



Machine Learning-Based Leaf Silica Content Estimation From Hyperspectral Reflectance

Adenan Yandra Nofrizal¹ · Rei Sonobe² · Hiroto Yamashita² · Akio Morita² · Takashi Ikka²

Received: 12 June 2024 / Accepted: 22 March 2026
© The Author(s) 2026

Abstract

Silica (Si) in plants has become a subject of increasing interest for various communities, e.g., those involved in global change biogeochemistry, agronomy and biotechnology, and many techniques have been applied to quantitatively determine the Si content of plants. However, there are few applicable techniques for the in-situ assessment of Si contents because many methods are destructive and expensive. As a recent alternative, spectrophotometric analyses using hyperspectral data are both non-destructive and comparatively cheap, and machine learning algorithms have been used to enhance hyperspectral remote sensing to evaluate biochemical properties. The objectives of this study were to examine the potential of hyperspectral remote sensing approaches to estimate the silica content of *Zizania latifolia* using conventional machine learning algorithms including random forests, extreme gradient boosting and Cubist. The results indicate that Cubist was the best algorithm, achieving a ratio of performance to deviation of 2.02 and a root mean square error of 42.52 $\mu\text{g cm}^{-2}$.

Keywords Hyperspectral Remote Sensing · Pre-processing · *Zizania latifolia*

Introduction

Slag is produced by melting waste materials, such as municipal solid waste and incinerated ash, at high temperatures, followed by cooling and solidification [21, 26, 49]. At present, most molten slag is utilized in construction and civil engineering, primarily as aggregate for concrete products, asphalt mixtures, and backfill materials [4, 34]. In contrast, there is growing interest in expanding its use to high-value-added industries with stable demand, particularly agriculture. Slag fertilizers, which are rich in silica and composed mainly of mixed metal oxides, have been reported to neutralize soil acidity, enhance crop productivity, mitigate greenhouse gas emissions, and immobilize heavy metals [9].

Although silica is not considered as an essential element for plant growth, it plays a crucial role in maintaining plant mechanical strength and physiological functions [13]. Its beneficial effects are especially pronounced under stress conditions [22]. Silicon accumulation in stems increases structural rigidity and improves lodging resistance [12, 23], thereby reducing harvest losses caused by strong winds or typhoons. In addition, silica accumulation enhances resistance to biotic stresses, including rice blast, powdery mildew, and insect pests such as rice leaf bugs and planthoppers [10, 28, 48]. However, the effects of slag fertilization depend strongly on application rate. Low rates have been shown to enhance antioxidant enzyme activity and chlorophyll content, whereas excessive application can suppress physiological performance, seedling growth, and yield [6]. Accurate quantification of silica content is therefore essential for evaluating fertilization effects and optimizing fertilizer management.

Gravimetric analysis has traditionally been used to quantify silica content, but this method is destructive and unsuitable for monitoring temporal changes in foliar silica during the growing season. Consequently, non-destructive approaches that allow frequent measurements are needed. Remote sensing provides an efficient means of acquiring consistent information on environmental and plant

Handling Editor: Rajender Parsad

✉ Rei Sonobe
sonobe.rei@shizuoka.ac.jp

¹ Department of Applied Geoinformatics and Cartography, Charles University, Praha, Czech Republic

² Faculty of Agriculture, Shizuoka University, Shizuoka, Japan

properties and is particularly effective for capturing spatial and temporal variability in vegetation phenology [3, 29, 43–45, 50]. Among these techniques, hyperspectral remote sensing has shown strong potential for estimating plant biochemical properties [24, 38, 41, 42].

Recent advances in data-driven approaches, including machine learning and artificial intelligence, have further expanded the analytical capacity of spectral data. In plant and agricultural sciences, these methods have been widely applied for genomic selection, multi-trait prediction, and the integration of complex datasets, including high-dimensional phenotypic and environmental information [11]. Machine learning algorithms are particularly well suited to hyperspectral analysis because they can capture non-linear relationships between spectral inputs and biochemical traits and automatically identify informative features while reducing redundancy [8, 15, 18].

Despite these advances, there remains a lack of systematic evaluation of machine learning algorithms for non-destructive estimation of foliar silica content under field conditions, particularly in graminoid species subjected to different slag fertilizer treatments. Addressing this gap is essential for translating spectral and data-driven approaches into practical fertilization management.

In this study, three widely used machine learning algorithms—Random Forest (RF), eXtreme Gradient Boosting (XGBoost), and Cubist—were evaluated for their ability to estimate silica content from hyperspectral reflectance data.

The objectives were to (1) quantify silica content in graminoid species using spectroscopic analysis and (2) identify the most effective machine learning algorithm for estimating silica content in wild rice plants subjected to slag fertilizer treatments.

Materials and Methods

Measurements

Manchurian wild rice (*Zizania latifolia* cv. Aogara), a silicicolous species that actively absorbs silica through its roots, was cultivated in a paddy field at Shizuoka University (Shizuoka, Japan). Plants were grown under flooded conditions with within-row and inter-row spacing of 100 cm (Fig. 1). Transplanting was conducted on 29 May 2020, and harvesting took place on 2 and 5 October 2020.

Molten slag was used as a siliceous fertilizer under the commercial name SK Keikaru, produced by a shaft-type gasification melting furnace and supplied by the Nishigaya Incineration Plant (Shizuoka, Japan). Basal fertilization consisted of 18 kg NH_4Cl , 12 kg P_2O_5 , and 12 kg K_2O per 1000 m^2 . An additional topdressing was applied, supplying 12 kg NH_4Cl , 12 kg P_2O_5 , and 6 kg K_2O per 1000 m^2 . The soluble silicic acid content of the molten slag was 32%, and the standard application rate was 127 kg per 1000 m^2 . Plants were cultivated under five slag fertilizer treatments: 0, 1, 2,



Fig. 1 *Zizania latifolia* cultivated in a paddy field

4, and 8 times the standard rate. Because the experiment was conducted in a continuously cultivated paddy field, residual silica in the soil was expected. Slag fertilizer was supplied by aerial spraying at the time of transplanting.

Leaf reflectance and silica content were measured for 200 leaves (40 leaves per treatment), randomly sampled on 2 and 5 October 2020. To minimize spatial variability, reflectance was measured 15 times per leaf, and the average spectrum was used. Hyperspectral reflectance was recorded at 1 nm intervals across the 400–2500 nm range using a leaf clip with a halogen light source attached to an ASD FieldSpec 4 Standard-Res Spectroradiometer (Malvern Panalytical Ltd., Malvern, UK), thereby reducing illumination-related errors.

Silica content was quantified using gravimetric analysis following dry ashing [32]. One gram of dried, ground plant material was placed in a 200-ml-tall beaker and capped with a watch glass. Then, 20 ml of 60% nitric acid was added, and the mixture was heated to 80 °C for 15 min and then to 100 °C for 15 min using a hot plate (HTP552AA HOT PLATE, ADVANTEC). Subsequently, the mixture was heated to 150 °C for three to four hours. Next, the watch glass was removed, and 10 ml of 60% perchloric acid was added. The mixture was heated to 80 °C for 15 min, 100 °C for 15 min and 150 °C for three to four hours using the hot plate (until the mixture became transparent). The mixture was cooled and diluted by a factor of 5 using Milli-Q water and was then filtered using filter paper (No. 6, ADVANTEC). Subsequently, we washed the beaker using Milli-Q water and washed the filter paper using 60 °C warm water to reduce the reactivity of chlorine. Finally, the paper was burned at 500 °C in a muffle furnace, and the ash weight was measured. The silica content was quantified using Eq. (1).

$$\text{Silica content } (\mu\text{g}/\text{cm}^2) = \text{Final weight (g)} / \text{Initial weight (g)} \quad (1)$$

$$\times \text{LMA} (\mu\text{g}/\text{cm}^2),$$

where, LMA is leaf mass per area.

Pre-Processing of Raw Reflectance Data

Previous studies have compared several spectral pre-processing methods, including continuum removal, de-trending (DT), first-derivative transformation, multiplicative scatter correction, and standard normal variate (SNV), for retrieving vegetation biochemical properties such as chlorophyll

and carotenoid contents. Among these methods, DT was reported to be the most effective for reflectance data acquired using the ASD FieldSpec® 4 spectroradiometer [24, 39, 40].

De-trending was therefore applied in this study to reduce noise and correct baseline and slope variations in spectral data. In this method, the baseline is modeled as a second-degree polynomial function of wavelength and subtracted from the original spectrum. DT has been widely used to account for baseline shifts and spectral curvilinearity [16].

Regression Models Based on Machine Learning Algorithms

The dataset was divided into training (50%), validation (25%), and test (25%) subsets using a stratified sampling strategy. Measurements were first grouped according to slag fertilizer treatments. Within each group, 50% of the samples were randomly assigned to the training set for model development. Half of the remaining samples were used for validation to optimize hyperparameters, and the rest were reserved for independent testing. This procedure was repeated 100 times to ensure robust model evaluation.

For Random Forest (RF), key hyperparameters included the number of trees, maximum number of features considered at each split (mtry), maximum tree depth, minimum samples required for node splitting, and minimum samples per leaf [2, 47]. The number of trees was set to 1000 to balance predictive performance and computational cost. Other parameters were tuned to control model complexity and reduce overfitting.

For eXtreme Gradient Boosting (XGBoost) [7], the number of boosting rounds was set to 1000. The learning rate (eta), maximum tree depth, minimum loss reduction for further partitioning, and L1 and L2 regularization parameters were optimized to improve generalization while avoiding overfitting.

Cubist model hyperparameters, including the number of committees, maximum number of rules, and maximum number of interactions allowed in each rule, were also optimized [20]. Increasing model complexity can improve accuracy but may lead to overfitting; therefore, these parameters were carefully tuned.

Bayesian optimization, a sequential model-based approach for optimizing black-box functions, was employed to optimize the hyperparameters of all machine learning algorithms [36]. All calculations were conducted using R v4.0.2 [30]. The R packages used for application of the various machine learning algorithms are listed in Table 1.

Table 1 Machine learning algorithms and packages used in this study

Algorithm	Package	Literature
Random Forest (RF)	randomForestSRC	[2, 47]
eXtreme gradient boosting (XGBoost)	xgboost	[7]
Cubist	Cubist	[20]

Performance assessment

To assess the performances of the regression models, the ratio of performance to deviation (RPD, Eq. (2)) [51] and the root-mean-square error (RMSE, Eq. (3)) were calculated. Each method was classified into three categories based on RPD: 'A' ($RPD > 2.0$), 'B' ($1.4 \leq RPD \leq 2.0$) or 'C' ($RPD < 1.4$). Models classified as 'A' or 'B' were assumed to have the potential to estimate silica contents [5]. The coefficient of determination (R^2 , Eq. (4)) was also calculated. The formulae for these metrics are as follows:

$$RPD = SD/RMSE \quad (2)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=0}^n (\hat{y}_i - y_i)^2} \quad (3)$$

$$R^2 = 1 - \left(\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \right), \quad (4)$$

where SD is the standard deviation of the silica content in the test data, n is the number of samples, y_i is the measured silica content, $\hat{y}_i$ is the estimated silica content and $\bar{y}$ is the mean value of the measured silica content.

Note: The performance metrics used in this study (RPD, RMSE, and R^2) are calculated under the standard assumptions that model residuals are approximately independent and identically distributed. While strict normality of residuals is not a prerequisite for model training in the applied machine learning algorithms, these metrics implicitly

assume that prediction errors are unbiased and not strongly autocorrelated. To enhance model robustness, all models were evaluated using repeated random stratified sampling and independent test datasets, which reduces the influence of data partitioning and potential violations of distributional assumptions. The sensitivity (S_i) for wavelength i (in nm) were calculated by Eq. (5) [17, 25]:

$$S_i = \frac{Var(f(X_{400}, \dots, X_i, \dots, X_{850})) - f(\bar{X})}{Var(Y)}, \quad (5)$$

where Var is the variation, $f()$ is the prediction of spectra resulting from the variation in wavelength i when other

wavelengths are held constant at their mean values, $f(\bar{X})$ is the estimated value based on the mean reflectance, and Y represents the measured carotenoid content. Once we had calculated S_i , we converted the scores to percentages.

Results

Measured Results

Figures 2 and 3 show the histograms of measured silica content and leaf mass per area (LMA). LMA ranged from 4.216 to 7.257 mg cm^{-2} , with a mean of 5.691 mg cm^{-2} and a standard deviation of 0.565 mg cm^{-2} . In contrast, silica content ranged from 101.04 to 570.68 $\mu\text{g cm}^{-2}$, with a mean of 243.37 $\mu\text{g cm}^{-2}$ and a standard deviation of 86.39 $\mu\text{g cm}^{-2}$. Silica accounted for 1.9–9.2% of LMA (mean: 4.2%). A significant positive correlation was observed between silica content and LMA ($r=0.701$, $p<0.001$; Fig. 4).

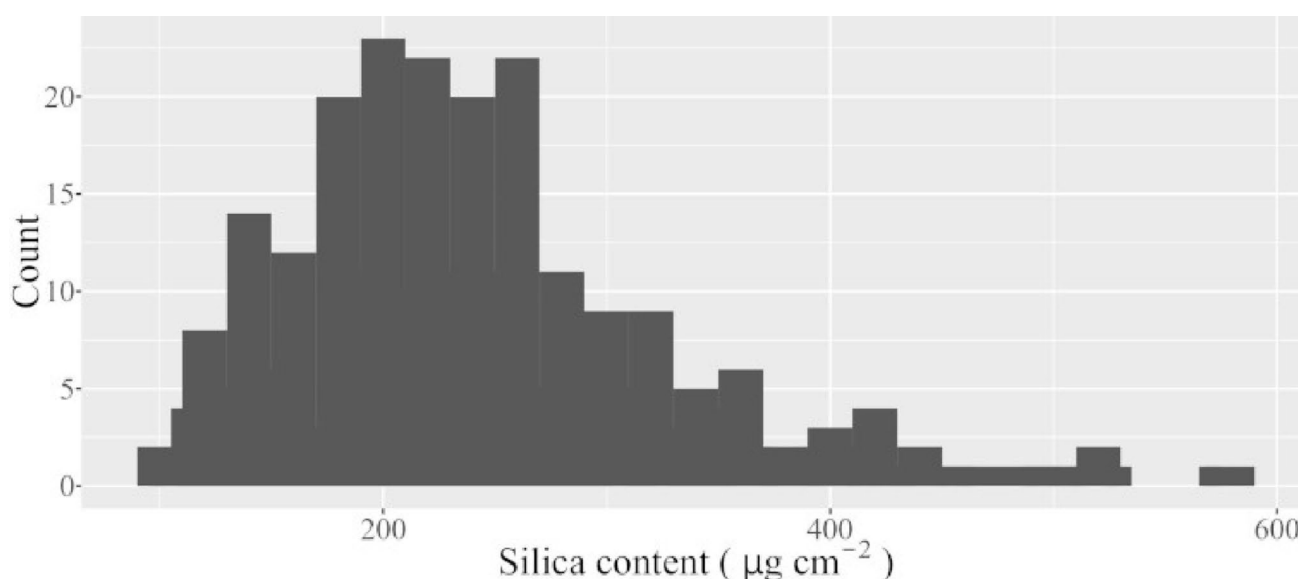


Fig. 2 Histograms of silica content

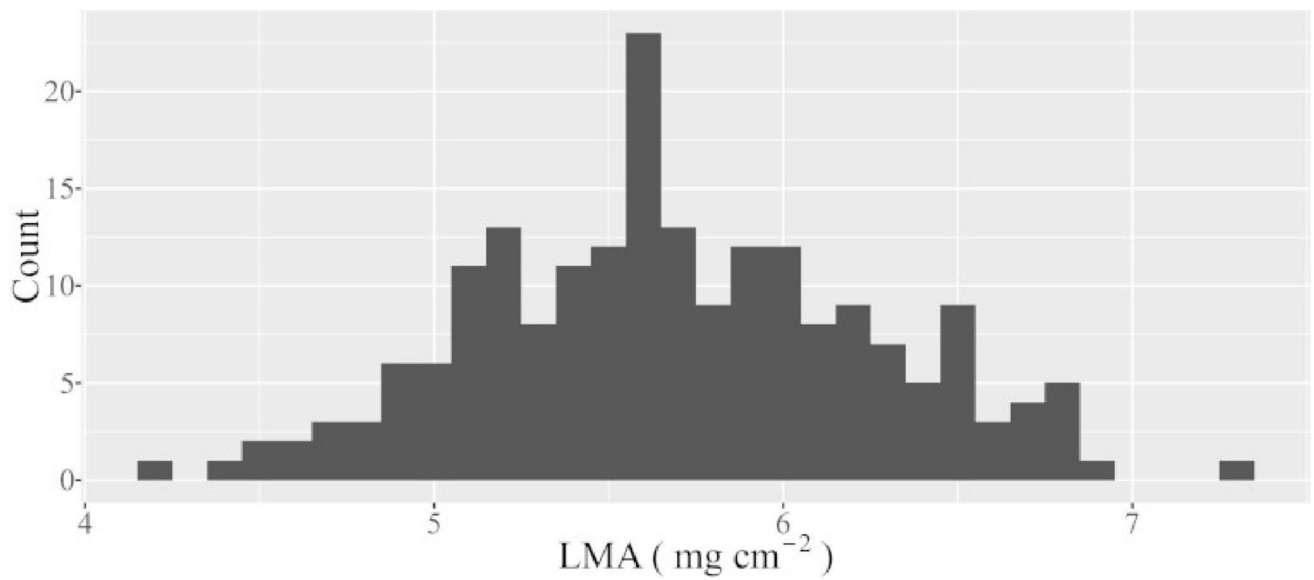


Fig. 3 Histograms of LMA

Fig. 4 Relationships between silica content and Leaf mass per area

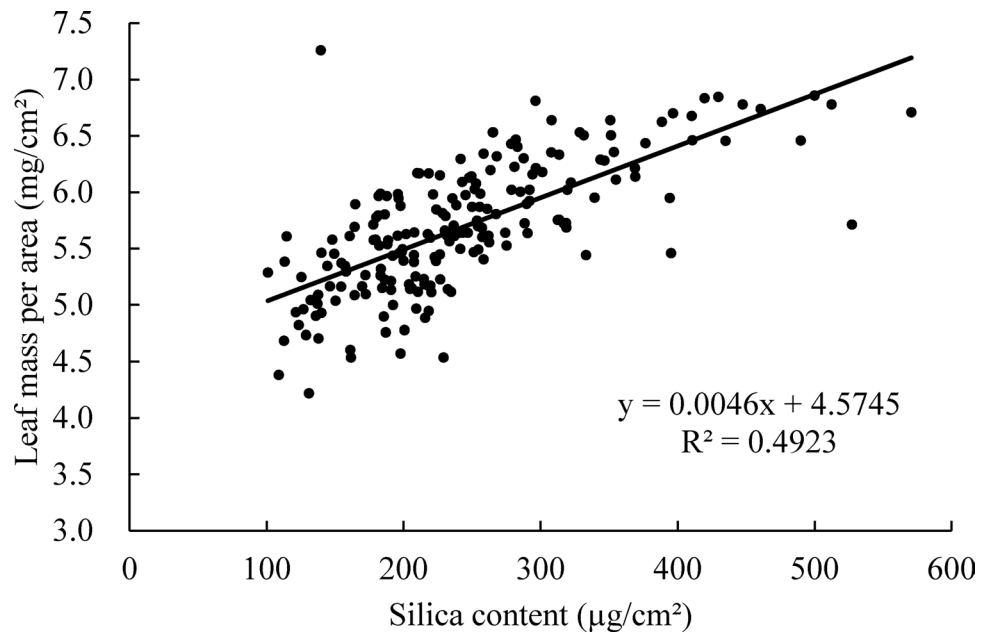


Figure 5 shows the correlation coefficients between the silica content and the reflectance at each wavelength for original reflectance and spectra after DT. After detrending, correlation coefficients decreased slightly in the 800–1400 nm region. Nevertheless, strong correlations were evident, including negative correlations near 520, 740, 1720, and 2315 nm and a positive correlation near 650 nm. The strongest positive correlation was observed at 654 nm ($r=0.425$), whereas the strongest negative correlation occurred at 1720 nm ($r=-0.498$).

Estimation Results

After 100 repetitions, model performance was evaluated using the test datasets (50 leaves $\times$ 100 iterations). Root mean square error (RMSE), residual prediction deviation (RPD), and coefficient of determination (R^2) were calculated for the three machine learning algorithms (Table 2). Both RF- and XGBoost-based models were classified as Category C, with RPD values below 1.4, indicating limited suitability for regression. In contrast, the Cubist-based model was consistently classified as Category A, with RPD values exceeding 2.0, indicating excellent predictive performance.

Fig. 5 Correlation coefficients between reflectance and silica content

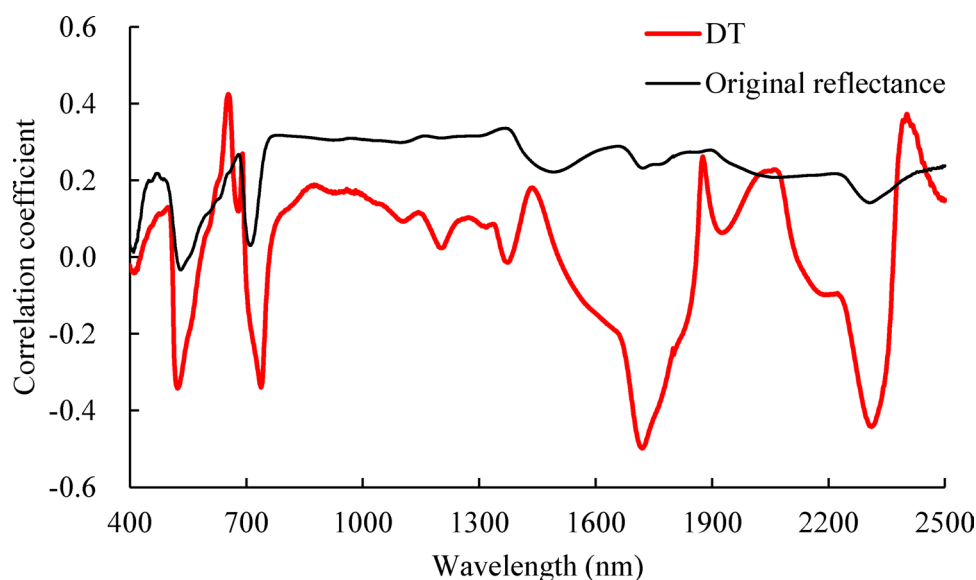


Table 2 Evaluation results of regression models. Values are expressed as mean $\pm$ standard deviation

Algorithm	RMSE ($\mu\text{g}/\text{cm}^2$)	RPD	R^2
RF	71.95 ± 8.12	1.20 ± 0.09	0.30 ± 0.10
XGBoost	65.09 ± 8.57	1.32 ± 0.12	0.43 ± 0.17
Cubist	42.52 ± 5.04	2.02 ± 0.25	0.76 ± 0.06

Figure 6 depicts the relationship between actual and estimated values, showing that both RF and XGBoost exhibited a significant decrease in estimation accuracy for silica content exceeding $400 \text{ mg}/\text{cm}^2$.

Table 3 summarizes the number of times each category (A, B, or C) was classified every iteration. When using RF, only one iteration was available (RPD = 1.43, Category B), and robustness could not be confirmed. In this case, the Cubist model was classified as B, while the XGBoost model was classified as C.

With XGBoost, depending on the partitioning method, 10 iterations were classified as Category A, indicating an excellent model, however, 47 iterations were classified as Category C, indicating unsuitability. On the other hand, the minimum RPD value when using Cubist was 1.47, and it was never judged as unsuitable for use.

Based on the evaluation using the RPD value, XGBoost-based model was determined to be a superior model to Cubist in 10 cases out of 100 iterations. Among these cases, RPD exceeded 2.0 in four iterations regardless of the algorithm used. In two iterations, XGBoost achieved Category A while Cubist remained in Category B.

Sensitivity Analysis

Figure 7 presents the results of the sensitivity analysis. In the XGBoost-based model, reflectance at 850 nm and 2300 nm

accounted for approximately 40% of total importance each. The importance of the 2300 nm band was also confirmed in the RF-based model, where it contributed nearly 40%. However, reflectance at 850 nm showed negligible importance ($<1\%$) in the RF model.

The Cubist-based model exhibited a distinct pattern. Reflectance in the 1675–1700 nm range, which showed a strong negative correlation with silica content, was identified as the most important predictor. However, the maximum importance did not exceed 25%, indicating that Cubist distributed importance across multiple bands rather than relying on a few dominant wavelengths.

Discussion

Relationship Between Reflectance and the Silica Content

Reflectance over the near-infrared domain (around 1900 and 2300 nm for graminoids) is effective in estimating the plant silica content, as this silica always occurs as hydrated condensates of orthosilicic acid ($\text{Si}(\text{OH})_4$) linked to organic biomolecules [35]. A significant ($p < 0.001$) positive correlation was confirmed between the silica content and reflectance at 680 nm, which is close to the red-edge inflection point, and this relationship was useful for estimating the chlorophyll content [52].

Slag fertilizer also influences the structure of leaves, such as the specific leaf area, thickness and leaf dry matter content [14]. Near-infrared and shortwave infrared reflectance have therefore been widely used to evaluate leaf dry matter [14], which likely contributes to the observed relationship between reflectance and silica content.

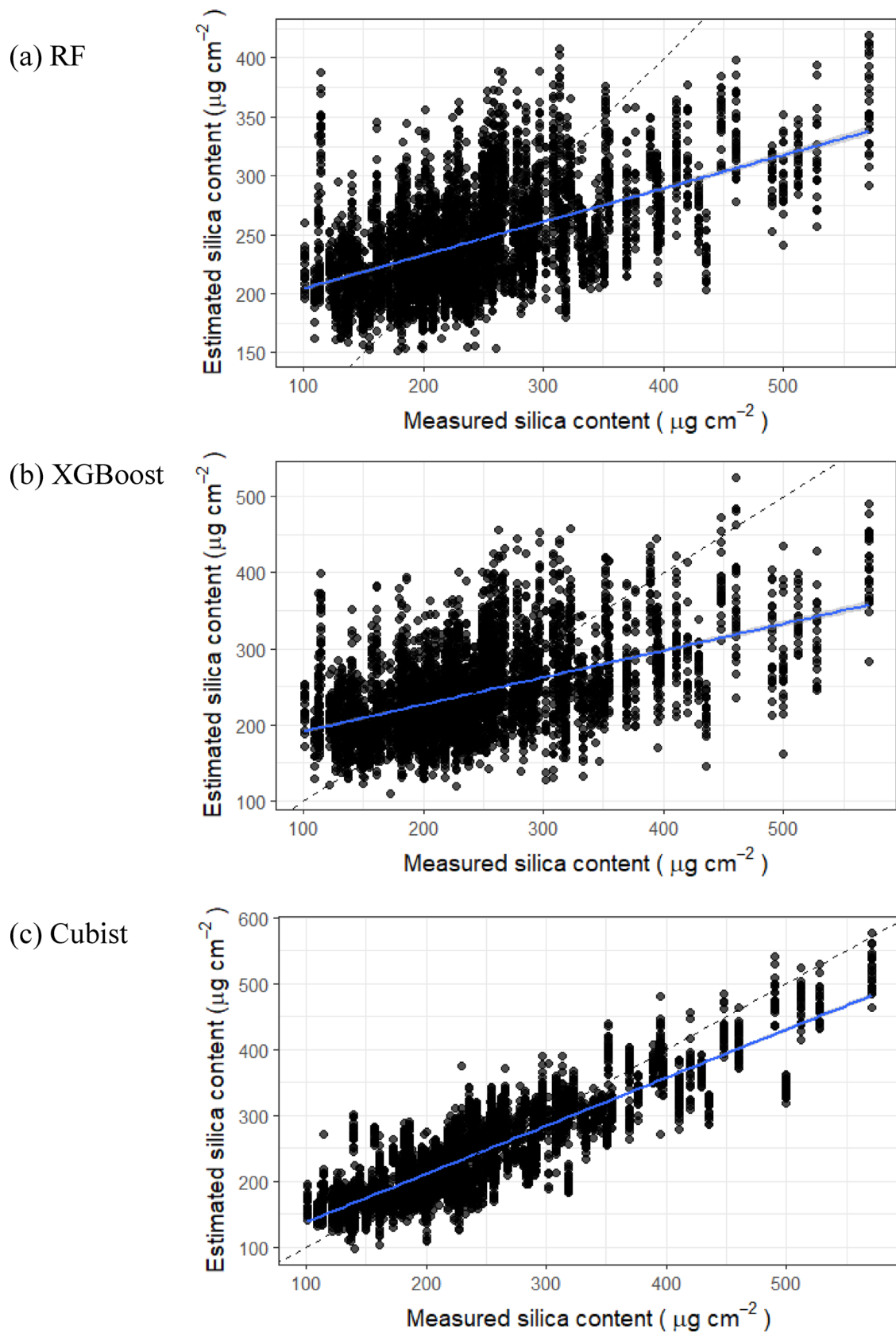


Fig. 6 Relationships between measured and estimated silica content

Table 3 Distribution of iteration cases in categories under different models

Category	RF	XGBoost	Cubist
A	0	10	60
B	1	43	40
C	99	47	0

Accuracy Assessment

Cubist has been reported to perform well with relatively small datasets [15], and the present study confirms its suitability for silica estimation. Although few studies have estimated plant silica content using in-field hyperspectral measurements, related research has proposed reflectance indices for evaluating silica in pest insects. Ahmed [1] reported that the index $(R_{750} - R_{560}) / (R_{750} + R_{560})$ effectively

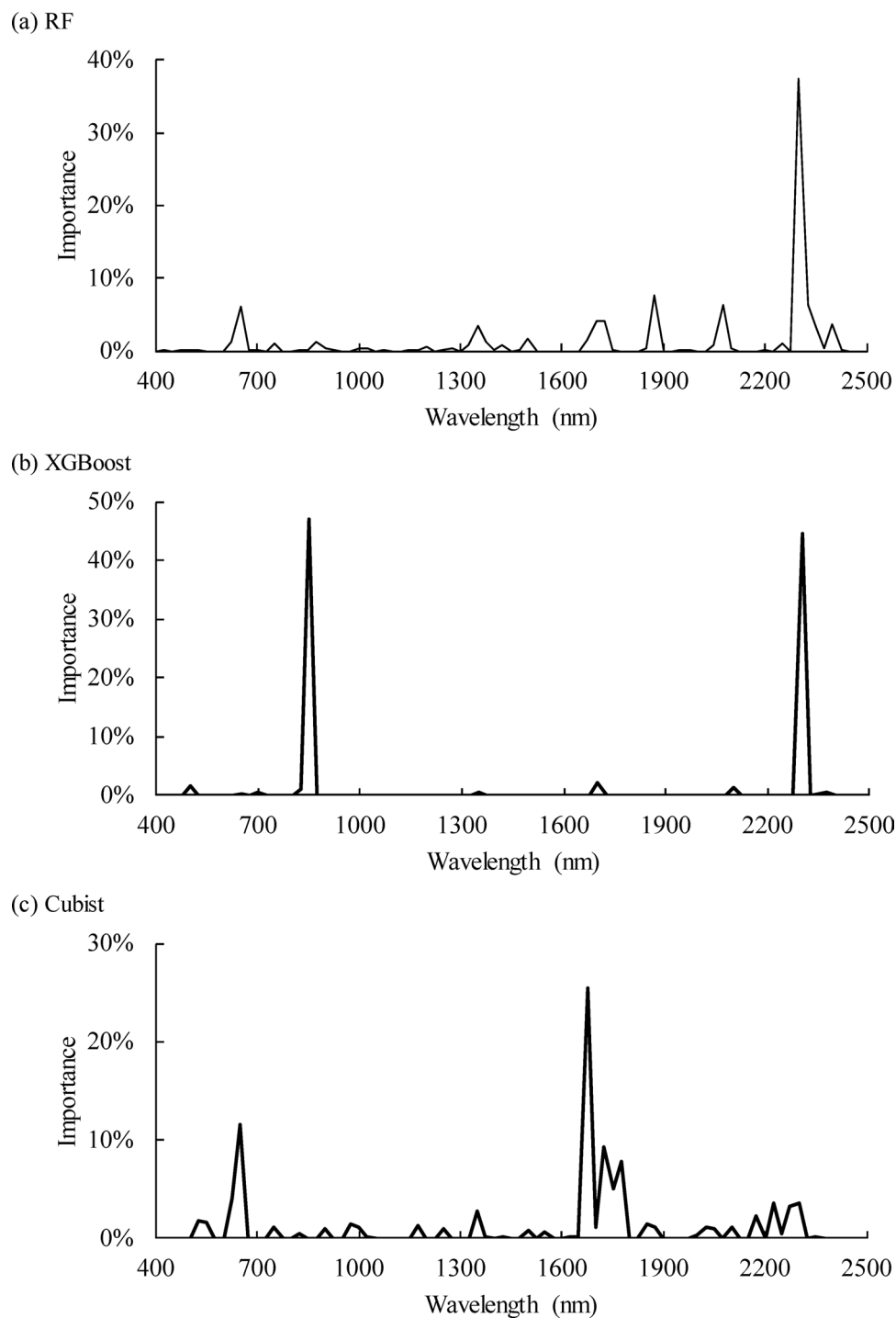
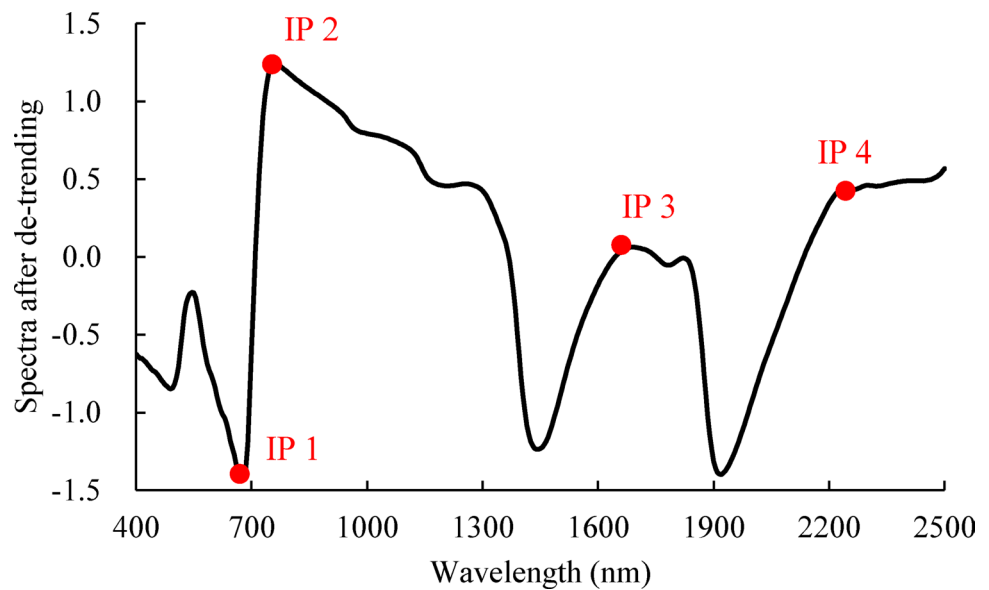
Fig. 7 Sensitivity analysis of regression models

Fig. 8 Sample spectra after de-trending

estimated silica concentration in *Eldana saccharina* ($R^2 = 0.53$). In addition, the nitrogen-to-silica ratio was correlated with $(R_{1075} - R_{730})/(R_{1075} + R_{730})$ ($R^2 = 0.67$).

For *Zizania latifolia*, significant correlations were observed between silica content and reflectance near 560 nm ($r = -0.342$ at 523 nm) and 730–750 nm ($r = -0.339$ at 737 nm), whereas no significant correlation was found at 1075 nm. However, these optical bands contributed little to the Cubist model. Reflectance in the visible region is strongly influenced by chlorophyll content [19, 33, 37, 46], and its confounding effects may be unavoidable in plant-based analyses. Consequently, machine learning models targeting plants selected different spectral features than those used in studies focusing on insects [1].

Figure 8 illustrates de-trended spectra with four inflection points (IP1–IP4). IP3 (1680–1700 nm) was most important for the Cubist model, while IP1 contributed to both RF and Cubist. IP2 was dominant in XGBoost, whereas IP4 was important for RF and XGBoost despite showing no significant correlation with silica content ($r = 0.004$). This suggests that models emphasizing IP4 may lack robustness.

Previous studies have linked near-infrared reflectance (e.g., 9.08 and 12.36 μm) to silica content [31]. Smis et al. [35] reported estimation errors of 10–24% relative to dry matter, while Quigley et al. [27] reported errors of 0.47–0.65% using partial least squares regression. In contrast, the present study achieved leaf-scale estimation of silica content ranging from 1.9% to 9.2% (relative to dry weights of 4.2–7.3 g) with an RMSE of 42.52 $\mu\text{g cm}^{-2}$. These results demonstrate that Cubist is a promising algorithm for estimating silica content from hyperspectral reflectance with higher accuracy than many previously reported approaches.

Limitations and Practical Implications

Despite the promising results, several limitations of this study should be acknowledged. First, the size of the training dataset was relatively small and derived from a single field experiment. Although repeated stratified sampling was applied to enhance robustness, larger and more diverse datasets are required to further improve model stability and generalization. Second, hyperspectral data processing and machine learning model optimization involve substantial computational costs, which may limit immediate large-scale implementation. Third, this study focused on *Zizania latifolia* under specific environmental and management conditions; therefore, the transferability of the developed models to other crops or agroecosystems should be validated in future studies.

From a practical perspective, the proposed framework offers valuable implications for crop management and breeding. Non-destructive estimation of foliar silica content enables rapid phenotyping of silica uptake efficiency, which can support the selection of genotypes with enhanced stress resistance and lodging tolerance. For breeders, this approach provides a complementary tool to conventional phenotyping by allowing repeated measurements throughout the growing season without destructive sampling. Moreover, integration of hyperspectral sensing with data-driven models can facilitate precision fertilization strategies, supporting sustainable use of slag-based silica fertilizers and improved nutrient management in paddy systems.

Conclusions

In this study, we evaluated the feasibility of estimating the silica content of *Zizania latifolia* cultivated under different levels of slag fertilization using hyperspectral reflectance data and machine learning approaches. Slag produced by shaft-type gasification melting furnaces is effective for waste volume reduction, and its conversion into fertilizer represents a promising pathway toward resource recycling and sustainable agriculture.

In paddy ecosystems, nutrient inputs originate not only from fertilizers and biological residues but also from rainfall and irrigation water. Silica, which often exhibits a negative nutrient balance in paddy fields, is a key element for rice cultivation, including *Zizania latifolia*, and its continuous supply is essential for maintaining crop health and stress resistance. Recent declines in soil silica content and silica input via irrigation water in Japan highlight the growing need for efficient and adaptive silica management strategies.

The results of this study demonstrate that hyperspectral reflectance, combined with appropriate machine learning algorithms, provides an effective and non-destructive means of evaluating silica uptake under field conditions. This approach offers a practical tool for assessing the effects of slag fertilization and optimizing fertilizer application schedules. More broadly, the proposed framework has the potential to support precision nutrient management, improve the sustainable reuse of industrial by-products, and contribute to environmentally sound agricultural practices across a range of cropping systems.

Acknowledgements We thank Mr. Keitaro Koike, Mr. Junji Shinada, Ms. Shiori Yonezawa, and Ms. Ibuki Yamazaki of Shizuoka University for assisting with the plant sampling and measurement.

Author Contribution H.Y., T.I. and A.M. designed the cultivation system; A.Y.N. and R.S. measured the hyperspectral reflectance; A.Y.N. and R.S. analyzed the data and wrote the manuscript.

Declarations

Conflict of Interest The authors declare no conflict of interest.

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

- Ahmed FB (2010) Estimation of leaf nitrogen and silicon using hyperspectral remote sensing. *J Appl Remote Sens* 4(1):043560. <https://doi.org/10.1117/1.3525241>
- Barnes RJ, Dhanoa MS, Lister SJ (1989) Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Appl Spectrosc* 43(5):772–777. <https://doi.org/10.1366/0003702894202201>
- Bijay-Singh V-S, Yadvinder-Singh, Thind HS, Ajay-Kumar, Saterpal-Singh, Choudhary OP, Gupta RK, Vashistha M (2013) Supplementing fertilizer nitrogen application to irrigated wheat at maximum tillering stage using chlorophyll meter and optical sensor. *Agric Res* 2(1):81–89. <https://doi.org/10.1007/s40003-013-0053-y>
- Češnovar M, Traven K, Horvat B, Ducman V (2019) The potential of ladle slag and electric arc furnace slag use in synthesizing alkali activated materials; the influence of curing on mechanical properties. *Mater* 12(7):1173. <https://doi.org/10.3390/ma12071173>
- Chang CW, Laird DA, Mausbach MJ, Hurburgh CR (2001) Near-infrared reflectance spectroscopy-principal components regression analyses of soil properties. *Soil Sci Soc Am J* 65(2):480–490. <https://doi.org/10.2136/sssaj2001.652480x>
- Chen D, Meng ZW, Chen YP (2019) Toxicity assessment of molybdenum slag as a mineral fertilizer: a case study with pak-choi (*Brassica Chinensis* L). *Chemosphere* 217:816–824. <https://doi.org/10.1016/j.chemosphere.2018.10.216>
- Chen T, He T (2014) Xgboost: Extreme Gradient Boosting. *R Lecture* 2016. <https://doi.org/10.32614/CRAN.package.xgboost>
- Dabek A, Mantovani L, Mirabella S, Vignati M, Cinquemani S (2025) Advancements in Non-Destructive Detection of Biochemical Traits in Plants Through Spectral Imaging-Based Algorithms: A Systematic Review. *Algorithms* 18:255. <https://doi.org/10.3390/a18050255>
- Das S, Kim GW, Hwang HY, Verma PP, Kim PJ (2019) Cropping with slag to address soil, environment, and food security. *Front Microbiol* 10:1320. <https://doi.org/10.3389/fmicb.2019.01320>
- Du J, Liu B, Zhao T, Xu X, Lin H, Ji Y, Li Y, Li Z, Lu C, Li P, Zhao H, Li Y, Yin Z, Ding X (2022) Silica nanoparticles protect rice against biotic and abiotic stresses. *J Nanobiotechnol* 20(1):197. <https://doi.org/10.1186/s12951-022-01420-x>
- Farooq M, Gao S, Hassan M, Huang Z, Rasheed A, Hearne S, Prasanna B, Li X, Li H (2024) Artificial intelligence in plant breeding. *Trends Genet* 40:891–908. <https://doi.org/10.1016/j.tig.2024.07.001>
- Fallah A (2012) Silicon effect on lodging parameters of rice plants under hydroponic culture. *Int J AgriSci* 2(7):630–634
- Fauteux F, Remus-Borel W, Menzies JG, Belanger RR (2005) Silicon and plant disease resistance against pathogenic fungi. *FEMS Microbiol Lett* 249(1):1–6. <https://doi.org/10.1016/j.femsle.2005.06.034>
- Feret JB, Francois C, Asner GP, Gitelson AA, Martin RE, Bidet LPR, Ustin SL, le Maire G, Jacquemoud S (2008) PROSPECT-4 and 5: Advances in the leaf optical properties model separating photosynthetic pigments. *Remote Sens Environ* 112(6):3030–3043. <https://doi.org/10.1016/j.rse.2008.02.012>
- Fernandez-Delgado M, Sirsat MS, Cernadas E, Alawadi S, Barro S, Febrero-Bande M (2019) An extensive experimental survey of regression methods. *Neural Netw* 111:11–34. <https://doi.org/10.1016/j.neunet.2018.12.010>
- Ishwaran H, Kogalur UB (2007) Random Survival Forests for R. *New Funct Multivar Anal* 7(2):25–31
- Kawamura K, Nishigaki T, Andriamananjara A, Rakotonindrina H, Tsujimoto Y, Moritsuka N, Rabenarivo M, Razafimbelo T

- (2021) Using a one-dimensional convolutional neural network on visible and near-infrared spectroscopy to improve soil phosphorus prediction in Madagascar. *Remote Sens* 13(8):1519. <https://doi.org/10.3390/rs13081519>
18. Khonina SN, Kazanskiy NL, Oseledets IV, Nikonorov AV, Butt MA (2024) Synergy between Artificial Intelligence and Hyperspectral Imaging-A Review. *Technologies* 12:163. <https://doi.org/10.3390/technologies120901631>
 19. Kira O, Linker R, Gitelson A (2015) Non-destructive estimation of foliar chlorophyll and carotenoid contents: Focus on informative spectral bands. *Int J Appl Earth Obs Geoinf* 38:251–260. <https://doi.org/10.1016/j.jag.2015.01.003>
 20. Kuhn M, Quinlan R (2025) Cubist: Rule- And Instance-Based Regression Modeling. R package version 0.5.0.9000. <https://github.com/topepo/cubist>
 21. Liu Y, Li B, Chai X, Zhang M (2025) Characteristics of Municipal Solid-Waste Incineration Fly Ash with High-Temperature Melting and Vitrification. *J Environ Eng* 151. <https://doi.org/10.1061/joeedu.eeeng-7991>
 22. Luyckx M, Hausman JF, Lutts S, Guerriero G (2017) Silicon and plants: Current knowledge and technological perspectives. *Front Plant Sci* 8:411. <https://doi.org/10.3389/fpls.2017.00411>
 23. Miao W, Li F, Lu J, Wang D, Chen M, Tang L, Xu Z, Chen W (2023) Biochar application enhanced rice biomass production and lodging resistance via promoting co-deposition of silica with hemicellulose and lignin. *Sci Total Environ* 855:158818. <https://doi.org/10.1016/j.scitotenv.2022.158818>
 24. Nofrizal AY, Sonobe R, Yamashita H, Seki H, Mihara H, Morita A, Ikka T (2022) Evaluation of a one-dimensional convolution neural network for chlorophyll content estimation using a compact spectrometer. *Remote Sens* 14(9):1997. <https://doi.org/10.3390/rs14091997>
 25. Pan L, Lu R, Zhu Q, Tu K, Cen H (2016) Predict compositions and mechanical properties of sugar beet using hyperspectral scattering. *Food Bioproc Tech* 9(7):1177–1186. <https://doi.org/10.1007/s11947-016-1710-5>
 26. Pioro LS, Pioro IL (2004) Reprocessing of metallurgical slag into materials for the building industry. *Waste Manag* 24(4):371–379. [https://doi.org/10.1016/S0956-053X\(03\)00071-0](https://doi.org/10.1016/S0956-053X(03)00071-0)
 27. Quigley KM, Althoff AG, Donati GL (2016) Inductively coupled plasma optical emission spectrometry as a reference method for silicon estimation by near infrared spectroscopy and potential application to global-scale studies of plant chemistry. *Microchem J* 129:231–235. <https://doi.org/10.1016/j.microc.2016.06.028>
 28. Rahardjo BT, Achadian EMA, Taufiqurrahman AF, Hidayat MR (2021) Silica fertilizer (Si) enhances sugarcane resistance to the sugarcane top borer scirpophaga excerptalis walker. *Agrivita* 43(1):37–42. <https://doi.org/10.17503/agrivita.v1i1.2654>
 29. Ravikanth L, Chelladurai V, Jayas DS, White NDG (2016) Detection of broken kernels content in bulk wheat samples using near-infrared hyperspectral imaging. *Agric Res* 5(3):285–292. <https://doi.org/10.1007/s40003-016-0227-5>
 30. R Core Team (2021) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing
 31. Ribeiro da Luz B, Crowley JK (2007) Spectral reflectance and emissivity features of broad leaf Plants: prospects for remote sensing in the thermal infrared (8.0–14.0 μm). *Remote Sens Environ* 109(4):393–405. <https://doi.org/10.1016/j.rse.2007.01.008>
 32. Sakumotsubunsekihouiinkai (1975) Analytical Methods for Cultivated Plants for Nutritional Diagnosis. Yokendo, pp 141–142
 33. Schlemmer M, Gitelson A, Schepers J, Ferguson R, Peng Y, Shanahana J, Rundquist D (2013) Remote estimation of nitrogen and chlorophyll contents in maize at leaf and canopy levels. *Int J Appl Earth Obs Geoinf* 25(1):47–54. <https://doi.org/10.1016/j.jag.2013.04.003>
 34. Serjun VZ, Mirtič B, Mladenović A (2013) Evaluation of ladle slag as a potential material for building and civil engineering. *Mater Tehnol* 47(5):543–550
 35. Smis A, Murguzur FJA, Struyf E, Soininen EM, Jusdado JGH, Meire P, Brathen KN (2014) Determination of plant silicon content with near infrared reflectance spectroscopy. *Front Plant Sci* 5:496. <https://doi.org/10.3389/fpls.2014.00496>
 36. Snoek J, Ripped O, Swersky K, Kiros R, Satish N, Sundaram N, Patwary MMA, Prabhat, Adams RP (2015) Scalable bayesian optimization using deep neural networks. *ICML'15: Proceedings of the 32nd International Conference on International Conference on Machine Learning* 37:2171–218
 37. Song D, Qiao L, Gao D, Li S, Li M, Sun H, Ma J (2021) Development of crop chlorophyll detector based on a type of interference filter optical sensor. *Comput Electr Agric* 187:106260. <https://doi.org/10.1016/j.compag.2021.106260>
 38. Sonobe R, Hirono Y (2023) Applying variable selection methods and preprocessing techniques to hyperspectral reflectance data to estimate tea cultivar chlorophyll content. *Remote Sens* 15(1):19. <https://doi.org/10.3390/rs15010019>
 39. Sonobe R, Yamashita H, Mihara H, Morita A, Ikka T (2020) Estimation of leaf chlorophyll a, b and carotenoid contents and their ratios using hyperspectral reflectance. *Remote Sens* 12(19):3265. <https://doi.org/10.3390/rs12193265>
 40. Sonobe R, Yamashita H, Mihara H, Morita A, Ikka T (2021) Hyperspectral reflectance sensing for quantifying leaf chlorophyll content in wasabi leaves using spectral pre-processing techniques and machine learning algorithms. *Int J Remote Sens* 42(4):1311–1329. <https://doi.org/10.1080/01431161.2020.1826065>
 41. Sonobe R, Yamashita H, Nofrizal AY, Seki H, Morita A, Ikka T (2022) Use of spectral reflectance from a compact spectrometer to assess chlorophyll content in *Zizania Latifolia*. *Geocarto Int* 37(18):5363–5375. <https://doi.org/10.1080/10106049.2021.1914747>
 42. Sonobe R, Hirono Y (2023) Carotenoid content estimation in tea leaves using noisy reflectance data. *Remote Sens* 15(17):4303. <https://doi.org/10.3390/rs15174303>
 43. Tsuchiya Y, Yoshida K, Ishiguro Y, Kawaii J, Yamashita H, Ikka T, Sonobe R (2026) Optimizing chlorophyll content prediction in tea leaves via spectral transformations and deep learning. *BMC Plant Biol*. <https://doi.org/10.1186/s12870-025-07863-2>
 44. Tsuchiya Y, Hirono Y, Seki H, Sonobe R (2025) Chlorophyll and carotenoid content estimation using hyperspectral reflectance data from a compact spectrometer and fractal-order derivative analysis. *Int J Remote Sens* 46:7258–7273. <https://doi.org/10.1080/01431161.2025.2555281>
 45. Tsuchiya Y, Hirono Y, Sonobe R (2025) Estimating chlorophyll content in tea leaves using spectral reflectance and deep learning methods. *Ecol Inf* 91:103399. <https://doi.org/10.1016/j.ecoinf.2025.103399>
 46. Tong A, He Y (2017) Estimating and mapping chlorophyll content for a heterogeneous grassland: Comparing prediction power of a suite of vegetation indices across scales between years. *ISPRS J Photogramm Remote Sens* 126:146–167. <https://doi.org/10.1016/j.isprsjprs.2017.02.010>
 47. Udaya BK (2023) Package ‘randomForestSRC’ - Fast Unified Random Forests for Survival, Regression, and Classification (RF-SRC). <https://doi.org/10.32614/CRAN.package.randomForestSRC>
 48. Vu Q, Dossa GS, Mundaca EA, Settele J, Crisol-Martinez E, Horgan FG (2022) Combined effects of soil silicon and host plant resistance on planthoppers, blast and bacterial blight in tropical rice. *Insects* 13(7):607. <https://doi.org/10.3390/insects13070604>
 49. Wang Y, Tao H, Yu Q, He Y, Li Y, Wang Z (2025) Co-gasification and melting behavior of municipal solid waste incineration fly

- ash and rice husk. *Fuel* 398:15. <https://doi.org/10.1016/j.fuel.2025.135548>
50. Wehrhan M, Puppe D, Kaczorek D, Sommer M (2021) Spatial patterns of aboveground phytogenic Si stocks in a grass-dominated catchment - Results from UAS-based high-resolution remote sensing. *Biogeosciences* 18(18):5163–5183. <https://doi.org/10.5194/bg-18-5163-2021>
51. Williams P, Norris K (2001) Near-infrared technology in the agricultural and food industries. *American Association of Cereal Chemist*
52. Wu B, Ye H, Huang W, Wang H, Luo P, Ren Y, Kong w (2021) Monitoring the vertical distribution of maize canopy chlorophyll content based on multi-angular spectral data. *Remote Sens* 13(5):987. <https://doi.org/10.3390/rs13050987>

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