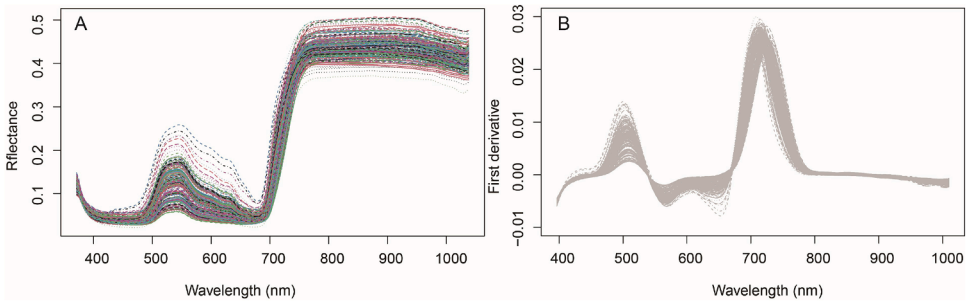


Fig. 3. Spectra of all hydrangea leaves. (A) The original spectra. (B) The spectra after FD transformation.

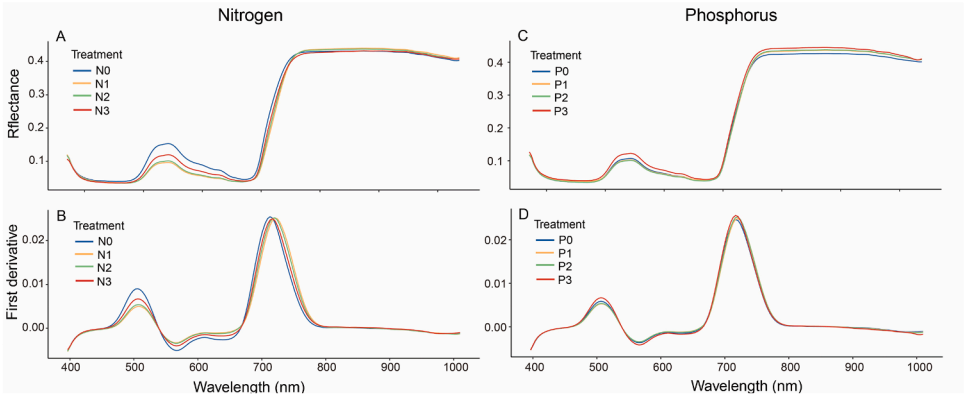


Fig. 4. Original spectra (A and C) and spectra after FD transformation (B and D) under different N and P treatments. N treatment: N_0 (0 g·pot^{-1}), N_1 (0.4 g·pot^{-1}), N_2 (0.8 g·pot^{-1}), and N_3 (1.2 g·pot^{-1}). P treatment: P_0 (0 g·pot^{-1}), P_1 (0.15 g·pot^{-1}), P_2 (0.3 g·pot^{-1}), and P_3 (0.45 g·pot^{-1}).

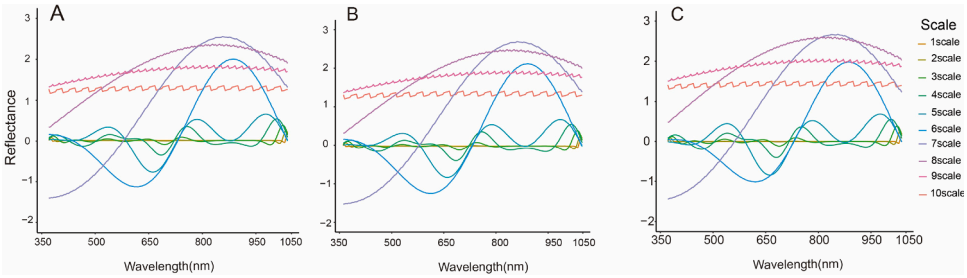


Fig. 5. Changes in the wavelet transform reflectance of leaf spectra under different nutritional statuses: (A) N treatment, (B) P treatment, and (C) control.

under the control ((Fig. 5C) at mid-range scales, particularly at scales [6] and [7], highlighting the transformative effect of these treatments on the spectral characteristics.

3.3. Correlation analysis

3.3.1. Correlation between nutrient content and first derivative spectra

To use the leaf spectra for predicting nutrient content, we conducted a correlation analysis between the FD spectra and nutrient content. The results revealed that more wavebands exhibited highly significant correlations after the raw spectra were processed via FD transformation (Fig. 6). For the N level, the highest correlation coefficient (r) of the raw spectra was found near 569 nm (–0.68) (Fig. 6A). In the FD spectra, the highest r (0.71) was observed near 782 nm (Fig. 6B). The correlation remained high ($r > 0.65$) in the range of 599–660 nm. For the P level, the highest r of the raw spectra was found near 652 nm (–0.30) (Fig. 6C). In the FD spectra, the highest r was found near 458 nm (–0.35). The correlation coefficient remained high (close to 0.3) in the range of 652–674 nm (Fig. 6D). Therefore, the FD spectra were selected as the independent variables for the prediction model based on their strong correlation with nutrient content.

3.3.2. Correlation between nutrient content and CWT

The correlation between hydrangea leaf spectra and nutrient contents after CWT is shown in Fig. 7. The r values were considerably greater for LNC than for LPC. LNC-sensitive bands ($P < 0.05$) were predominantly found at scales [5, 6, 7], whereas LPC-sensitive bands ($P < 0.05$) were concentrated at scales [6, 7]. Specifically, at the N level, a wavelet feature at 462 nm and scale [6] showed the strongest negative correlation (r= –0.90). At the P level, the strongest negative correlation (r= –0.87) was observed at 809 nm and scale [1].

These findings suggested that compared to the FD transform, CWT

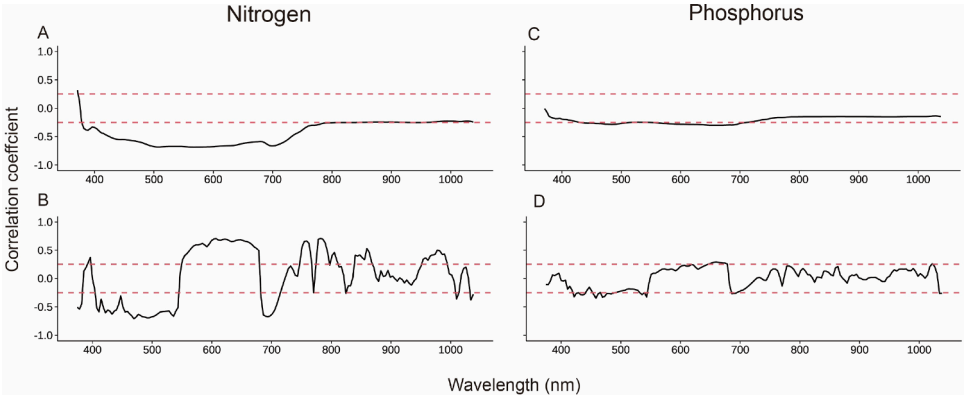


Fig. 6. Correlation analysis between different spectral data and the contents of N and P. Original spectra (A and C) and the first derivative spectra (B and D). The two red dashed lines below and above indicate a significance level of 0.05.

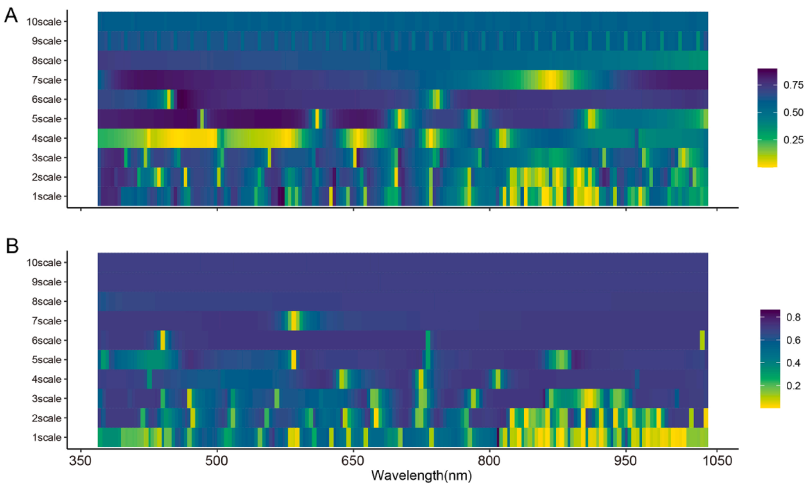


Fig. 7. Absolute values of correlation coefficients of different wavelet coefficients with leaf nutrient contents: (A) N treatment; (B) P treatment.

effectively increased the correlation values. The magnitude of the correlation coefficients and the number of sensitive bands significantly increased at scale [6] under the N and P treatments, highlighting that the wavelet transform was effective at this scale for analysing the original spectrum.

3.4. Single leaf estimation models based on hyperspectral images

Models were developed to predict the LNC and LPC using full spectra processed by FD transformation; three machine-learning techniques were used to construct the models. The scatter plots of the predicted

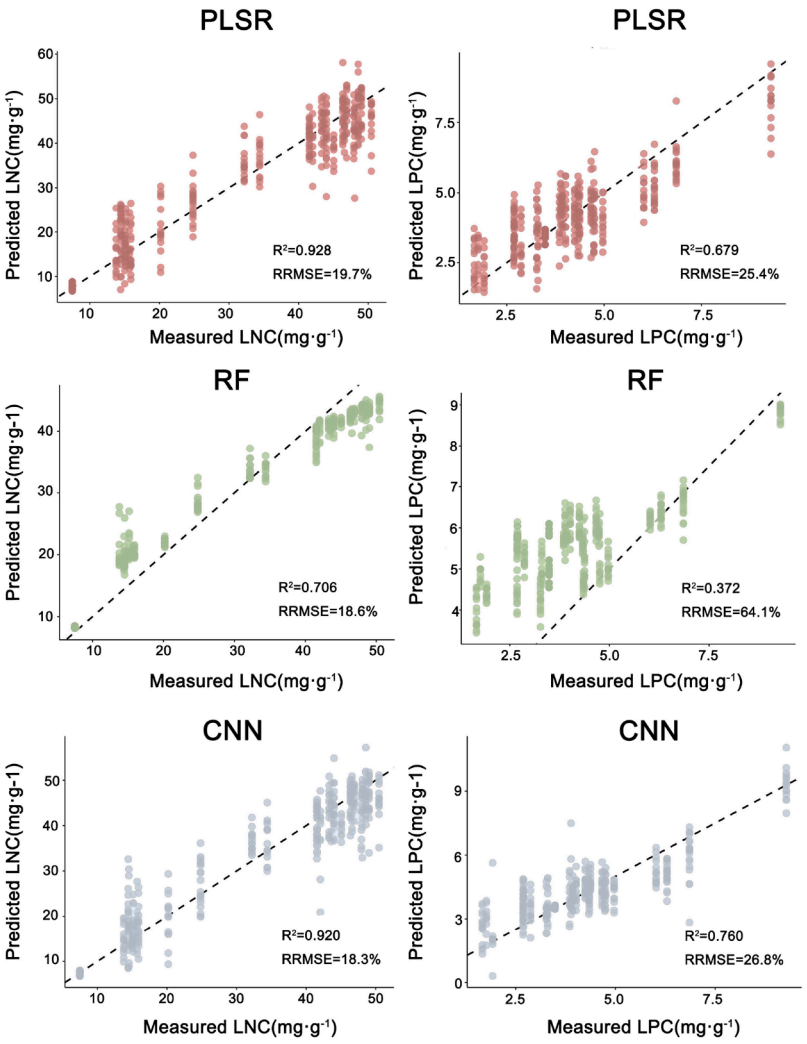


Fig. 8. Relationships between the predicted and measured variables (LNC and LPC) using three models. The dashed line corresponds to a 1:1 linear fit.

versus measured LNC and LPC values are shown in Fig. 8, with regression plots illustrating the relationship between the predicted and actual measurements. The PLSR model used 50 latent variables to capture the relationships effectively between the predictor and response variables.

To assess model convergence and performance, Fig. 9 shows the loss during training and validation for the CNN model, with LNC prediction used as an example. The training loss (red) and validation loss (black) were plotted as functions of the number of epochs. A small difference between the training and validation losses indicated effective model training, which helped avoid overfitting. Optimal results were achieved at about 100 epochs.

For LNC predictions, the models achieved R^2 values ranging from 0.706 to 0.928, and the RRMSE values ranged from 18.3 % to 19.7 %. The PLSR model had the highest R^2 of 0.928, followed by the CNN model (R^2= 0.920) and the RF model (R^2= 0.706). The CNN model exhibited the lowest RRMSE at 18.3 %. The LPC predictions performed worse than the LNC predictions, with R^2 values ranging from 0.372 to 0.760. The PLSR model for LPC had the lowest RRMSE of 25.4 %. Overall, the results showed that the RF model performed the poorest, whereas the PLSR model performed well, similar to the CNN model.

To assess whether the wavelet transform could further improve prediction accuracy, models incorporating CWT were developed using PLSR. Correlation coefficients (r values) across various scales were determined to create a correlation scalogram. To reduce redundancy, the r values obtained were sorted in descending order, retaining only the top 5 % as significant wavelet feature regions and yielding 88 key features. All selected features showed strong correlations with LNC and LPC ($P < 0.05$). The threshold for the top 5 % absolute r value was 0.823 for LNC and 0.684 for LPC. Model performance improvements were notable (Fig. 10). For LNC prediction, the model achieved an R^2 of 0.947 and an RRMSE of 9.6 %, whereas for LPC prediction, the R^2 was 0.827 and the RRMSE was 10.6 %.

The contributions of the top 10 sensitive wavelet features of the PLSR model are shown in Fig. 11. For N, the feature importance of 550 nm was the highest (0.06), representing 6 % of the contribution. For P, the feature importance of 1030 nm was the highest (0.08), representing 8 % of the contribution.

3.5. Discrimination model validation

To illustrate how the hydrangea leaves were classified into different nutrient states, confusion matrices were created using the PNN model. The N and P levels were categorised into three classes, i.e., deficiency, medium, and surplus, based on the measured LNC, LPC, biomass, and plant growth state. The leaf samples were marked as deficient (N_0), medium (N_1–N_2), or surplus (N_3) for the N level and as deficient (P_0), medium (P_1–P_2), or surplus (P_3) for the P level.

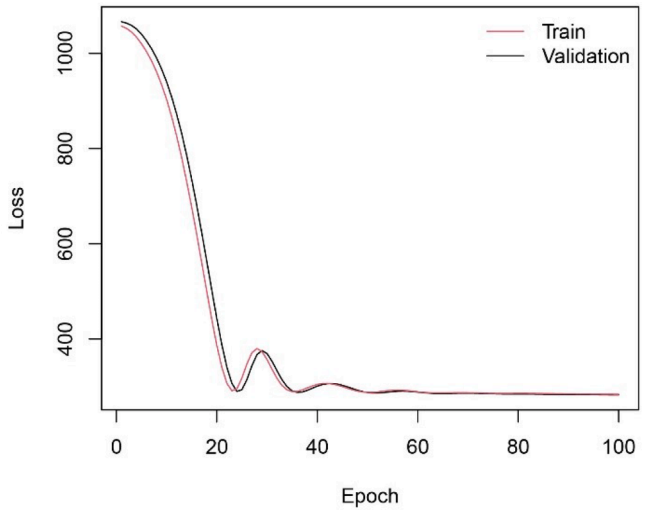


Fig. 9. CNN training loss curves.

Table 2 summarises the confusion matrices, overall accuracy (OA), and kappa values for different nutrient states. All categories had prediction accuracies higher than 95 % in all classes. The N classification of the model achieved higher OA and kappa values than the P classification, with OA of 99.8 % and 99.4 %, respectively, and kappa values of 0.997 and 0.989, respectively. N deficiency had the highest prediction accuracy at 100 %, while the performance of the model in predicting excessive N had the lowest prediction accuracy at 97.8 %. The results confirmed that with the addition of N, the N status was more difficult to recognise via spectral reflectance.

There were some differences in the classification of P. The highest percentage of correctly classified cases was observed under P surplus (100 %). This may be due to excessive P causing stress, as indicated by the lowest biomass and highest spectral reflectance under the different P treatments. In the other categories, the accuracy for the medium P level was lower than that for P deficiency. This could be explained by the fact that the medium P level is closer to the high and low levels, making it more difficult to distinguish than the other two classes.

4. Discussion

Studies have shown that spectral reflectance can indicate plant nutrient conditions. The visible (VIS, 400–760 nm) and near-infrared (NIR, 760–900 nm) spectra are often used to distinguish between N and P. As reported by Zhang et al. (2023), our study confirmed that N and P deficiencies increased reflectance in the VIS and decreased reflectance in the NIR. Other studies have reported that the red-edge region is important for distinguishing N and P levels (Mutanga and Kumar, 2007; Cho and Skidmore, 2006). We also found that the red-edge position for N was around 710 nm, which is strongly linked to chlorophyll (Abdel-Rahman et al., 2010). For the P level, the red-edge positions were identified at 713 nm and 728 nm, which matched the findings reported by Guo et al. (2018). The relationship between the LPC and red-edge wavelengths was also influenced by chlorophyll. When LPC is low or absent, the chlorophyll content decreases, and the anthocyanin content increases, which can cause the leaf surface to become purple (Li et al., 2016). These findings suggest that changes in red-edge wavelengths are associated with the chlorophyll content and nutrient statuses.

We also compared the results of nutrient content with those of two spectral transformations: FD and CWT. After FD treatment, correlation analysis revealed that the wavelengths most strongly correlated with LNC (599–660 nm) and LPC (653–674 nm) were located primarily in the red-edge region. The highest correlation for LNC was found at 782 nm (r= 0.71), whereas 458 nm had a negative correlation with LPC (r= –0.35). These findings highlight the importance of red and blue regions for nutrient prediction, with 782 nm being closely associated with the absorption of chlorophyll a (Osborne et al., 2002) and 458 nm being close to the chlorophyll b absorption characteristic (Curran, 1989). Compared to FD, CWT provided even more effective dimensionality reduction, denoising, and enhanced correlation with N and P levels (Cheng et al., 2012), achieving correlations of r= –0.90 and r= –0.87, respectively. These results suggest that CWT can effectively extract more reliable spectral features, making it an ideal preprocessing technique for hyperspectral data analysis.

While comparing the PLSR, RF, and CNN algorithms for estimating LNC and LPC, we found that RF was the least effective, whereas PLSR showed performance similar to that of CNN. After using CWT, the PLSR models outperformed the other algorithms, achieving improved prediction accuracy while reducing the number of input variables. This matched the findings of Yang et al. (2021), who emphasised the robustness and stability of PLSR. For LNC and LPC estimation, the VIS and NIR regions were identified as important, with sensitive bands at

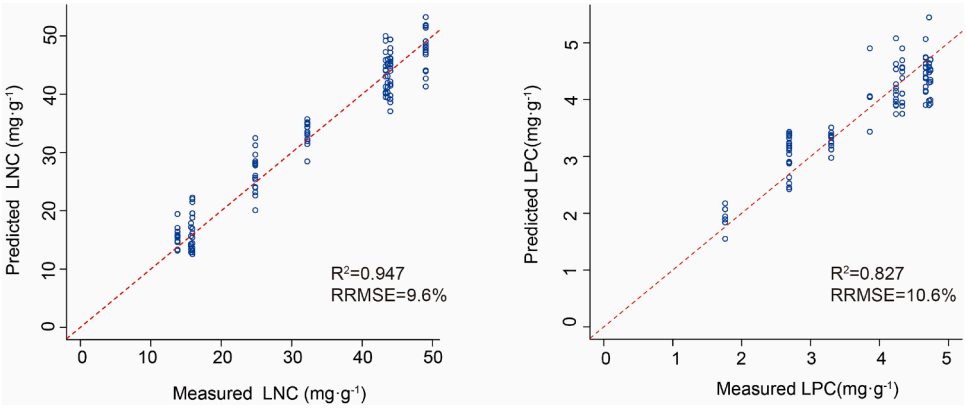


Fig. 10. Relationships between the predicted and observed variables (LNC and LPC) using the PLSR model. The dashed line corresponds to a 1:1 linear fit.

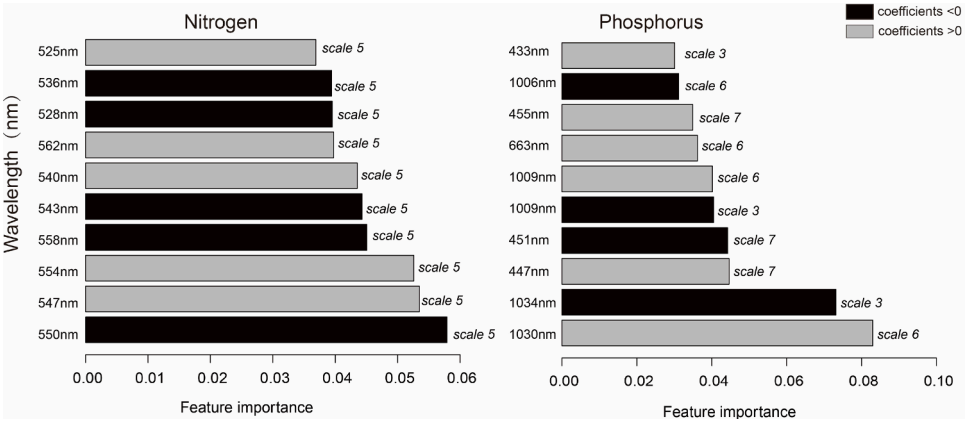


Fig. 11. The importance of the top 10 features of PLSR models based on multiple wavelet features.

Table 2
Confusion matrices, OA, and kappa coefficients of the PNN model under different nutrient statuses.

Model (PNN)		Predicted Class			Precision	Recall	OA	Kappa
		Deficiency	Medium	Surplus				
Actual	Deficiency	804	0	0	100 %	100 %	99.8 %	0.997
Class	Medium	0	630	0	99.5 %	100 %		
(N)	Surplus	0	3	132	100 %	97.8 %		
Actual	Deficiency	801	3	0	99.3 %	99.6 %	99.4 %	0.989
Class	Medium	6	687	0	99.6 %	99.1 %		
(P)	Surplus	0	0	75	100 %	100 %		

about 550 nm at scale [5] for LNC and 1030 nm at scale [6] for LPC, confirming the relevance of these wavelengths for plant nutrient prediction (Zhang et al., 2023).

Overall, the prediction of LNC was better than that of LPC in all models, probably because of a weaker correlation of spectral reflectance with LPC within the 400–1000 nm range. This indicates that detecting P deficiencies is challenging, as they rarely show visible symptoms like chlorosis at early stages. In contrast, N deficiencies are typically easier to identify (Pinit et al., 2022). This finding aligns with the understanding that P shortage not only reduces LNC but also hinders ATP synthase activity, disrupting the electron transport chain and leading to transient changes in chlorophyll a fluorescence (Carstensen et al., 2018). Our discrimination model, which uses PNN algorithms, confirmed these observations, showing a slightly lower overall accuracy in identifying P than in identifying N.

By combining hyperspectral spectroscopy with CWT, we developed a reliable, efficient method for estimating plant nutrients, particularly for hydrangea cultivation. Compared to other studies, our study more effectively addressed the challenge of weak spectral correlations with LPC. The models developed in this study, particularly the CWT-PLSR approach, can be used in future precision agriculture applications, offering improved nutrient management and potentially extending to other crops.

5. Conclusion

In this study, we successfully used hyperspectral spectroscopy to characterise N and P in hydrangea. By applying spectral preprocessing techniques such as FD and CWT, we extracted more detailed spectral information, improving the ability to characterise plant nutrition. By comparing different preprocessing methods, we found that CWT outperforms FD in terms of spectral data quality and prediction accuracy. The PLSR model, coupled with CWT, provided the best predictions for LNC and LPC, achieving high accuracy (R^2= 0.947, RRMSE= 9.6 % for LNC; R^2= 0.827, RRMSE= 10.6 % for LPC). This method significantly decreased the number of input variables without decreasing prediction

performance. Additionally, the classification model (PNN) achieved over 95 % accuracy in distinguishing between N and P nutrient statuses. The findings of this study can provide a robust foundation for the use of hyperspectral spectroscopy in precision nutrient management for the cultivation of hydrangea and other crops. The CWT-PLSR model developed in this study offers a promising approach to improve crop management through more efficient and accurate nutrient estimation; future studies should focus on refining these models for field conditions.

CRediT authorship contribution statement

Jun Yang: Writing – original draft, Formal analysis, Data curation. **Guochun Shen:** Writing – review & editing, Methodology, Formal analysis. **Jun Qin:** Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was funded by Science and Technology Research Project of Shanghai Greening and City Appearance Administration in 2024 (G242423), and Shanghai Agriculture Applied Technology Development Program, China (Grant No. TS20210101).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.scienta.2025.114144.

Data availability

The data used to support the findings of this study are available from the corresponding author upon request.

References

Abdel-Rahman, E.M., Ahmed, F.B., van den Berg, M., 2010. Estimation of sugarcane leaf nitrogen concentration using in situ spectroscopy. Int. J. Appl. Earth Obs. Geoinf. 12, S52–S57. https://doi.org/10.1016/j.jag.2009.11.003.

Abdulazeez, A.M., Zeebaree, D.Q., Zebari, D.A., Hameed, T.H., 2021. Leaf identification based on shape, color, texture and vines using probabilistic neural network. Comp. y Sist. 25. https://doi.org/10.13053/cys-25-3-3470.

Amiri, M., Zinatloo-Ajabshir, S., Ahmadi-Zeidabadi, M., Sharifianjazi, F., 2025. Innovative ultrasound assisted synthesis of sponge like cerium dioxide nanostructure using Rosa Damascena extract and its efficient performance for cancer therapy. Sci. Rep. 15, 933. https://doi.org/10.1038/s41598-025-85137-5.

Anderson, N., Weiland, J., Pharis, J., Gagné, W., Janiga, E., Rosenow, M.J., 2009. Comparative forcing of *Hydrangea macrophylla* 'Bailer' as a florist's hydrangea. Sci. Hortic. 122, 221–226. https://doi.org/10.1016/j.scienta.2009.05.003.

Azlah, M.A.F., Chua, L.S., Rahmad, F.R., Abdullah, F.I., Wan Alwi, S.R., 2019. Review on techniques for plant leaf classification and recognition. Computers 8, 77. https://doi.org/10.3390/computers8040077.

Ball, K.R., Liu, H., Brien, C., Berger, B., Power, S.A., Pendall, E., 2022. Hyperspectral imaging predicts yield and nitrogen content in grass–legume polycultures. Precis. Agric. 23, 2270–2288. https://doi.org/10.1007/s11119-022-09920-4.

Baranowski, P., Jedryczka, M., Mazurek, W., Babula-Skowronska, D., Siedliska, A., Kaczmarek, J., 2015. Hyperspectral and thermal imaging of oilseed rape (*Brassica napus*) response to fungal species of the genus Alternaria. PLoS One 10, e0122913. https://doi.org/10.1371/journal.pone.0122913.

Carstensen, A., Herdean, A., Schmidt, S.B., Sharma, A., Spetea, C., Pribil, M., Husted, S., 2018. The impacts of phosphorus deficiency on the photosynthetic electron transport chain. Plant Physiol. 177, 271–284. https://doi.org/10.1104/pp.17.01624.

Cheng, T., Rivard, B., Sánchez-Azofeifa, A., 2011. Spectroscopic determination of leaf water content using continuous wavelet analysis. Remote Sens. Environ. 115, 659–670. https://doi.org/10.1016/j.rse.2010.11.001.

Cheng, T., Rivard, B., Sánchez-Azofeifa, A.G., Féret, J.-B., Jacquemoud, S., Ustin, S.L., 2012. Predicting leaf gravimetric water content from foliar reflectance across a range of plant species using continuous wavelet analysis. J. Plant Physiol. 169, 1134–1142. https://doi.org/10.1016/j.jplph.2012.04.006.

Cheng, X., Feng, Y., Guo, A., Huang, W., Cai, Z., Dong, Y., Guo, J., Qian, B., Hao, Z., Chen, G., Liu, Y., 2024. Detection of rubber tree powdery mildew from leaf level hyperspectral data using continuous wavelet transform and machine learning. Remote Sens. 16, 105. https://doi.org/10.3390/rs16010105.

Cho, M.A., Skidmore, A.K., 2006. A new technique for extracting the red edge position from hyperspectral data: The linear extrapolation method. Remote Sens. Environ. 101, 181–193. https://doi.org/10.1016/j.rse.2005.12.011.

Curran, P.J., 1989. Remote sensing of foliar chemistry. Remote Sens. Environ. 30, 271–278. https://doi.org/10.1016/0034-4257(89)90069-2.

Ghorbani, M.A., Shamshirband, S., Zare Haghi, D., Azani, A., Bonakdari, H., Ebtehaj, I., 2017. Application of firefly algorithm-based support vector machines for prediction of field capacity and permanent wilting point. Soil Tillage Res. 172, 32–38. https://doi.org/10.1016/j.still.2017.04.009.

Gu, C., Ji, S., Xi, X., Zhang, Z., Hong, Q., Huo, Z., Li, W., Mao, W., Zhao, H., Zhang, R., Li, B., Tan, C., 2022. Rice yield estimation based on continuous wavelet transform with multiple growth periods. Front. Plant Sci. 13. https://doi.org/10.3389/fpls.2022.931789.

Guo, D., Juan, J., Chang, L., Zhang, J., Huang, D., 2017. Discrimination of plant root zone water status in greenhouse production based on phenotyping and machine learning techniques. Sci. Rep. 7, 8303. https://doi.org/10.1038/s41598-017-08235-z.

Guo, P.-T., Shi, Z., Li, M.-F., Luo, W., Cha, Z.-Z., 2018. A robust method to estimate foliar phosphorus of rubber trees with hyperspectral reflectance. Ind. Crops. Prod. 126, 1–12. https://doi.org/10.1016/j.indcrop.2018.09.055.

Hanafi, M.R., Rahimpour, H., Zinatloo-Ajabshir, S., Moodi, F., Fahmi, A., 2024. Performance enhancement, life cycle assessment, and feature analysis of wheat starch-based NaCl-binder as a sustainable alternative to OPC mortar. Results Eng. 24, 103281. https://doi.org/10.1016/j.rineng.2024.103281.

Imani, S., Soltani-Jigheh, H., Kafil, H.S., Zinatloo-Ajabshir, S., Fahmi, A., 2025. Environmental sustainability in combating wind erosion: chemical stabilization of sand dunes using acidic mulching. Results Eng. 25, 103758. https://doi.org/10.1016/j.rineng.2024.103758.

Li, D., Wang, X., Zheng, H., Zhou, K., Yao, X., Tian, Y., Zhu, Y., Cao, W., Cheng, T., 2018. Estimation of area- and mass-based leaf nitrogen contents of wheat and rice crops from water-removed spectra using continuous wavelet analysis. Plant Methods 14, 76. https://doi.org/10.1186/s13007-018-0344-1.

Li, M.-F., Tang, X.-P., Wu, W., Liu, H.-B., 2013. General models for estimating daily global solar radiation for different solar radiation zones in mainland China. Energy Convers. Manage. 70, 139–148. https://doi.org/10.1016/j.enconman.2013.03.004.

Li, L., Wang, S., Ren, T., Ma, Y., Wei, Q., Gao, W., Lu, J., 2016. Evaluating models of leaf phosphorus content of winter oilseed rape based on hyperspectral data. Trans. Chin. Soc. Agric. Eng. 32, 209–218. https://doi.org/10.11975/j.issn.1002-6819.2016.14.028.

Li, T., Zhu, Z., Cui, J., Chen, J., Shi, X., Zhao, X., Jiang, M., Zhang, Y., Wang, W., Wang, H., 2021. Monitoring of leaf nitrogen content of winter wheat using multi-angle hyperspectral data. Int. J. Remote Sens. 42, 4672–4692. https://doi.org/10.1080/01431161.2021.1899333.

Liu, Z.-Y., Qi, J.-G., Wang, N.-N., Zhu, Z.-R., Luo, J., Liu, L.-J., Tang, J., Cheng, J.-A., 2018. Hyperspectral discrimination of foliar biotic damages in rice using principal component analysis and probabilistic neural network. Precision Agric. 19, 973–991. https://doi.org/10.1007/s11119-018-9567-4.

Lou, S., Hu, R.-Q., Liu, Y., Zhang, W., Yang, S.-Q., 2022. The formulation of irrigation and nitrogen application strategies under multi-dimensional soil fertility targets based on preference neural network. Sci. Rep. 12, 20918. https://doi.org/10.1038/s41598-022-25133-1.

Mutanga, O., Kumar, L., 2007. Estimating and mapping grass phosphorus concentration in an African savanna using hyperspectral image data. Int. J. Remote Sens. 28, 4897–4911. https://doi.org/10.1080/01431160701253253.

Osborne, S.L., Schepers, J.S., Francis, D.D., Schlemmer, M.R., 2002. Detection of phosphorus and nitrogen deficiencies in corn using spectral radiance measurements. Agron. J. 94, 1215–1221. https://doi.org/10.2134/agronj2002.1215.

Pinit, S., Ruengchaijatuporn, N., Sriswasdi, S., Buaboocha, T., Chadchawan, S., Chaiwanon, J., 2022. Hyperspectral and genome-wide association analyses of leaf phosphorus status in local Thai indica rice. PLoS One 17, e0267304. https://doi.org/10.1371/journal.pone.0267304.

Pourdarbani, R., Sabzi, S., Rohban, M.H., García-Mateos, G., Arribas, J.I., 2021. Nondestructive nitrogen content estimation in tomato plant leaves by Vis-NIR hyperspectral imaging and regression data models. Appl. Opt. 60, 9560. https://doi.org/10.1364/AO.431886.

Shreckhise, J.H., Owen, J.S., Niemiera, A.X., 2019. Growth response of *Hydrangea macrophylla* and *Ilex crenata* cultivars to low-phosphorus controlled-release fertilizers. Sci. Hortic. 246, 578–588. https://doi.org/10.1016/j.scienta.2018.10.045.

Sun, D., Cen, H., Weng, H., Wan, L., Abdalla, A., El-Manawy, A.I., Zhu, Y., Zhao, N., Fu, H., Tang, J., Li, X., Zheng, H., Shu, Q., Liu, F., He, Y., 2019. Using hyperspectral analysis as a potential high throughput phenotyping tool in GWAS for protein content of rice quality. Plant Methods 15, 54. https://doi.org/10.1186/s13007-019-0432-x.

USDA, N., 2019. Census of Horticultural Specialties.

Velumani, K., Lopez-Lozano, R., Madec, S., Guo, W., Gillet, J., Comar, A., Baret, F., 2021. Estimates of maize plant density from UAV RGB images using Faster-RCNN detection model: Impact of the spatial resolution. Plant Phenomics 2021, 9824843. https://doi.org/10.34133/2021/9824843.

Wang, J., Tian, T., Wang, H., Cui, J., Zhu, Y., Zhang, W., Tong, X., Zhou, T., Yang, Z., Sun, J., 2021. Estimating cotton leaf nitrogen by combining the bands sensitive to

nitrogen concentration and oxidase activities using hyperspectral imaging. Comput. Electron. Agric. 189, 106390. https://doi.org/10.1016/j.compag.2021.106390.

Wang, X., Hu, Y., Liaquat, F., Zhang, X., Ye, K., Qin, J., Liu, Q., 2022. Effects of nitrogen exponential fertilization on growth and nutrient concentration of *Hydrangea macrophylla* seedlings. Phyton 91, 395–407. https://doi.org/10.32604/phyton.2022.017981.

Wang, Z., Chen, J., Fan, Y., Cheng, Y., Wu, X., Zhang, J., Wang, B., Wang, X., Yong, T., Liu, W., Liu, J., Du, J., Yang, W., Yang, F., 2020. Evaluating photosynthetic pigment contents of maize using UVE-PLS based on continuous wavelet transform. Comput. Electron. Agric. 169, 105160. https://doi.org/10.1016/j.compag.2019.105160.

Wang, Y., Xiong, F., Zhang, Y., Wang, S., Yuan, Y., Lu, C., Nie, J., Nan, T., Yang, B., Huang, L., Yang, J., 2023. Application of hyperspectral imaging assisted with integrated deep learning approaches in identifying geographical origins and predicting nutrient contents of Coix seeds. Food Chem. 404, 134503. https://doi.org/10.1016/j.foodchem.2022.134503.

Watt, M.S., Buddenbaum, H., Leonardo, E.M.C., Estarija, H.J.C., Bown, H.E., Gomez-Gallego, M., Hartley, R., Massam, P., Wright, L., Zarco-Tejada, P.J., 2020. Using hyperspectral plant traits linked to photosynthetic efficiency to assess N and P partition. ISPRS J. Photogram. Remote Sens. 169, 406–420. https://doi.org/10.1016/j.isprsjprs.2020.09.006.

Yang, H., Li, F., Hu, Y., Yu, K., 2021. Hyperspectral indices optimization algorithms for estimating canopy nitrogen concentration in potato (Solanum tuberosum L.). Int. J. Appl. Earth Observ. Geoinfo. 102, 102416. https://doi.org/10.1016/j.jag.2021.102416.

Zhang, Y., Wang, Teng, Li, Z., Wang, Tianli, Cao, N., 2023. Based on machine learning algorithms for estimating leaf phosphorus concentration of rice using optimized spectral indices and continuous wavelet transform. Front. Plant Sci. 14. https://doi.org/10.3389/fpls.2023.118591.

Zhuang, T., Zhang, Y., Li, D., Schmidhalter, U., Ata-UI-Karim, S.T., Cheng, T., Liu, X., Tian, Y., Zhu, Y., Cao, W., Cao, Q., 2023. Coupling continuous wavelet transform with machine learning to improve water status prediction in winter wheat. Precision Agric. 24, 2171–2199. https://doi.org/10.1007/s11119-023-10036-6.

Zinatloo-Ajabshir, S., Yousefi, A., Jekle, M., Sharifianjazi, F., 2025. Ingenious wheat starch/lepidium perfoliatum seed mucilage hybrid composite films: synthesis, incorporating nanostructured Dy2Ce2O7 synthesized via an ultrasound-assisted approach and characterization. Carbohydrate Polym. Technol. Appl. 9, 100657. https://doi.org/10.1016/j.carpta.2024.100657.

Zonarsaghar, A., Mousavi-Kamazani, M., Zinatloo-Ajabshir, S., 2022. Sonochemical synthesis of CeVO4 nanoparticles for electrochemical hydrogen storage. Int. J. Hydrogen Energ. 47, 5403–5417. https://doi.org/10.1016/j.ijhydene.2021.11.183.