Journal of Forestry Research (2024) 35:116
https://doi.org/10.1007/s11676-024-01769-9

ORIGINAL PAPER

Deciphering nitrogen concentrations in *Metasequoia glyptostroboides*: a novel approach using RGB images and machine learning

Cong Ma[1,2] · Ran Tong[1] · Nianfu Zhu[1] · Wenwen Yuan[1] · Yanji Li[1] · G. Geoff Wang[3] · Tonggui Wu[1]

Received: 21 October 2023 / Accepted: 4 April 2024 / Published online: 3 August 2024
© The Author(s) 2024

Abstract Recent advances in spectral sensing techniques and machine learning (ML) methods have enabled the estimation of plant physiochemical traits. Nitrogen (N) is a primary limiting factor for terrestrial forest growth, but traditional methods for N determination are labor-intensive, time-consuming, and destructive. In this study, we present a rapid, non-destructive method to predict leaf N concentration (LNC) in *Metasequoia glyptostroboides* plantations under N and phosphorus (P) fertilization using ML techniques and unmanned aerial vehicle (UAV)- based RGB (red, green, blue) images. Nine spectral vegetation indices (VIs) were extracted from the RGB images. The spectral reflectance and VIs were used as input features to construct models for estimating LNC based on support vector machine, random forest (RF), and multiple linear regression, gradient boosting regression and classification and regression trees (CART). The results show that RF is the best fitting model for estimating LNC with a coefficient of determination (R^2) of 0.73. Using this model, we evaluated the effects of N and P treatments on LNC and found a significant increase with N and a decrease with P. Height, diameter at breast height (DBH), and crown width of all *M. glyptostroboides* were analyzed by Pearson correlation with the predicted LNC. DBH was significantly correlated with LNC under N treatment. Our results highlight the potential of combining UAV RGB images with an ML algorithm as an efficient, scalable, and cost-effective method for LNC quantification. Future research can extend this approach to different tree species and different plant traits, paving the way for large-scale, time-efficient plant growth monitoring.

Keywords RGB images · Random forest · LNC · N and P addition · *Metasequoia*

Introduction

Leaves, as the organs of photosynthesis and carbon assimilation, often have priority in nutrient allocation under environmental changes (Urban et al. 2021). Leaf nitrogen concentration (LNC) directly determines the productivity of plants, and is also the result of nutrient content adjustment by plants in response to environmental changes (Tang et al. 2018). As the most essential nutrient in plants, N is involved in almost all physiological processes and plays a very important role (Tong et al. 2022). Monitoring N levels can help to understand the metabolic processes and health status of plants at an early stage (Wen et al. 2020). In fact, leaf N serves as a robust indicator of both plant growth and its N nutritional status, profoundly influencing plant development (Xiao et al. 2022a; Wu et al. 2023).

Project funding: This research was supported by the "Pioneer" and "Leading Goose" R&D Program of Zhejiang (2022C02053), and National Natural Science Foundation of China (NSFC) (32201632).

The online version is available at https://link.springer.com/.

Corresponding editor: Lei Yu.

✉ Tonggui Wu
wutonggui@caf.ac.cn

[1] East China Coastal Forest Ecosystem Long-Term Research Station, Research Institute of Subtropical Forestry, Chinese Academy of Forestry, Hangzhou 311400, People's Republic of China

[2] Nanjing Forestry University, Nanjing 210037, People's Republic of China

[3] Department of Forestry and Environmental Conservation, Clemson University, Clemson, SC 29634-0317, USA

Typically, N is determined by the Kjeldahl method, which is expensive, complex and time-consuming, and prevents comprehensive, large-scale, rapid nutrient diagnosis (Li et al. 2022). Therefore, non-destructive techniques can be used that are rapid and inexpensive (Banerjee et al. 2020). Unmanned aerial vehicle (UAV) monitoring technology offers numerous advantages, including flexibility, cost-effectiveness, and high resolution. As UAV remote sensing technology has evolved in recent years, the collection of vegetation spectral information via spectral cameras mounted on drone platforms has emerged as a significant advancement in the field of precision forestry. This cutting-edge technology enables the acquisition of vital crop growth information through the analysis of vegetation spectral data. In particular, it provides rapid, non-invasive and accurate results, revolutionizing traditional practices (Ji et al. 2021). Therefore, using UAV remote sensing to replace traditional collection methods can effectively obtain growth information in a timely and effective manner. Over the past decade, the number of studies using spectral data to ascertain leaf N in agricultural crops has increased. While there has been limited research on determining LNC in trees, RGB cameras, which are more cost-effective than hyperspectral and multispectral cameras, have gained importance. Vital vegetation indices (Vis) derived from RGB images are critical for estimating plant growth parameters (Yang et al. 2019). In a notable study, Kou et al. (2022) used a UAV digital camera to capture RGB images of cotton canopies at 20 m. By experimenting with two cotton varieties and six N gradients, they aimed to predict canopy N concentrations. They used linear regression, support vector machine, and one-dimensional convolutional neural network (CNN) regression models to accurately predict cotton canopy N levels. Impressively, the model based on the UAV RGB images proved to be highly predictive.

Machine learning (ML) models can quickly process, analyze and model image information through programming and automated processing, thereby analyzing a large amount of information contained in images (Jordan and Mitchell 2015). ML models have been established as effective tools for dealing with the nonlinear relationships between independent and dependent variables. They have become increasingly popular in research on smart forestry applications, including species, structure, and biomass monitoring (Torresan et al. 2019), foliar pest and disease identification (Zhang et al. 2018), and forest management (Liu et al. 2018). ML algorithms are increasingly being applied to crop growth monitoring and nutrient diagnosis (Ghosal et al. 2018; Zhang et al. 2016). The combination of UAV-acquired spectral images with ML models to predict LNC has been validated as feasible (Liu et al. 2016; Yamashita et al. 2020; Noguera et al. 2021). Several studies have combined spectral metrics with a variety of machine learning algorithms, such as SVM, RF, MLR, partial least squares regression (PLSR), GBR, and CART, to predict LNC in maize, citrus, and grapes, and compared the performance of several machine learning algorithms, finding that the random forest algorithm was more accurate than the other algorithms (Prado Osco et al. 2019; Barzin et al. 2021; Tang et al. 2022).

Nitrogen deposition has long been of interest as an important component in the N cycle of ecosystems. Due to global climate change and human activities, atmospheric N deposition has been increasing, and eastern China is one of the regions with high N deposition. In the eastern coastal region, high N deposition and high P leaching are important factors that alter soil nutrients and plant growth (Li et al. 2019). *Metasequoia glyptostroboides* Hu & Cheng, the dawn redwood, has become one of the pioneer afforestation species in the Eastern Coastal Protection Forest due to its excellent ability to prevent wind erosion, stabilize sand, thrive in barren soils, and resist waterlogging.

For *M. glyptostroboides*, few studies have evaluated LNC with remote sensing data, and no study has been found using spectral indices with ML models (Lin et al. 2014; Shen et al. 2020; Onishi and Ise 2021). The objectives of this study were to develop and compare different ML-based models to estimate the LNC of *M. glyptostroboides* using VIs and spectral bands acquired from a visible-light camera mounted on a UAV. Using the optimal model, the LNC was estimated across the experimental site. This approach allowed us to investigate the effect of N and P fertilization on the LNC and the correlation of height, DBH and crown width with LNC for each tree in the sample plot. This will provide theoretical support for non-destructive and rapid bulk estimation of *M. glyptostroboides* leaf stoichiometry and will contribute to the enhancement of the value of *M. glyptostroboides* plantations and the efficient exploitation of their potential.

Materials and methods

Experimental design

The experiment was carried out at the end of June and November 2022 at the Dongtai Forest Farm (121° 450 E, 33° 420 N) on the coast in the central part of Jiangsu Province. The forest farm is in the northern subtropical zone with an annual average temperature of 14.6 °C. The soil texture is sandy loam and the surface soil is alkaline. This research site is a pioneering area for *M. glyptostroboides* plantations along the coast of China. In 2014, N and P experiments were initiated in a 6-year 300 m × 100 m M. *glyptostroboides* plantation. Ten 20 m × 30 m subplots were randomly selected in each block for N and P fertilization, with five subplots for the N treatment, and five for the P treatment. Each plot was separated by a buffer zone.

N treatments were zero (N0), 0.8 (N0.8), 2.4 (N2.4), 4 (N4) and 4.8 (N4.8) mol m^{-2} in the form of urea [$CO(NH_2)_2$]; P treatments were zero (P0), 0.05 (P0.05), 0.2 (P0.2), 0.6 (P0.6) and 1.0 (P1) mol m^{-2} in the form of monocalcium phosphate [$Ca(H_2PO_4)_2$]. The fertilization gradient was based on a pretest and previous seedling pot studies (Wang et al. 2018). The fertilization applications were: the first between late March and early April, accounting for 60% of the total; the second was in mid-June, accounting for 40% of the total. The experimental design is shown in Fig. 1.

Sampling and data collection

To avoid errors in remote sensing data collection due to rainy or cloudy weather, data collection was carried out on clear days with less than 20% cloud cover. A Phantom 4 PRO quadcopter drone manufactured by DJI Innovations in Shenzhen, China and equipped with a visible light RGB camera was used to acquire RGB image data. Agisoft Photoscan Professional software was employed for image stitching and 3D reconstruction. The images were then imported into ENVI 5.3 (Exelis Visual Information Solutions, Inc., USA) for data extraction and analysis. ENVI 5.3 software segmented the image data collected at each sampling. The region of interest (ROI) of the target image was manually marked and a mask was applied to output the image data of the target image. Due to slight errors in the selection of the ROI, soil pixels between rows were marked as plant areas, which would affect the detection accuracy of the model. Therefore, this study used a threshold segmentation method to remove soil and other noise from the target area (Wang et al. 2013), leaving only the *M. glyptostroboide* canopy pixels.

At the same time as the UAV collection, an average of three trees in each plot were selected for leaf collection, sampling the upper and outer fully expanded leaves of the canopy. After drying at 75 °C for 72 h, the leaves were ground using a mechanical grinder. The N concentration was determined using the Kjeldahl method (Wen et al. 2021).

DBH was measured for all trees in each plot. The height of a single tree was extracted using point cloud data generated by unmanned aerial surveys, and the crown width of single trees was extracted from orthophotos acquired by the UAV.

Data analysis

Reflectance data from three visible light spectral bands obtained by the UAV were used to calculate nine VIs (Table 1). Subsequently, models were created using the

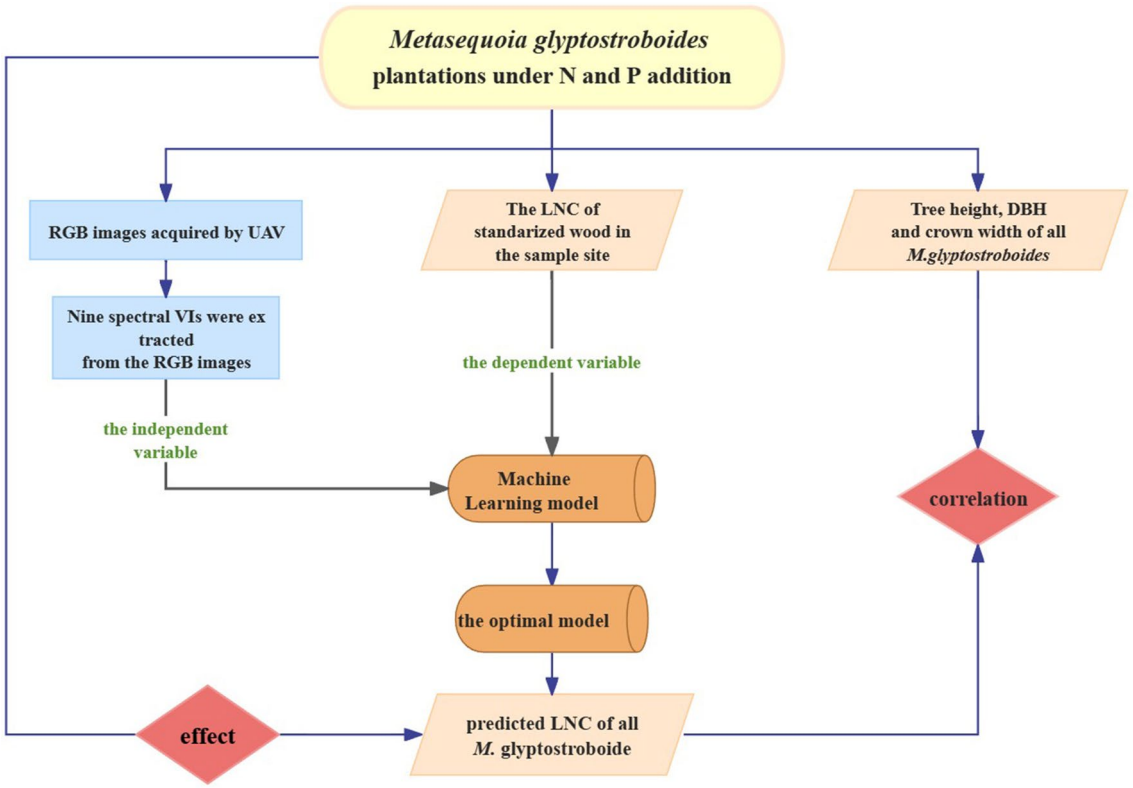


Fig. 1 Experimental flowchart

Table 1 Visible-band vegetation indexes

Vegetation index	Variable	Vegetation index	Variable
R, G, B	V1, V2, V3	G/(R+B)	V8
R/(R+G+B)	V4	B/(R+G)	V9
G/(R+G+B)	V5	(G-R)/(G+R)	V10
B/(R+G+B)	V6	GMR=G-R	V11
R/(G+B)	V7	GDR=G/R	V12

Table 2 The average of R^2 and *RMSE* for five machine learning models

ML algorithm	R^2 train	*RMSE* train	R^2 test	*RMSE* test
SVM	0.61	1.77	0.48	2.11
CART	0.67	1.62	0.47	2.16
GBR	0.81	1.26	0.46	2.13
RF	0.91	0.91	0.51	2.03
MLR	0.58	1.85	0.45	2.32

original reflectance data of the three bands and the VIs. The data obtained in June and in November 2022 were pooled and divided into training (80%) and test (20%) datasets based on different N and P fertilization gradients. In total, 180 data points were obtained. All graphs were generated using Graphpad Prism 9.3.0 (Swift 1997) and R Studio (Racine 2012).

Machine learning methods

Five ML methods were used to find an estimation algorithm to explain the relationship between the dependent variable (LNC) and the 12 independent variables (12 VIs) to estimate LNC. The ML methods included RF (Breiman 2001), GBR (Friedman 2002), SVM (Smola and Scholkopf 2004), CART (Bertsimas and Dunn 2017) and multiple linear regression (MLR) (McCann et al. 2010). All ML models were performed using R Studio (Racine 2012) and the "randomforest" package (Genuer and Poggi 2020), the "gbm" package (Ridgeway and Ridgeway 2004), the "e1071" package (Dimitriadou et al. 2009), the "rpart" package (Therneau et al. 2015), and the "lm" function (Elzhov et al. 2016) in R Studio.

In the GBR model, the "n.trees" parameter was set to 100, the "interaction.depth" parameter to 3, and the rest of the parameters were library defaults. In the random forest model, n.tree represents the number of decision trees, m.try the number of nodes in the decision tree. The choice of values for n.tree and m.try directly affects the final accuracy of the model. This study conducted several experiments and determined that the setting values for ntree and mtry should be 500 and 4, respectively. Kernel in the SVM model was set to "radial" and the rest of the parameters were library defaults. The library default values were adopted for the number and depth of trees, nodes, and leaves in CART models.

To evaluate the performance of each model as well as the relationship between the predicted and observed variables, R^2 and the root mean square error (*RMSE*) metrics were used.

These parameters were calculated as follows:

$$R^2 = \frac{\sum_{i=0}^{n}\left(X_i-\overline{X}\right)^2\left(Y_i-\overline{Y}\right)^2}{n\sum_{i=0}^{n}\left(X_i-\overline{X}\right)^2\sum_{i=0}^{n}\left(Y_i-\overline{Y}\right)^2} \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i-X_i)^2} \quad (2)$$

where X_i and Y_i are the predicted and measured N concentration, respectively, and X and Y their means, respectively. n is the number of samples.

Results

Model performances for LNC

We executed each of the five machine learning (ML) models 50 times, assessing their performance via both R^2 and *RMSE* metrics on the training and test datasets (Table 2). The RF model demonstrated superior performance, achieving an average R^2 of 0.51 and an *RMSE* of 2.03 on the test dataset. Consequently, we selected the RF model for subsequent prediction of leaf nitrogen content.

N estimation model and predict LNC

The R^2 of the RF model on the training data set ranged from 0.90 to 0.93, indicating a good fit. The *RMSE* ranged from 0.80 to 1.00, further demonstrating the excellent predictive ability of the model on the training data. However, when applied to the test dataset, the performance of the RF model was slightly lower, with R^2 ranging from 0.23 to 0.73 and *RMSE* from 1.46 to 2.83 (Fig. 2). Of all the iterations, the one that produced the highest R^2 on the test dataset was selected as the optimal model. This model achieved an R^2 of 0.73 and an RMSE of 1.66 on the test data set (Fig. 3b). The corresponding values for the training data set were an R^2 of 0.90 and an *RMSE* of 0.96 (Fig. 3a). The larger value of residual prediction deviation (RPD) indicates that the model is able to reliably predict its chemical composition (Wager

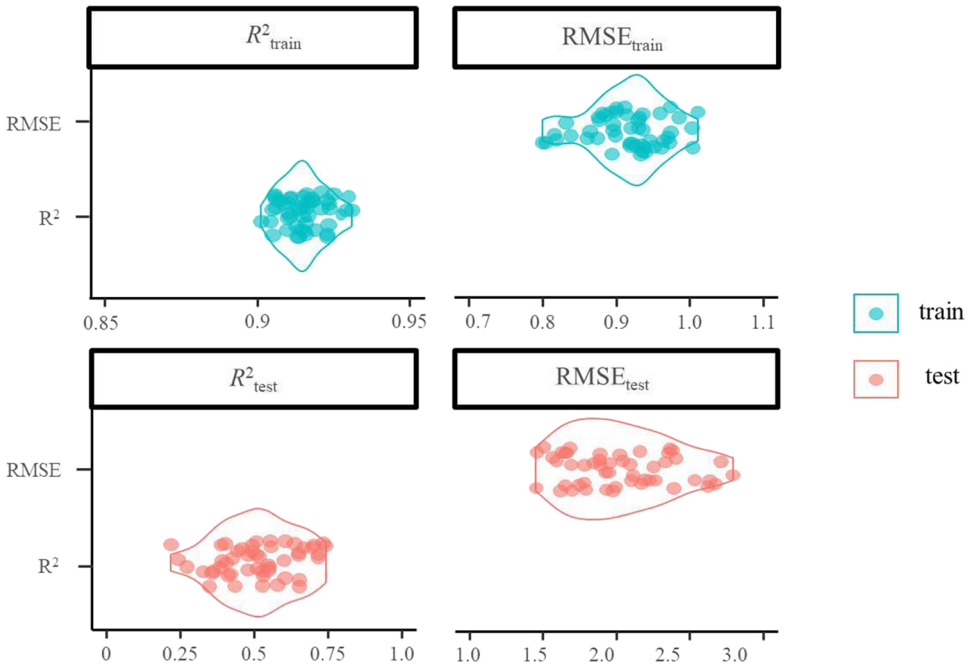


Fig. 2 R^2 and *RMSE* values derived from 50 runs of the random forest model for the estimation of LNC in the training and test datasets

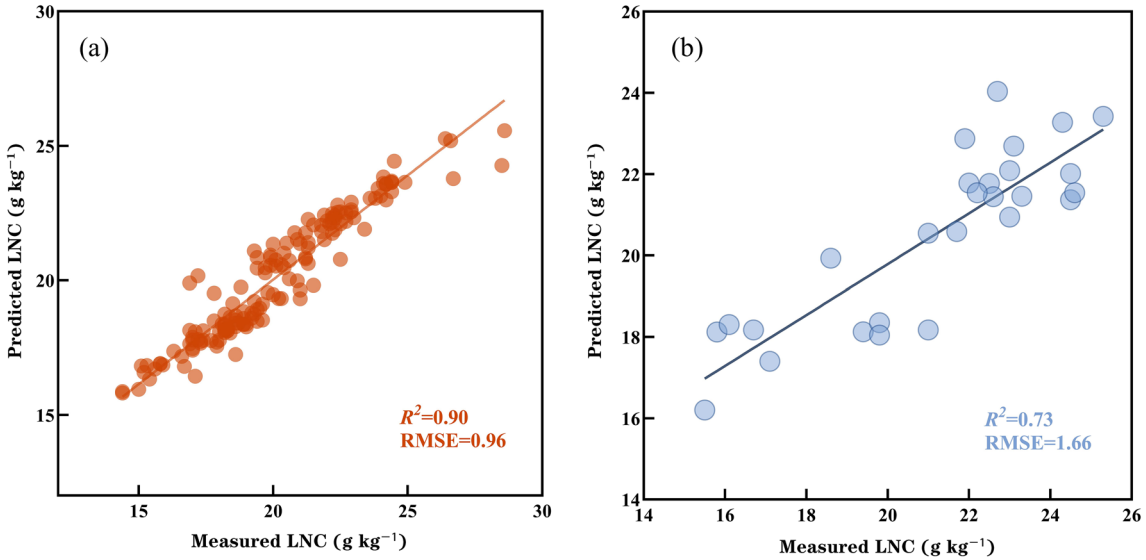


Fig. 3 The relationship between measured and predicted values obtained from the optimal random forest model. **a** N concentration in the training data set; **b** N concentration in the test data set

et al. 2014). The RPD of the model is 2.4 and 1.28 for the test and training data sets, respectively. The residual and quantile–quantile plots for the training and test data sets of the model are shown in Figs. 4 and 5. In the optimal model, the significance of each vegetation index variable is depicted in Fig. 6a. The results show that the vegetation indices V8, V10, and V11 contribute significantly to the construction of a random forest model for predicting leaf nitrogen content, with replacement accuracy correlation errors (IncMSE) of 12.2%, 11.6%, and 11.2%. Thus, these indices emerged as key predictors within our model. Conversely, the V1, V3, and V7 indices had the lowest IncMSE values of all indices examined, 2.6%, 5.6%, and 5.9%, respectively, highlighting their limited contribution. Figure 7 shows the correlation between all vegetation indices and LNC, revealing a lack of correlation between V1 and V4, V6, V11, further emphasizing the minimal role of V1 in predicting LNC. In addition, LNC showed significant negative correlations with all vegetation indices, especially with V8 and V10. To improve the predictive ability of the model, we excluded V1, V3, and

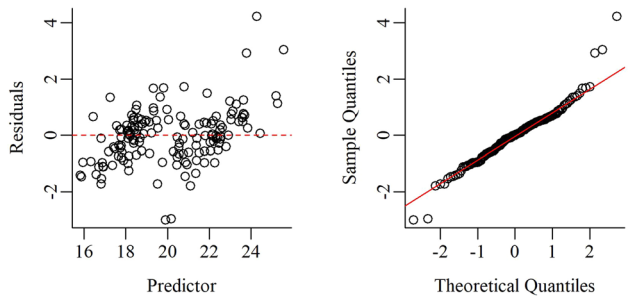


Fig. 4 Residual and quantile–quantile plots for the optimal model for the training data set

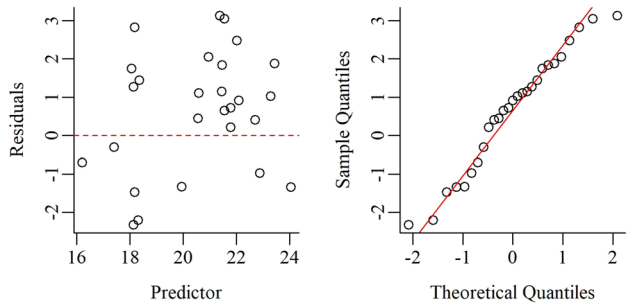


Fig. 5 Residual and quantile–quantile plots for the optimal model for the test data set

V7 due to their minimal IncMSE values, and after rebuilding and rerunning the Random Forest model (50 iterations), we observed an improved maximum R^2 of 0.76 (previously 0.73) and a reduced RMSE of 1.58 (previously 1.66). This refinement underscores the improved predictive performance of our model. Among the best performing models, V8 and V10 maintained their prominence with IncMSE values of 12.7% and 13.2%, respectively, as shown in Fig. 6b. We used the optimal model to predict the LNC of the remaining *M. glyptostroboides* (870 in total) in the sample area.

Effects of N and P fertilization on LNC of *M. glyptostroboides* plantations

Under N treatment, LNC at the N4.8 level was significantly higher than for the N2.4, N0.8 and N0 levels. Leaf nitrogen content at the N4 and N2.4 levels was significantly higher than for the N0 and N0.8 levels (Fig. 8a). With P fertilization, the LNC at the P0 level (control) was significantly higher than for the P0.2, P0.6 and P1 levels (Fig. 8b).

Pearson correlation analysis was carried out between the LNC, model-predicted, and the height, DBH and crown width for each tree. Under N treatment, the LNC was not significantly correlated with other indicators (Fig. 9a). With P fertilization, LNC was not significantly correlated with other indicators (Fig. 9b).

Discussion

Machine learning models for estimating LNC

Spectral remote sensing technology has become an effective research tool for quantitative estimation of vegetation parameters and monitoring growth and other related fields. Our study has demonstrated that the combination of

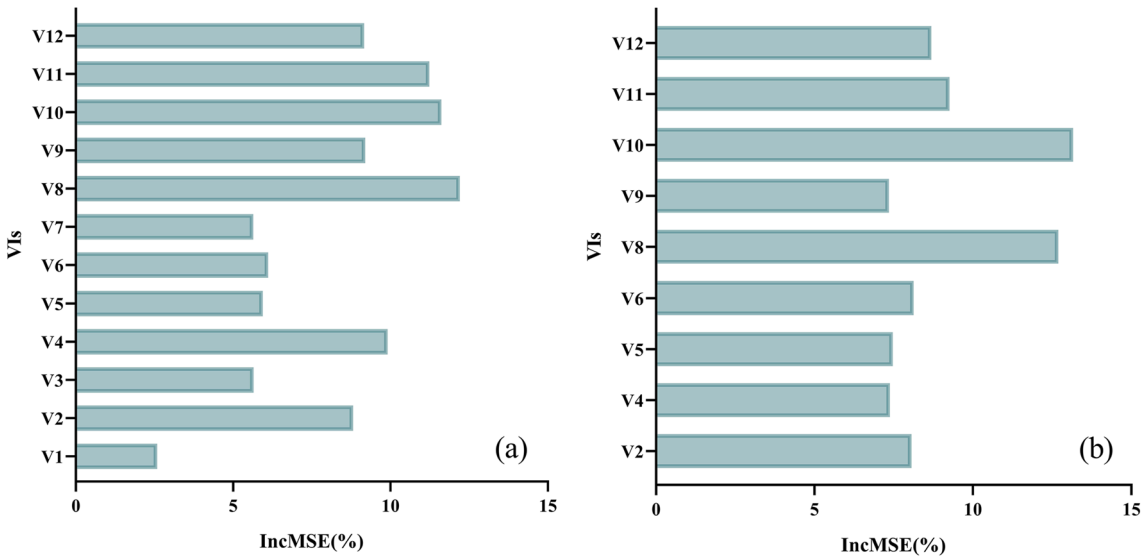


Fig. 6 Variable importance in predicting LNC of *M. glyptostroboides* using the optimal RF model

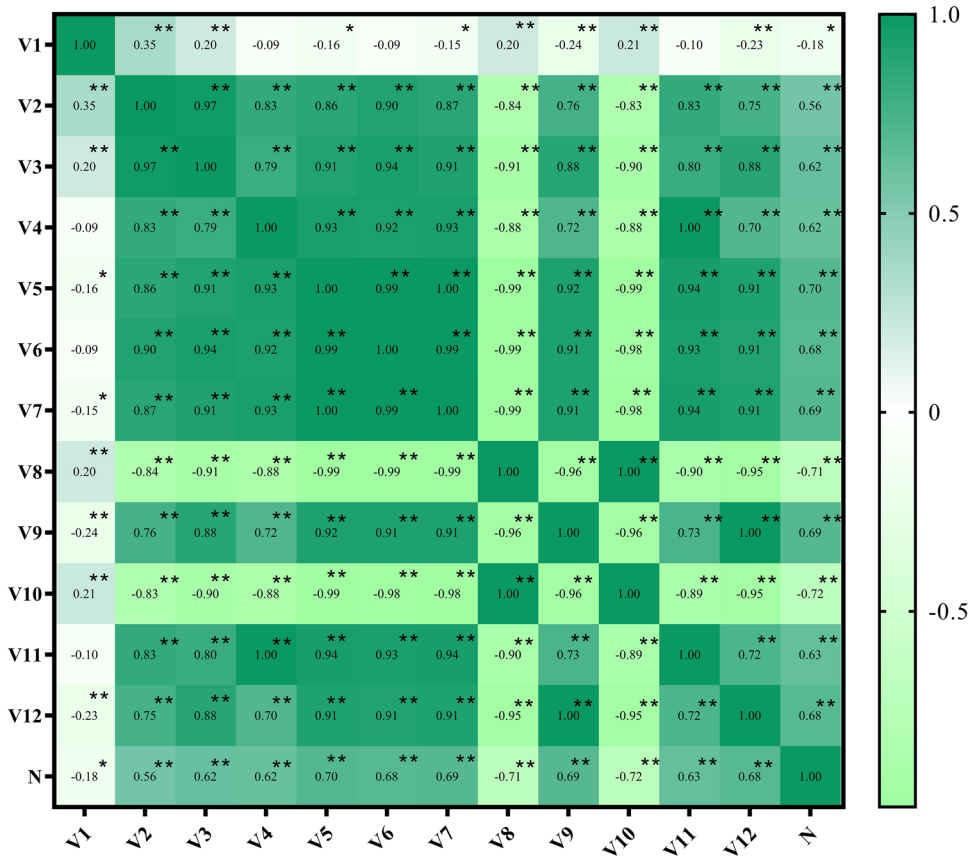


Fig. 7 Correlations between the various vegetation indices and the LNC

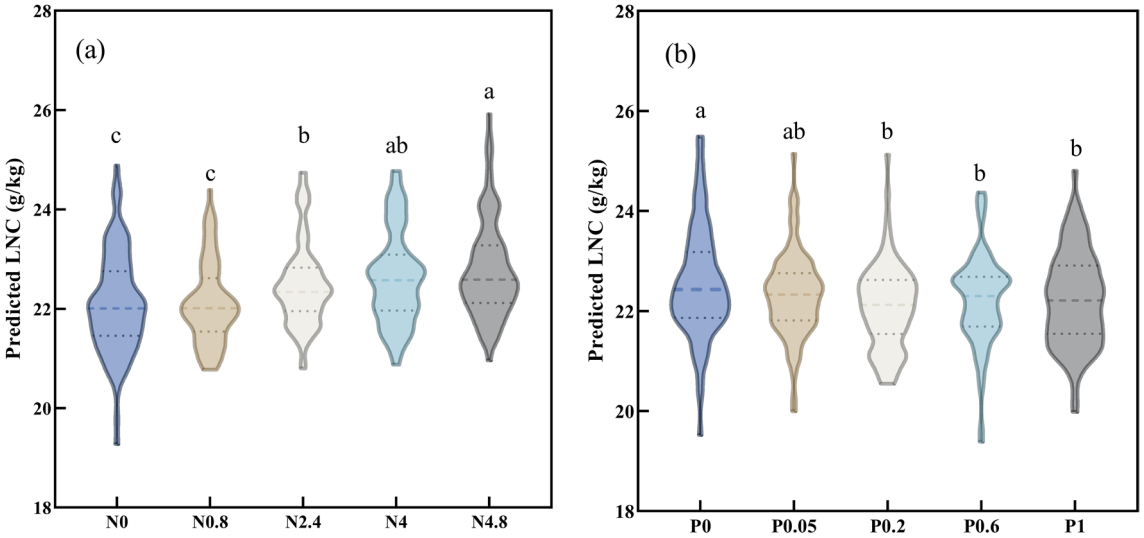


Fig. 8 Effect of N and P addition on the LNC of *M. glyptostroboides*

visible light spectral VIs and the ML algorithm is a powerful tool for estimating LNC in *M. glyptostroboides*. By comparing the performance of the five machine learning models, the random forest model performed the best. This is consistent with the results of other studies using RGB imagery in combination with ML models to estimate rice N concentration that found that the random forest algorithm performed the best and had the highest accuracy (Shi et al. 2021; Qiu et al. 2021).

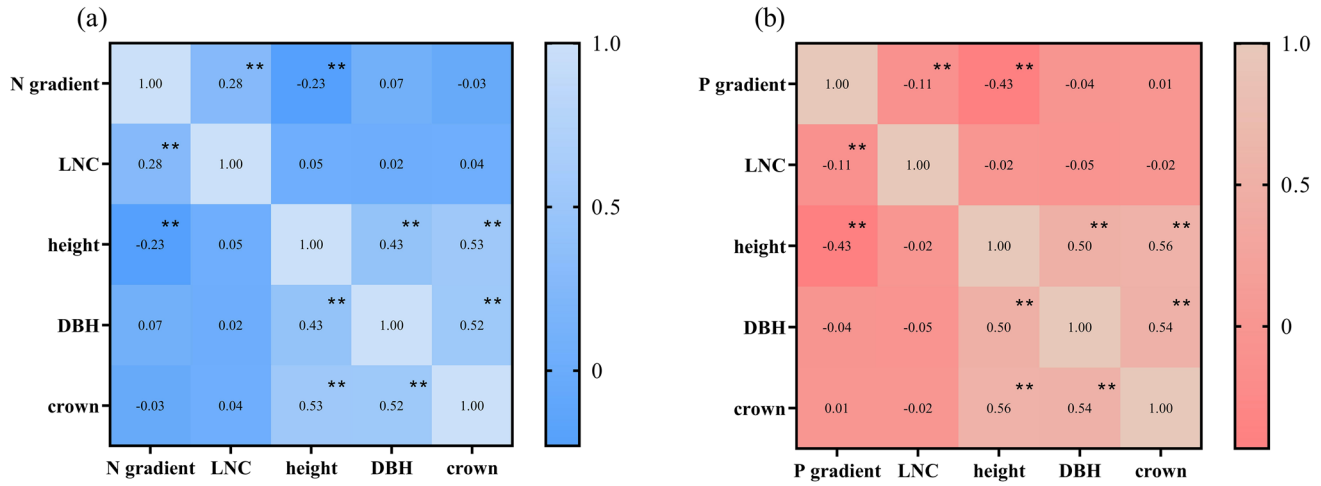


Fig. 9 Correlation analysis among indicators of *M. glyptostroboides* under N and P addition

We also attempted to estimate leaf P and carbon concentrations in *M. glyptostroboides* using visible light vegetation indices combined with the ML model, but the results were unsatisfactory, with an R^2 only around 0.2. This could be related to the selection of visible light VIs and the presence of certain interfering factors in the acquired UAV imagery, such as incomplete soil darkening. The neglect of this influence during color information extraction may result in less precise models (Liang et al. 2020). Therefore, for different output variables, it is necessary to use appropriate spectral VIs as input variables. At the beginning of the experiment, a well-designed remote sensing image acquisition method should be established to improve the accuracy of spectral information. At the same time, the image processing system should be further improved, and more pre-processing such as noise reduction and land darkening should be carried out in the image processing stage, so as to minimize the influence of the extra factors on the color information as much as possible (Li et al. 2007).

However, in recent studies, a BP neural network (BPNN) model constructed using canopy hyperspectral data to invert leaf carbon concentration (LCC) (Nie 2023) and VIs from multispectral images used as input variables in a random forest model to predict leaf phosphorus concentration (LPC) (Peng et al. 2022), both achieved an R^2 of about 0.7. Multispectral and hyperspectral cameras typically include multiple red-edge, near-infrared, and infrared bands that are very sensitive to leaf pigments, leaf structure, and canopy structure. RGB cameras, on the other hand, lack these bands. This may be the reason why chlorophyll indices calculated from multispectral and hyperspectral cameras have higher estimation accuracy compared to those from RGB cameras.

It has been shown that combining visible spectral data with weather and geographic data or with plant height and canopy cover can improve the accuracy of LNC estimation (Lu et al. 2021; Roth et al. 2023). Therefore, in future, spectral indices can be combined with plant structure parameters and weather and geographic data to improve the utility of visible UAV in forestry (Xiao et al. 2022b; Tang et al. 2022; Zhang et al. 2022).

Significant correlations were found between leaf nitrogen content and the twelve vegetation indices, indicating a link between nitrogen assimilation in plants and their spectral characteristics. In particular, V8 and V10 showed highly significant negative correlations with LNC, highlighting their substantial contributions to the predictive model.

Studies by Zheng et al. (2018) and Yang et al. (2020) confirmed that the V8 and V10 indices serve as comparatively better predictors of LNC. A low LNC indicates reduced chlorophyll in leaves, resulting in reduced light absorption and increased reflectance, thereby increasing visible light reflectance. Chlorophyll exhibits different absorption characteristics in the green and red spectral bands, with stronger absorption in the green band and weaker absorption in the red band. Since V1 and V3 represent values within the red and blue bands, respectively, their limited association with chlorophyll content may explain their low IncMSE values. Future research should focus on examining additional vegetation indices or their combinations to further refine the accuracy and robustness of LNC prediction models.

Effects of N and P fertilization on LNC in *M. glyptostroboides*

Leaves are the most sensitive organs of the plant in responding to changes in external conditions, and external nutrient supplies can reorder the allocation of plant nutrients by promoting plant growth and other internal homeostatic abilities (Tian et al. 2018; You et al. 2018). In this study, both N and P treatments significantly affected the LNC of

M. glyptostroboides predicted using the optimal model. Nitrogen increased the LNC, which is consistent with other research conducted on species such as *Quercus*, *Alnus*, *Larix* (Wang et al. 2022; Deng et al. 2015; Bai et al. 2015). The addition of nitrogen directly enhances the supply of N in the soil and promotes its absorption and utilization by *M. glyptostroboides*. Consequently, the LNC of *M. glyptostroboides* increased.

Our data showed that the addition of phosphorous significantly reduced LNC, which is different from the results of other studies. A meta-analysis found that P fertilization had either a positive effect or no significant effect on LNC (Yuan and Chen 2015; Sun et al. 2022). The results may be attributed to previous research findings from the sample site indicating that *M. glyptostroboides* on the coast had a leaf N/P ratio < 14, indicating N limitation. Previous studies have shown that P fertilization increases the N limitation in plants (Xu et al. 2021), and plants may decrease leaf nitrogen levels to adjust the balance between N and P; N and P are competitive for uptake and transport processes in plants (Tognetti et al. 2021). When the supply of P increases, plants may adjust the absorption and transport mechanisms in their roots, leading to LNC decreases.

Conclusion

This study utilized RGB spectral images acquired from an unmanned aerial vehicle to establish machine learning models relating leaf nitrogen content in coastal *M. glyptostroboides* plantations under N and P treatments to visible spectral vegetation indices. The performance of five machine learning models was compared, and it was concluded that the combination of visible spectral VIs and the random forest model can effectively predict LNC in *M. glyptostroboides*. Leaf nitrogen levels changed with the addition of N and P. This study validates the feasibility of using visible light spectral indices in conjunction with the random forest algorithm to establish prediction models for LNC in *M. glyptostroboides*, providing support for the non-destructive real-time monitoring of vegetation stoichiometry.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Bai XF, Xu FL, Wang WL, Zhao YF, Wang LL, Sun PY (2015) Ecological stoichiometry of soil carbon, nitrogen and phosphorus in a *Larix principis-rupprechtii* plantation. Sci Soil Water Conserv 13(6):68–75. https://doi.org/10.16843/j.sswc.2015.06.011 (**in Chinese**)

Banerjee BP, Joshi S, Thoday-Kennedy E, Pasam RK, Tibbits J, Hayden M, Spangenberg G, Kant S (2020) High-throughput phenotyping using digital and hyperspectral imaging-derived biomarkers for genotypic nitrogen response. J Exp Bot 71(15):4604–4615. https://doi.org/10.1093/jxb/eraa143

Barzin R, Kamangir H, Bora GC (2021) Comparison of machine learning methods for leaf nitrogen estimation in corn using multispectral UAV images. Trans ASABE 64(6):2089–2101. https://doi.org/10.13031/trans.14305

Bertsimas D, Dunn J (2017) Optimal classification trees. Mach Learn 106(7):1039–1082. https://doi.org/10.1007/s10994-017-5633-9

Breiman L (2001) Random forests. Mach Learn 45:5–32. https://doi.org/10.1023/A:1010933404324

Deng HJ, Zhang LN, Zhang GS, Lin YM, Wu CZ, Hong W (2015) Effects of nitrogen deposition on leaf elements and their stoichiometric ratios in *Schima superba* and *Pinus massoniana* mixed forest. J for Environ 35(2):118–124. https://doi.org/10.13324/j.cnki.jfcf.2015.02.004 (**in Chinese**)

Dimitriadou E, Hornik K, Leisch F, Meyer D, Weingessel A, Leisch MF (2009) Package 'e1071' R Software package, available at: https://cran.r-project.org/web/packages/e1071/index.html

Elzhov TV, Mullen KM, Spiess A, Bolker B (2016) Package 'minpack. lm'. Title R Interface Levenberg-Marquardt Nonlinear Least-Sq. Algorithm Found MINPACK Plus Support Bounds. Available at: https://rdrr.io/cran/minpack.lm/

Friedman JH (2002) Stochastic gradient boosting. Comput Stat Data An 38:367–378

Genuer R, Poggi JM (2020) Random forests with R. Springer, Berlin

Ghosal S, Blystone D, Singh AK, Ganapathysubramanian B, Singh A, Sarkar S (2018) An explainable deep machine vision framework for plant stress phenotyping. PNAS 115(18):4613–4618. https://doi.org/10.1073/pnas.1716999115

Jordan MI, Mitchell TM (2015) Machine learning: Trends, perspectives, and prospects. Science 349(6245):255–260. https://doi.org/10.1126/science.aaa8415

Ji JY, Luo YP, Sun XS, Chen FH, Luo G, Wu YJ, Gao Y, Ji RR (2021) Improving image captioning by leveraging intra- and inter-layer global representation in transformer network. AAAI21 35:1655–1663. https://doi.org/10.1609/aaai.v35i2.16258

Kou JM, Duan L, Yin CX, Ma LL, Chen XY, Gao P, Lv X (2022) Predicting leaf nitrogen content in cotton with UAV RGB images. Sustainability 14(15):9259. https://doi.org/10.3390/su14159259

Li XZ, Xie RZ, Wang KR, Bai ZY, Li SK, Wang FY, Gao SJ (2007) Acquiring nitrogen quantity in digital image of cotton leaf by artificial neutral network model. Acta Agron Sin 33(10):1662–1666 (**in Chinese**)

Li RR, Lu Y, Wang YM, Wan FX (2019) Effects of N addition on C, N and P stoichiometry and soil enzyme activities in *Cupressus lusitanica* Mill. plantation. Chin J Ecol 38(2):384–393. https://doi.org/10.13292/j.1000-4890.201902.028 (**in Chinese**)

Li W, Zhu X, Yu X, Li M, Tang X, Zhang J, Xue Y, Zhang C, Jiang Y (2022) Inversion of nitrogen concentration in apple canopy based on UAV hyperspectral images. Sensors (basel) 22(9):3503. https://doi.org/10.3390/s22093503

Liang J, Liu SY, Wang SB, Huang LP, Zhang JL, Wu QB, Guo F, Meng WW, Chen TT, Qi HX, Wang LD, Zhang Z, Wan SB, Zhang L (2020) Peanut nitrogen nutrition inversion based on unmanned

aerial vehicle remote sensing. Chin J Oil Crop Sci 42(6):1043–1050. https://doi.org/10.19802/j.issn.1007-9084.2019234 (**in Chinese**)

Lin H, Yan EP, Wang GX, Song RF (2014) Analysis of hyperspectral bands for the health diagnosis of tree species. In: 2014 third international workshop on earth observation and remote sensing applications (EORSA), Changsha. IEEE, pp 448–451

Liu XF, Lyu Q, He SL, Yi SL, Hu DY, Wang ZT, Xie RJ, Zheng YQ, Deng L (2016) Estimation of carbon and nitrogen contents in citrus canopy by low-altitude remote sensing. Int J Agric Biol Eng 9(5):149–157. https://doi.org/10.3965/j.ijabe.20160905.2246

Liu ZL, Peng CH, Work T, Candau JN, DesRochers A, Kneeshaw D (2018) Application of machine-learning methods in forest ecology: recent progress and future challenges. Environ Rev 26(4):339–350. https://doi.org/10.1139/er-2018-0034

Lu JS, Cheng DL, Geng CM, Zhang ZT, Xiang YZ, Hu TT (2021) Combining plant height, canopy coverage and vegetation index from UAV-based RGB images to estimate leaf nitrogen concentration of summer maize. Biosyst Eng 202:42–54. https://doi.org/10.1016/j.biosystemseng.2020.11.010

McCann CM, Baylis M, Williams DJL (2010) The development of linear regression models using environmental variables to explain the spatial distribution of *Fasciola hepatica* infection in dairy herds in England and Wales. Int J Parasitol 40(9):1021–1028. https://doi.org/10.1016/j.ijpara.2010.02.009

Nie LC (2023) Hyperspectral inversion of carbon, nitrogen and phosphorus stoichiometry of dominant plants in Yancheng Coastal Wetland. Acta Eco Sin 43(12):5173–5185. https://doi.org/10.5846/stxb202202280471(inChinese)

Noguera M, Aquino A, Ponce JM, Cordeiro A, Silvestre J, Arias-Calderón R, da Encarnação MM, Jordão P, Andújar JM (2021) Nutritional status assessment of olive crops by means of the analysis and modelling of multispectral images taken with UAVs. Biosyst Eng 211:1–18. https://doi.org/10.1016/j.biosystemseng.2021.08.035

Onishi M, Ise T (2021) Explainable identification and mapping of trees using UAV RGB image and deep learning. Sci Rep 11:903. https://doi.org/10.1038/s41598-020-79653-9

Peng XL, Chen DY, Zhou ZJ, Zhang ZT, Xu C, Zha Q, Wang F, Hu XT (2022) Prediction of the nitrogen, phosphorus and potassium contents in grape leaves at different growth stages based on UAV multispectral remote sensing. Remote Sens 14(11):2659. https://doi.org/10.3390/rs14112659

Prado Osco L, Marques Ramos AP, Roberto Pereira D, Akemi Saito Moriya É, Nobuhiro Imai N, Takashi Matsubara E, Estrabis N, de Souza M, Marcato J Jr, Gonçalves WN, Li J, Liesenberg V, Eduardo Creste J (2019) Predicting canopy nitrogen content in *Citrus*-trees using random forest algorithm associated to spectral vegetation indices from UAV-imagery. Remote Sens 11(24):2925. https://doi.org/10.3390/rs11242925

Qiu ZC, Ma F, Li ZW, Xu XB, Ge HX, Du CW (2021) Estimation of nitrogen nutrition index in rice from UAV RGB images coupled with machine learning algorithms. Comput Electron Agric 189:106421. https://doi.org/10.1016/j.compag.2021.106421

Racine JS (2012) RStudio: a platform-independent IDE for R and sweave. J Appl Econom 27(1):167–172. https://doi.org/10.1002/jae.1278

Ridgeway G, Ridgeway M (2004) The gbm package. R Foundation for Statistical Computing, Vienna, Austria 5(3).

Roth RT, Chen KR, Scott JR, Jung J, Yang Y, Camberato JJ, Armstrong SD (2023) Prediction of cereal rye cover crop biomass and nutrient accumulation using multi-temporal unmanned aerial vehicle based visible-spectrum vegetation indices. Remote Sens 15(3):580. https://doi.org/10.3390/rs15030580

Shen X, Cao L, Coops NC, Fan HC, Wu XQ, Liu H, Wang GB, Cao FL (2020) Quantifying vertical profiles of biochemical traits for forest plantation species using advanced remote sensing approaches. Remote Sens Environ 250:112041. https://doi.org/10.1016/j.rse.2020.112041

Shi PH, Wang Y, Xu JM, Zhao YL, Yang BL, Yuan ZQ, Sun QY (2021) Rice nitrogen nutrition estimation with RGB images and machine learning methods. Comput Electron Agric 180:105860. https://doi.org/10.1016/j.compag.2020.105860

Smola AJ, Schölkopf B (2004) A tutorial on support vector regression. Stat Comput 14(3):199–222. https://doi.org/10.1023/B:STCO.0000035301.49549.88

Sun Y, Wang C, Chen X, Liu S, Lu X, Chen HYH, Ruan H (2022) Phosphorus additions imbalance terrestrial ecosystem C: N: P stoichiometry. Glob Chang Biol 28(24):7353–7365. https://doi.org/10.1111/gcb.16417

Swift ML (1997) GraphPad prism, data analysis, and scientific graphing. J Chem Inf Comput Sci 37(2):411–412. https://doi.org/10.1021/ci960402j

Tang ZY, Xu WT, Zhou GY, Bai YF, Li JX, Tang XL, Chen DM, Liu Q, Ma WH, Xiong GM, He HL, He NP, Guo YP, Guo Q, Zhu JL, Han WX, Hu HF, Fang JY, Xie ZQ (2018) Patterns of plant carbon, nitrogen, and phosphorus concentration in relation to productivity in China's terrestrial ecosystems. PNAS 115(16):4033–4038. https://doi.org/10.1073/pnas.1700295114

Tang X, Gao H, Gao X, Li H, Jia S, Zhang X (2022) Monitoring of grape nitrogen content based on UAV multispectrum. For Mach Woodwork Equip 50(06):63−68(75). https://doi.org/10.13279/j.cnki.fmwe.2022.0082 (**in Chinese**)

Therneau T, Atkinson B, Ripley B (2015) Package 'rpart'. Available at: https://cran.ma.ic.ac.uk/web/packages/rpart/rpart.pdf. Accessed on 20 April 2016

Tian D, Du EZ, Jiang L, Ma SH, Zeng WJ, Zou AL, Feng CY, Xu LC, Xing AJ, Wang W, Zheng CY, Ji CJ, Shen HH, Fang JY (2018) Responses of forest ecosystems to increasing N deposition in China: a critical review. Environ Pollut 243:75–86. https://doi.org/10.1016/j.envpol.2018.08.010

Tognetti PM, Prober SM, Báez S, Chaneton EJ, Firn J, Risch AC, Schuetz M, Simonsen AK, Yahdjian L, Borer ET, Seabloom EW, Arnillas CA, Bakker JD, Brown CS, Cadotte MW, Caldeira MC, Daleo P, Dwyer JM, Fay PA, Gherardi LA, Hagenah N, Hautier Y, Komatsu KJ, McCulley RL, Price JN, Standish RJ, Stevens CJ, Wragg PD, Sankaran M (2021) Negative effects of nitrogen override positive effects of phosphorus on grassland legumes worldwide. PNAS 118(28):e2023718118. https://doi.org/10.1073/pnas.2023718118

Tong R, Wen YX, Wang JY, Lou CY, Ma C, Zhu NF, Yuan WW, Geoff Wang G, Wu TG (2022) Root nutrient capture and leaf resorption efficiency modulated by different influential factors jointly alleviated P limitation in Quercus acutissima across the North-South Transect of Eastern China. For Res. https://doi.org/10.48130/fr-2022-0007

Torresan C, Benito Garzón M, O'grady M, Robson TM, Picchi G, Panzacchi P, Tomelleri E, Smith M, Marshall L, Tognetti R, Rustad LE, Kneeshaw D (2021) A new generation of sensors and monitoring tools to support climate-smart forestry practices. Can J Forest Res 51(12):1751–1765

Urban A, Rogowski P, Wasilewska-Dębowska W, Romanowska E (2021) Understanding maize response to nitrogen limitation in different light conditions for the improvement of photosynthesis. Plants (basel) 10(9):1932. https://doi.org/10.3390/plants10091932

Wager S, Hastie T, Efron B (2014) Confidence intervals for random forests: the jackknife and the infinitesimal jackknife. J Mach Learn Res 15(1):1625–1651

Wang Y, Wang DJ, Zhang G, Wang J (2013) Estimating nitrogen status of rice using the image segmentation of G-R thresholding method. Field Crops Res 149:33–39. https://doi.org/10.1016/j.fcr.2013.04.007

Wang JY, Wang JN, Guo WH, Li YG, Wang GG, Wu TG (2018) Stoichiometric homeostasis, physiology, and growth responses of three tree species to nitrogen and phosphorus addition. Trees 32(5):1377–1386. https://doi.org/10.1007/s00468-018-1719-7

Wang RZ, Mao YX, Yun LL, You WZ, Zhang HD (2022) Effects of nitrogen addition on leaf carbon, nitrogen and phosphorus stoichiometry and nonstructural carbohydrates in Mongolian oak (*Quercus mongolica*). Chin J Ecol 41(7):1369–1377. https://doi.org/10.13292/j.1000-4890.202207.009

Wen BB, Xiao W, Mu Q, Li DM, Chen XD, Wu HY, Li L, Peng FT (2020) How does nitrate regulate plant senescence? Plant Physiol Biochem 157:60–69. https://doi.org/10.1016/j.plaphy.2020.08.041

Wen YX, Tong R, Zhang H, Feng KQ, Song R, Wang GG, Wu TG (2021) N addition decreased stand structure diversity in young but increased in middle-aged *Metasequoia glyptostroboides* plantations. Glob Ecol Conserv 30:e01803. https://doi.org/10.1016/j.gecco.2021.e01803

Wu L, Gong YJ, Bai XP, Wang W, Wang Z (2023) Nondestructive determination of leaf nitrogen content in corn by hyperspectral imaging using spectral and texture fusion. Appl Sci 13(3):1910. https://doi.org/10.3390/app13031910

Xiao Q, Tang W, Zhang C, Zhou L, Feng L, Shen J, Yan T, Gao P, He Y, Wu N (2022a) Spectral preprocessing combined with deep transfer learning to evaluate chlorophyll content in cotton leaves. Plant Phenomics 2022:9813841. https://doi.org/10.34133/2022/9813841

Xiao Q, Wu N, Tang W, Zhang C, Feng L, Zhou L, Shen J, Zhang Z, Gao P, He Y (2022b) Visible and near-infrared spectroscopy and deep learning application for the qualitative and quantitative investigation of nitrogen status in cotton leaves. Front Plant Sci 13:1080745. https://doi.org/10.3389/fpls.2022.1080745

Xu M, Zhu Y, Zhang S, Feng Y, Zhang W, Han X (2021) Global scaling the leaf nitrogen and phosphorus resorption of woody species: Revisiting some commonly held views. Sci Total Environ 788:147807. https://doi.org/10.1016/j.scitotenv.2021.147807

Yamashita H, Sonobe R, Hirono Y, Morita A, Ikka T (2020) Dissection of hyperspectral reflectance to estimate nitrogen and chlorophyll contents in tea leaves based on machine learning algorithms. Sci Rep 10:17360. https://doi.org/10.1038/s41598-020-73745-2

Yang B, Wang M, Sha Z, Wang B, Chen J, Yao X, Cheng T, Cao W, Zhu Y (2019) Evaluation of aboveground nitrogen content of winter wheat using digital imagery of unmanned aerial vehicles. Sensors (basel) 19(20):E4416. https://doi.org/10.3390/s19204416

Yang MJ, Hassan MA, Xu KJ, Zheng CY, Rasheed A, Zhang Y, Jin XL, Xia XC, Xiao YG, He ZH (2020) Assessment of water and nitrogen use efficiencies through UAV-based multispectral phenotyping in winter wheat. Front Plant Sci 11:927. https://doi.org/10.3389/fpls.2020.00927

You C, Wu F, Yang W, Xu Z, Tan B, Yue K, Ni X (2018) Nutrient-limited conditions determine the responses of foliar nitrogen and phosphorus stoichiometry to nitrogen addition: a global meta-analysis. Environ Pollut 241:740–749. https://doi.org/10.1016/j.envpol.2018.06.018

Yuan ZY, Chen HY (2015) Negative effects of fertilization on plant nutrient resorption. Ecology 96(2):373–380. https://doi.org/10.1890/14-0140.1

Zhang LP, Zhang LF, Du B (2016) Deep learning for remote sensing data: a technical tutorial on the state of the art. IEEE Geosci Remote Sens Mag 4(2):22–40. https://doi.org/10.1109/MGRS.2016.2540798

Zhang XH, Qiao Y, Meng FF, Fan CG, Zhang MM (2018) Identification of maize leaf diseases using improved deep convolutional neural networks. IEEE Access 6:30370–30377. https://doi.org/10.1109/ACCESS.2018.2844405

Zhang Y, Wu JB, Wang AZ (2022) Comparison of various approaches for estimating leaf water content and stomatal conductance in different plant species using hyperspectral data. Ecol Indic 142:109278. https://doi.org/10.1016/j.ecolind.2022.109278

Zheng HB, Cheng T, Li D, Zhou X, Yao X, Tian YC, Cao WX, Zhu Y (2018) Evaluation of RGB, color-infrared and multispectral images acquired from unmanned aerial systems for the estimation of nitrogen accumulation in rice. Remote Sens 10(6):824. https://doi.org/10.3390/rs10060824

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.