50 YEARS
eISSN 1981-1829
Ciência e Agrotecnologia
Science and Agrotechnology

Proximal hyperspectral sensing and PLSR for non-destructive estimation of leaf chlorophyll content in citrus during the fruit ripening

Sensoriamento hiperespectral proximal e PLSR para a estimativa não destrutiva do teor de clorofila foliar em citros durante o amadurecimento dos frutos

Mesut Çoşlu[1*], Namık Kemal Sönmez[2]

ABSTRACT

Chlorophyll content is a key indicator of plant physiological status and photosynthetic activity; therefore, its rapid, non-destructive, and cost-effective assessment is of considerable importance for crop monitoring and management. This study aims to estimate the chlorophyll content of Satsuma mandarin and Valencia orange leaves using five wavelengths (blue, green, red, red-edge and near infrared) and ten different vegetation indices. The study was conducted during the fruit ripening period of citrus trees. Spectroradiometric and SPAD (The Soil Plant Analysis Development) measurements were performed within the scope of proximal remote sensing. Ten vegetation indices, including LCI, CI-G, GNDVI, NDVI, SR, TGI, VARI, NGRDI, NGBDI, and GLI, were derived from hyperspectral (HS) data obtained through spectroradiometric measurements. The statistical analyses estimated SPAD values by Partial Least Squares Regression (PLSR) analysis. According to the findings, the model developed for Satsuma mandarin achieved R^2, RMSE, and MAPE values of 0.89, 1.79, and 2.16%, respectively, while the corresponding values for the Valencia orange were 0.90, 1.60, and 2.01%. These results demonstrate that the relative chlorophyll content during fruit ripening period can be successfully estimated using HS data and the PLSR algorithm. The proposed approach may support rapid chlorophyll monitoring during fruit ripening, a critical period for nutrient management and orchard decision-making, thereby contributing to precision agriculture applications in citrus production.

Index terms: Precision agriculture; proximal sensing; SPAD; vegetation index.

RESUMO

O teor de clorofila é um importante indicador do estado fisiológico das plantas e da atividade fotossintética; portanto, sua avaliação rápida, não destrutiva e de baixo custo é de grande importância para o monitoramento e o manejo das culturas. Este estudo teve como objetivo estimar o teor de clorofila das folhas de tangerineira Satsuma e laranjeira Valencia utilizando cinco comprimentos de onda (azul, verde, vermelho, red-edge e infravermelho próximo) e dez diferentes índices de vegetação. O estudo foi conduzido durante o período de amadurecimento dos frutos dos citros. As medições espectrorradiométricas e de SPAD (The Soil Plant Analysis Development) foram realizadas no contexto do sensoriamento remoto proximal. Dez índices de vegetação, incluindo LCI, CI-G, GNDVI, NDVI, SR, TGI, VARI, NGRDI, NGBDI e GLI, foram derivados de dados hiperespectrais (HS) obtidos por meio de medições espectrorradiométricas. As análises estatísticas estimaram os valores de SPAD por meio da Regressão por Mínimos Quadrados Parciais (PLSR). De acordo com os resultados, o modelo desenvolvido para a tangerineira Satsuma apresentou valores de R^2, RMSE e MAPE de 0,89, 1,79 e 2,16%, respectivamente, enquanto os valores correspondentes para a laranjeira Valencia foram de 0,90, 1,60 e 2,01%. Esses resultados demonstram que o teor relativo de clorofila durante o período de amadurecimento dos frutos pode ser estimado com sucesso utilizando dados hiperespectrais e o algoritmo PLSR. A abordagem proposta pode apoiar o monitoramento rápido da clorofila durante o amadurecimento dos frutos, um período crítico para o manejo nutricional e a tomada de decisões em pomares, contribuindo para aplicações da agricultura de precisão na produção de citros.

Termos para indexação: Agricultura de precisão; sensoriamento remoto proximal; SPAD; índice de vegetação.

Introduction

Remote sensing, defined as the acquisition of information about objects without physical contact, has been increasingly applied in agricultural studies in recent years. In addition to conventional satellite-based systems, advances in sensor technologies have enabled the widespread use of alternative platforms, including proximal remote sensing systems. Proximal remote sensing refers to measurements obtained through sensors placed in contact with or in close proximity to plants. It offers important advantages for plant phenotyping due to its high spatial, spectral, and temporal resolution (Navalgund, Jayaraman & Roy, 2007; Toth & Jóźków, 2016; Tao et al., 2022).

Within proximal remote sensing applications, HS sensors provide high spectral resolution by collecting reflectance information across hundreds of narrow and contiguous bands in

Agricultural Sciences

Ciênc. Agrotec., 50:e006126, 2026
http://dx.doi.org/10.1590/1413-7054202650006126

Editor: Renato Paiva

[1]Burdur Mehmet Akif Ersoy University, Faculty of Arts and Sciences, Department of Precision Agriculture and Agricultural Robots, Burdur, Türkiye
[2]Akdeniz University, Faculty of Science, Department of Space Sciences and Technologies, Antalya, Türkiye

*Corresponding author: mcoslu@mehmetakif.edu.tr

Received in April 23, 2026 and approved in July 16, 2026

All the contents of this journal, except where otherwise noted, is licensed under a Creative Commons Attribution BY.

the electromagnetic spectrum. HS and portable spectroradiometer systems capture reflected or emitted radiation in the ultraviolet, visible, and near-infrared (NIR) regions, enabling detailed characterization of vegetation condition and physiological status (Im & Jensen, 2008; Terentev et al., 2022). Using HS data, some studies have successfully estimated sugar content in apples at different phenological stages, leaf nitrogen content in wheat, and yield in winter wheat (Zhang et al., 2015; Liang et al., 2018).

In citrus research, HS-based studies have predominantly focused on disease detection, including citrus canker and greening, as well as insect-induced damage (Qin et al., 2009; Weng et al., 2018; Abdulridha, Batuman & Ampatzidis, 2019; Feng et al., 2022). Beyond disease-oriented applications, Li et al. (2023) integrated HS imaging with machine learning approaches to evaluate soluble solid content in Tribute Citrus, while Osco et al. (2020) proposed a machine learning framework for estimating nutrient content in Valencia oranges using HS measurements.

Chlorophyll is a key pigment responsible for the green appearance of plants and plays a central role in photosynthesis by absorbing light energy and converting it into biochemical energy (Porcar-Castell et al., 2014; Croft et al., 2017). Although laboratory-based chemical extraction methods provide accurate measurements of chlorophyll concentration, these approaches are destructive and time-consuming (Arnon, 1949; Porra, Thompson & Kriedemann, 1989; Wellburn, 1994; Kamarianakis & Panagiotakis, 2023). Consequently, the development of rapid and non-destructive approaches for leaf chlorophyll assessment has attracted considerable research attention.

Leaf chlorophyll content is widely used as an indicator of plant physiological status and photosynthetic capacity, and portable SPAD meters are commonly employed for rapid, non-destructive measurements (Ling et al., 2011; Zhang et al., 2022; Guermazi et al., 2023). Several studies have combined SPAD measurements with spectral data to estimate relative chlorophyll content. For example, Wu et al. (2023) used multispectral (MS) vegetation indices derived from UAV imagery and machine learning algorithms, including PLSR, to estimate SPAD values in wheat. Yuan et al. (2022) reported significant correlations between SPAD values and RGB- and MS-based vegetation indices for Hopea hainanensis under different shading conditions, while also highlighting multicollinearity issues in spectral datasets. In citrus-specific studies, Garza et al. (2020) emphasized the importance of the Triangular Greenness Index (TGI) derived from RGB imagery for assessing citrus tree health, and Wang et al. (2018) demonstrated significant relationships between SPAD values obtained from different proximal and UAV-based remote sensing systems in mandarin orchards.

Recent studies have highlighted the potential of hyperspectral and multi-source remote sensing data for retrieving vegetation functional traits and biochemical parameters across different ecosystems. For example, Fu et al. (2024) successfully retrieved mangrove functional traits using field hyperspectral and satellite data, while Fu et al. (2025a) developed a mechanism-guided framework for chlorophyll content retrieval by integrating spectral information, leaf area index, and active learning techniques. Similarly, multi-layer remote sensing approaches have been successfully applied for monitoring bread wheat growth and crop condition by integrating data acquired from different sensing platforms, demonstrating the value of combining multiple sources of remotely sensed information for crop monitoring (Sönmez et al., 2026). Furthermore, HS data combined with PLSR and machine learning algorithms have been successfully used for SPAD estimation in wheat under CO_2 stress (Zhang et al., 2024), for evaluating relationships between SPAD and leaf area index using optimal HS band combinations (Aydoğdu & Yıldız, 2024), and for estimating SPAD values of maize using UAV-based MS data and different regression models (Ma et al., 2024). In addition, Jin et al. (2025) reported high prediction accuracies for SPAD estimation in rapeseed by integrating HS imaging with ensemble learning techniques.

In citrus production, the fruit ripening period is a critical phenological stage during which changes in leaf chlorophyll content may reflect plant nutritional status, photosynthetic activity, and overall tree health. Monitoring chlorophyll dynamics during this period can provide valuable information for nutrient management decisions and precision agriculture practices aimed at maintaining fruit quality and orchard productivity. Despite its agronomic importance, chlorophyll monitoring during fruit ripening has received considerably less attention than studies focusing on stress detection, disease assessment, or general vegetation monitoring in citrus orchards.

This study addresses a clear research gap by focusing on non-destructive chlorophyll estimation during the fruit ripening period of citrus, a phenological stage that has received limited attention in HS remote sensing studies. Although SPAD meters provide rapid chlorophyll assessment, their point based and labor-intensive nature limits their applicability in large-scale citrus orchards. Moreover, most existing studies focus on stress-induced chlorophyll variation, whereas the non-destructive estimation of relative chlorophyll content during the fruit ripening period of different citrus species using proximal HS data has received limited attention. Therefore, the objective of this study was to estimate the relative chlorophyll content of Satsuma mandarin (*Citrus unshiu* Marc.) and Valencia orange [*Citrus sinensis* (L.) Osb.] during the fruit ripening period using proximal hyperspectral measurements, vegetation indices, and PLSR.

Material and Methods

Study area

This study was conducted in a field with known ground truth, located at coordinates 36°55’27’’N and 31°00’27’’E, covering an area of approximately 115 hectares, in the Serik

District of Antalya Province (Figure 1). Located 35 km from the city center of Antalya, the area includes not only citrus trees but also collection orchards of avocado, pomegranate, and loquat. Within the citrus parcels of the study area, a total of 60 citrus trees were selected as sample trees, which was considered adequate to represent field conditions while ensuring the practical implementation of proximal HS and SPAD measurements. The study area was selected due to its association with an official agricultural research facility, where field management and data collection are conducted in a systematic and controlled manner.

Material

HS remote sensing enables the determination of spectral characteristics among plant varieties, facilitating tasks such as species differentiation, nutrient content analysis, and yield estimation. Proximal devices like spectroradiometers are commonly used in such studies (Ferreira et al., 2013; Dedeoglu, 2020). In this study, the ASD/FieldSpec®3 Plant Probe+Leaf Clip (Analytical Spectral Devices (ASD), Inc., Boulder, Colorado) was used as the proximal HS sensor for measuring the spectral values of leaves. The ASD FieldSpec HandHeld is a spectroradiometer with a wavelength range of 325-1075 nm, a sampling interval of 1.5 nm (bandwidth), a resolution of 3.5 nm, and scan times as short as 17 ms.

SPAD measurement devices are portable instruments used to measure the relative chlorophyll content in plant leaves. These devices offer several significant advantages, including the ability to perform simple and non-destructive measurements, monitor plant nutrition and health, use in small measurement areas, and have water resistance, which makes them reliably usable in outdoor environments. In this study, the SPAD-502+ (Konica Minolta, Inc., Tokyo 100-7015, Japan) chlorophyll meter was used to measure leaf chlorophyll content in orange trees. The SPAD-502+ is a portable chlorophyll measurement device that provides readings of -9.9 to 199.9 SPAD units. It features two light-emitting diodes and a silicon photodiode receptor for measuring leaf transmittance in the red (650 nm) and NIR (940 nm) regions. The SPAD-502+ chlorophyll meter can store up to 30 measurement values, as well as averaging and deleting measurement values (Konica Minolta, 2022). In this study, SPAD measurements were not considered as an ultimate target variable, but rather as a reliable proxy for relative chlorophyll content, enabling the calibration and validation of spectral models that can be extended to non-contact and spatially continuous monitoring systems. Previous studies have demonstrated strong relationships between SPAD values and chlorophyll concentration in citrus and other woody species (Wang et al., 2025; Qi et al., 2026).

This study was conducted using key phases including leaf sample collection, spectroradiometric measurements and chlorophyll readings (SPAD values), data organization and index calculations, statistical analyses, and the development of estimation models for SPAD values.

Collection of leaf samples, spectroradiometric measurements and SPAD readings

For the Satsuma mandarin in the study area, a total of 240 randomly selected leaf samples were collected on October 12, 2021, with four leaves taken from each tree from four different directions. For the Valencia orange, the samples were collected on March 16, 2022, under the same conditions. Although the fruit set periods of the two citrus species are similar, their fruit ripening periods differ due to species-specific phenological development.

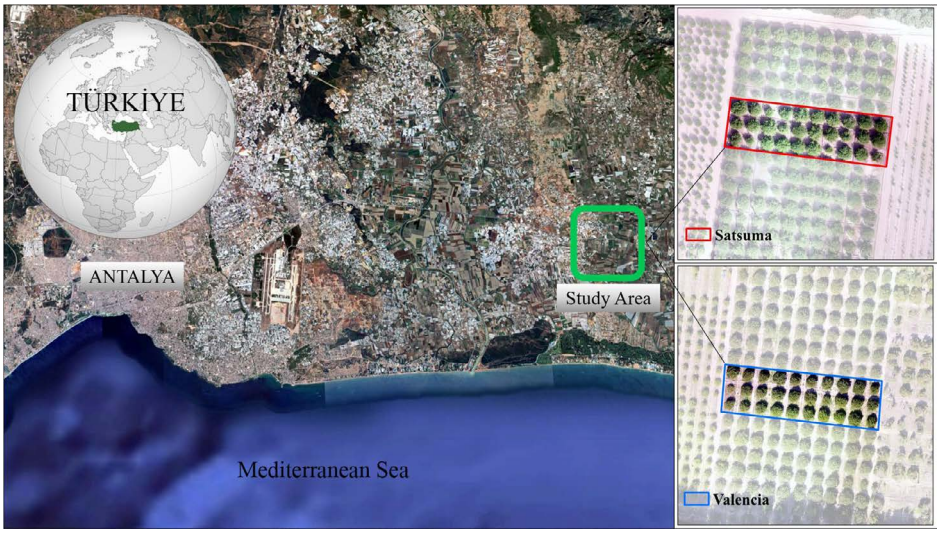


Figure 1: Study area.

Therefore, sampling dates were selected to coincide with comparable fruit ripening stages rather than identical calendar dates, to ensure physiologically consistent leaf chlorophyll conditions. Leaf samples collected from the selected trees were transported to the remote sensing laboratory under cold-chain conditions, where spectroradiometric and SPAD measurements were conducted. The reflectance values of the leaf samples were measured using the ASD/FieldSpec3 Plant Probe+Leaf Clip spectroradiometer and RS3 software. Before starting the spectral readings for each citrus variety, the calibration of the ASD/FieldSpec spectroradiometer was performed using a white reference spectralon block made of gypsum. The spectral values obtained from the HS measurements of the leaves were displayed as graphical data and exported in Excel format using Viewspec software. Before starting readings with the SPAD502+ chlorophyll meter used for measuring leaf chlorophyll values, calibration and a reading checker were used to verify that the measured values fell within the value range on the reading checker. The calibration and reading verification procedures were carried out following the steps outlined in the device's user manual. After the calibration processes, chlorophyll readings were conducted on the same leaves used for spectroradiometric measurements.

Data organization and index calculation

After the spectral measurements of Satsuma mandarin and Valencia orange leaves were conducted using the ASD/FieldSpec3 Plant Probe+Leaf Clip spectroradiometer, a total dataset of 1274 rows was exported and organized in Microsoft Excel. Within this framework, the dataset was first divided into B (400-500 nm), G (501-599 nm), R (600-679 nm), RE (680-750 nm), and NIR (751-900 nm) regions, and the average values for each region were calculated. Subsequently, a total of ten vegetation indices were computed, including five for the visible region indices (TGI, VARI, NGRDI, NGBDI, GLI) and five for the MS region indices (CI-G, LCI, GNDVI, NDVI, SR) (Table 1).

Statistical analyses and SPAD estimation models

In the statistical analysis of the study, normality tests of the datasets were conducted along with correlation and regression analyses. Before performing correlation analyses, descriptive statistics for each variable were calculated, and it was tested whether the data followed a normal distribution. Normality tests for the datasets included skewness-kurtosis coefficients, the Shapiro-Wilk test, and the creation of histograms and QQ plots. For the correlation analysis, Spearman's rank correlation coefficients were computed, as they are used for non-parametric data, particularly when the data does not follow a normal distribution (Ampatzidis et al., 2019; Myers & Sirois, 2006).

Before regression analyses, multicollinearity in the datasets was checked using correlation analysis results and variance inflation factor (VIF) values. PLSR analysis was applied in the development of estimation models in the study. In the PLSR model, which has advantages such as producing a simpler model with fewer factors or components in cases of high multicollinearity among variables, the structures of both the independent and dependent variables are modeled, unlike in traditional regression (Wold, Sjöström & Eriksson, 2001; Nengsih et al., 2019). In the PLSR analysis, the NIPALS algorithm was used, and for factor selection, the leave-one-out (LOO) method, one of the most popular cross-validation techniques, was preferred (Abdi, 2003; Abdi, 2010). In estimating the relative chlorophyll content of Satsuma mandarin and Valencia orange leaves, two different SPAD estimation models were created for each experimental area (Table 2). Model-1 includes all variables (B, G, R, RE, and NIR bands, as well as CI-G, LCI, GNDVI, NDVI, SR, TGI, VARI, NGRDI, NGBDI, and GLI). Model-2 consists of variables that exceeded Wold's criterion value of 0.8 in the PLSR analysis of Model-1, where Wold's criterion was used as a variable selection approach to retain predictors with high explanatory power while reducing model complexity and avoiding overfitting.

To determine the most suitable model obtained from the statistical analyses conducted for estimating SPAD values, the percentages of variance explained by the independent variables, the number of factors, and the predicted residual error sum of squares (RMPRESS) were calculated. Additionally, the R^2 (Equation 1), RMSE (Equation 2), and MAPE (Equation 3) model evaluation parameters, calculated using the equations provided below, were used to assess the fit of the SPAD estimation models to the dataset.

$$R^2 = \frac{(\sum_{i=1}^{n}(y_i - \bar{y})\left(f_i - \bar{f}\right))}{(\sum_{i=1}^{n}(y_i - \bar{y})^2 \sum_{i=1}^{n}\left(f_i - \bar{f}\right)^2)} \quad (1)$$

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

$$MAPE = 100 \times \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - f_i}{y_i}\right| \quad (3)$$

In the equations provided, n represents the number of observations, *yi* represents the actual values, *fi* represents the estimated values, $\bar{y}$ ve $\bar{f}$ are the mean values of the actual and estimated data, respectively. The closer the R^2 value calculated from Equation 1 is to 1, the higher the estimation performance of the model is. The smaller values of RMSE and MAPE, calculated from Equations 2 and 3, indicate that there is less inconsistency between the actual value and the estimated value.

Table 1: Vegetation indices.

Vegetation Indices	Equations	References
Chlorophyll Index-Green	$CI-G=\left(\frac{NIR}{Green}\right)-1$	(Gitelson et al., 2003)
Leaf Chlorophyll Index	$LCI=\left(\frac{(NIR-Red\ Edge)}{(NIR+Red)}\right)$	(Datt et al., 2003)
Green Normalized Difference Vegetation Index	$GNDVI=\left(\frac{(NIR-Green)}{(NIR+Green)}\right)$	(Tucker, 1979)
Normalized Difference Vegetation Index	$NDVI=\left(\frac{(NIR-Red)}{(NIR+Red)}\right)$	(Rouse et al., 1974)
Simple Ratio	$SR=\left(\frac{NIR}{Red}\right)$	(Birth & McVey, 1968)
Triangular Greenness Index	$TGI=\left(Green-(0.39\times Red)-(0.61\times Blue)\right)$	(Hunt et al., 2013)
Visible Atmospheric Resistant Index	$VARI=\left(\frac{(Green-Red)}{(Green+Red-Blue)}\right)$	(Gitelson et al., 2002)
Normalized Green-Red Difference Index	$NGRDI=\left(\frac{(Green-Red)}{(Green+Red)}\right)$	(Tucker, 1979)
Normalized Green-Blue Difference Index	$NGBDI=\left(\frac{(Green-Blue)}{(Green+Blue)}\right)$	(Meyer & Neto, 2008)
Green Leaf Index	$GLI=\left(\frac{2\times Green-Red-Blue}{(2\times Green+Red+Blue)}\right)$	(Louhaichi et al., 2001)

Table 2: SPAD estimation models.

Model	Independent variables	Dependent variable
Model-1	Band average values (B, G, R, RE, NIR) Vegetation index (CI-G, LCI, GNDVI, NDVI, SR, TGI, VARI, NGRDI, NGBDI, GLI)	Chlorophyll Measurement Values (SPAD)
Model-2	Variables greater than Wold's criterion (0.8)	

RESULTS AND DISCUSSION

Since the normality tests indicated that the data did not follow a normal distribution for either citrus species, Spearman's rank correlation coefficients were calculated (Figure 2). For Satsuma mandarin, an examination of the pairwise correlation plots for all variables used in the study revealed high and statistically significant relationships among variables (Figure 2a). Strong correlations were observed between GNDVI and CI-G, as well as between SR and NDVI ($p < 0.0001$). In addition, the highest positive correlation with the dependent variable SPAD was found for the MS index LCI ($\rho = 0.79$), whereas the strongest negative correlation was observed for the RE band ($\rho = -0.79$). Wang et al. (2018) similarly reported significant correlations between SPAD values derived from different remote sensing datasets.

For the Valencia orange experimental field, correlation analysis also revealed strong relationships among the examined variables (Figure 2b). Similar to Satsuma mandarin, high correlations were observed between MS indices such as GNDVI and CI-G, as well as between SR and NDVI ($p < 0.0001$). Notably, the chlorophyll indices CI-G and NDVI, together with SR, exhibited the highest positive correlations ($\rho = 0.98$). Furthermore, the strongest positive and negative correlations with the dependent variable SPAD were associated with the MS index LCI ($\rho = 0.78$) and the RE band ($\rho = -0.86$), respectively. Given the substantial multicollinearity among variables and the non-

normal distribution of the data, a non-parametric PLSR approach was therefore adopted for SPAD estimation in this study.

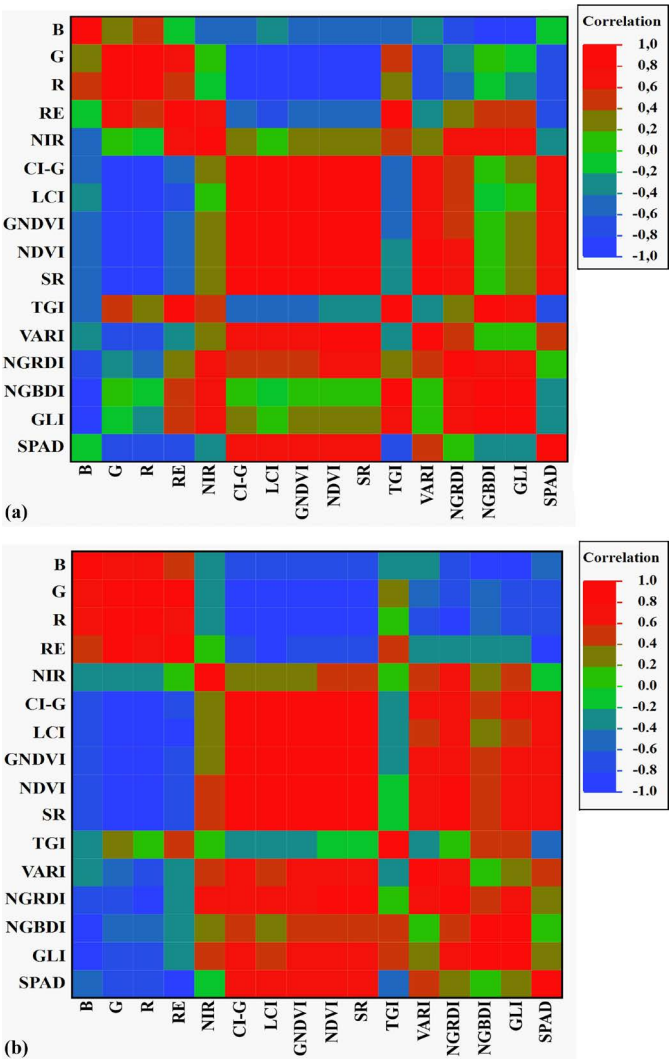


Figure 2: Pairwise correlation matrices for (a) Satsuma mandarin and (b) Valencia orange.

Considering the strong relationships and pronounced multicollinearity observed among the variables, PLSR analysis was employed for SPAD estimation. For the Satsuma mandarin dataset, Model-1 yielded the lowest RMPRESS value (0.46) with five latent factors, and therefore the results of the five-factor model were evaluated. The five-factor PLSR model explained 87.23% of the variation in SPAD values. As illustrated in Figure 3, the highest VIP scores for this model were associated with the RE band, LCI, and TGI. In contrast, the B band, NDVI, NGRDI, and GLI exhibited values below Wold's criterion in Model-1.

For Model-2, comprising 11 variables selected based on Wold's criterion, the lowest RMPRESS value (0.45) was achieved with seven latent factors. The seven-factor PLSR model explained 88.62% of the variance in SPAD values. Consistent with Model-1, the RE band, LCI, and TGI again exhibited the highest VIP values in Model-2. Furthermore, only the NIR band was found to have a value below Wold's criterion in this model. Similarly, Garza et al. (2020) reported that the TGI index derived from RGB imagery was particularly prominent for chlorophyll estimation in citrus using SPAD measurements. The correlation loading plots presented in Figure 3 demonstrate that SPAD values are well captured by both Model-1 and Model-2 and indicate that MS indices are positively associated with SPAD values in both models.

Following the PLSR analysis conducted for Satsuma mandarin, the same modeling framework was applied to the Valencia orange dataset to evaluate its performance under comparable conditions. For Model-1 applied to Valencia oranges, the lowest RMPRESS value (0.64) was obtained with seven latent factors (Figure 4). According to the seven-factor PLSR analysis results, Model-1 was able to explain 86.48% of the variance in relative chlorophyll content. As shown in Figure 4, similar to the results for Satsuma mandarins, the variables with the highest VIP values for this model were the RE band, along with the vegetation indices LCI and TGI. For Model-1, the independent variables B band, NGRDI, and GLI were found to have values below Wold's criterion.

For Model-2, which consists of 12 variables, the PLSR analysis indicated that the lowest RMPRESS value (0.58) was achieved with nine factors. According to the nine-factor PLSR analysis results, the model explained 90.34% of the variance in SPAD values. In this model, where all variables were found to be significant, the highest VIP values were again obtained for the RE band, the MS index LCI, and the RGB-based index TGI, consistent with the results observed for Satsuma mandarin and Model-1. The correlation loading plots presented in Figure 4 further indicate that SPAD values are well explained by both models and that MS indices are positively associated with SPAD values in both cases.

The model evaluation metrics (R^2, RMSE, and MAPE) for the SPAD estimation models developed for Satsuma mandarin and Valencia orange are presented in Figure 5. Furthermore, the observed versus predicted scatter plots and residual distributions for the best-performing models (Model-2) are illustrated in Figure 6. For Satsuma mandarin, Model-2 achieved the highest predictive accuracy, yielding the highest coefficient of determination ($R^2 = 0.89$) along with the lowest RMSE (1.79) and MAPE (2.39%) values. Similarly, for Valencia orange, Model-2 outperformed Model-1, providing the highest R^2 value (0.90) and the lowest RMSE (1.60) and MAPE (2.01%). Although the differences between Model-1 and Model-2 were relatively small for both citrus species, Model-2 achieved slightly higher R^2 values and lower RMSE and MAPE values. Therefore, Model-2 was selected as the preferred model for SPAD estimation in both experimental areas.

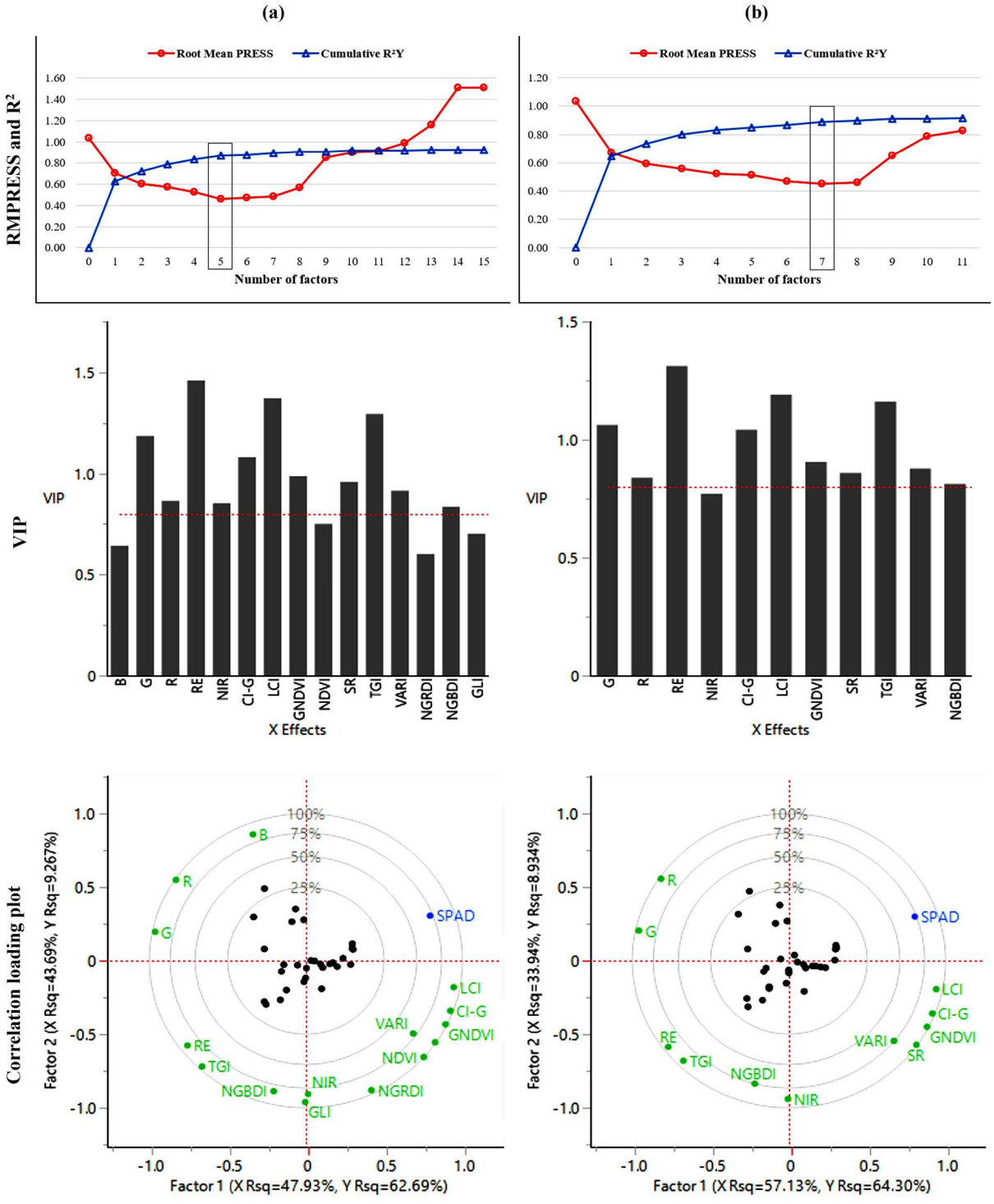


Figure 3: PLSR analysis results for Satsuma mandarin (a) Model-1 and (b) Model-2.

The rapid and non-destructive characterization of spectral responses reflecting plant canopy structure is essential for chlorophyll assessment and plant nutrition studies (Battude et al., 2016; Du & Noguchi, 2017; Qiao et al., 2020). In this study, the spectral behavior of the B, G, R, RE, and NIR bands in Satsuma mandarin and Valencia orange leaves exhibited typical vegetation reflectance patterns, characterized by lower reflectance in the blue and red regions and higher reflectance

in the green, RE, and NIR regions. This behavior is consistent with strong pigment absorption in the blue and red wavelengths and enhanced scattering in the RE and NIR regions, which are well-recognized spectral signatures of healthy vegetation (Qiao et al., 2020). Comparable spectral trends have been reported in maize, cotton, and other crop canopies used for SPAD estimation (Qiao et al., 2022; Yuan et al., 2023). All five spectral bands and ten vegetation indices used in this study showed strong

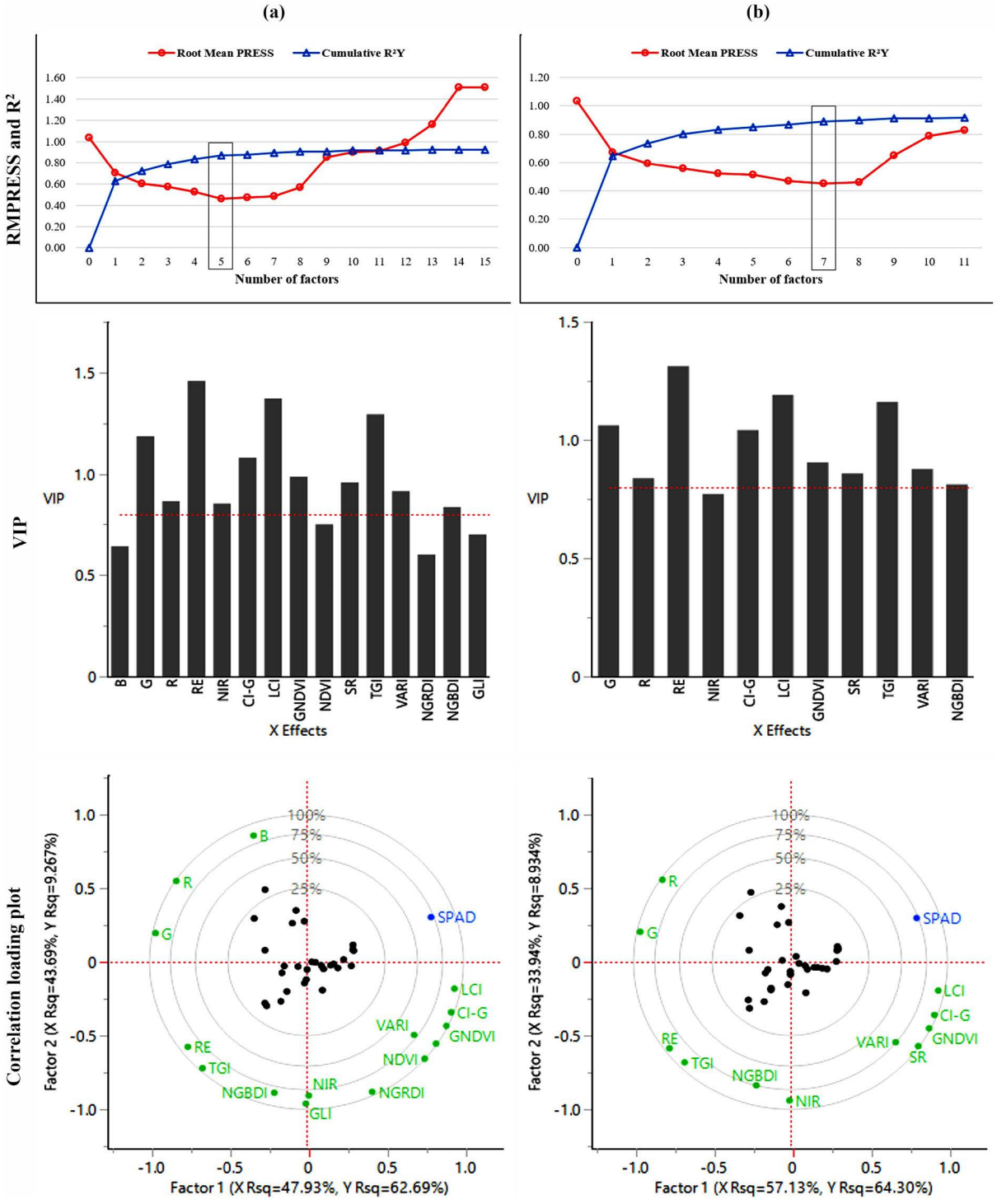


Figure 4: PLSR analysis results for Valencia orange. (a) Model-1 and (b) Model-2.

and significant relationships with SPAD values for both citrus species. Among individual bands, the RE band exhibited the strongest correlation with SPAD, while the LCI from MS indices and the TGI from RGB-based indices showed the highest correlations among vegetation indices. Similar findings have been reported in wheat and citrus-related studies, highlighting the robustness of these indices across growth stages, crops, and sensing platforms (Yin et al., 2023; Wang et al., 2023; Ding et al., 2023). The PLSR-based models developed in this study demonstrated strong predictive performance for both citrus

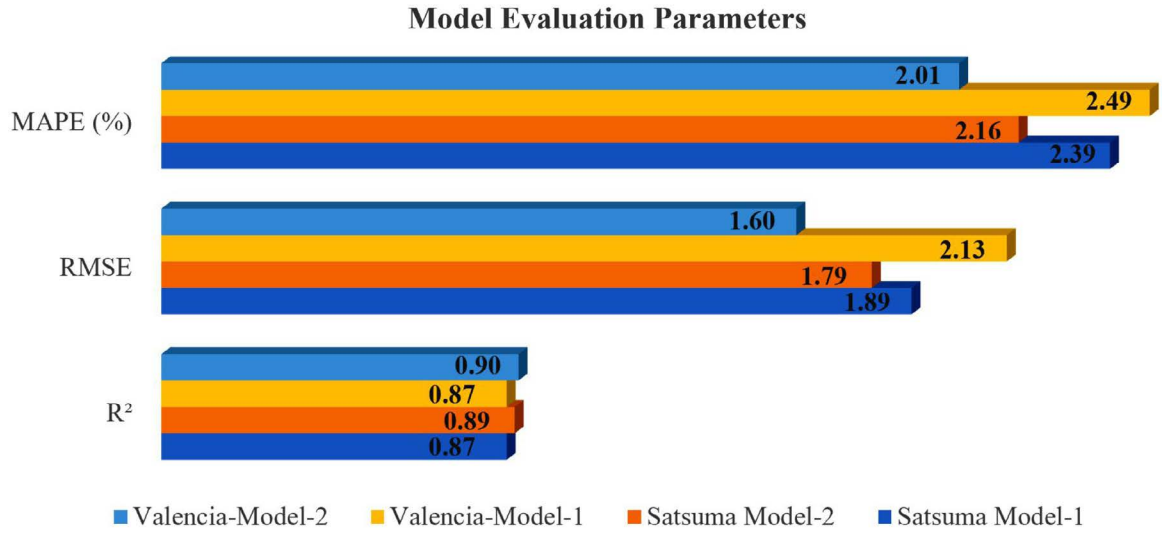


Figure 5: Model evaluation parameters.

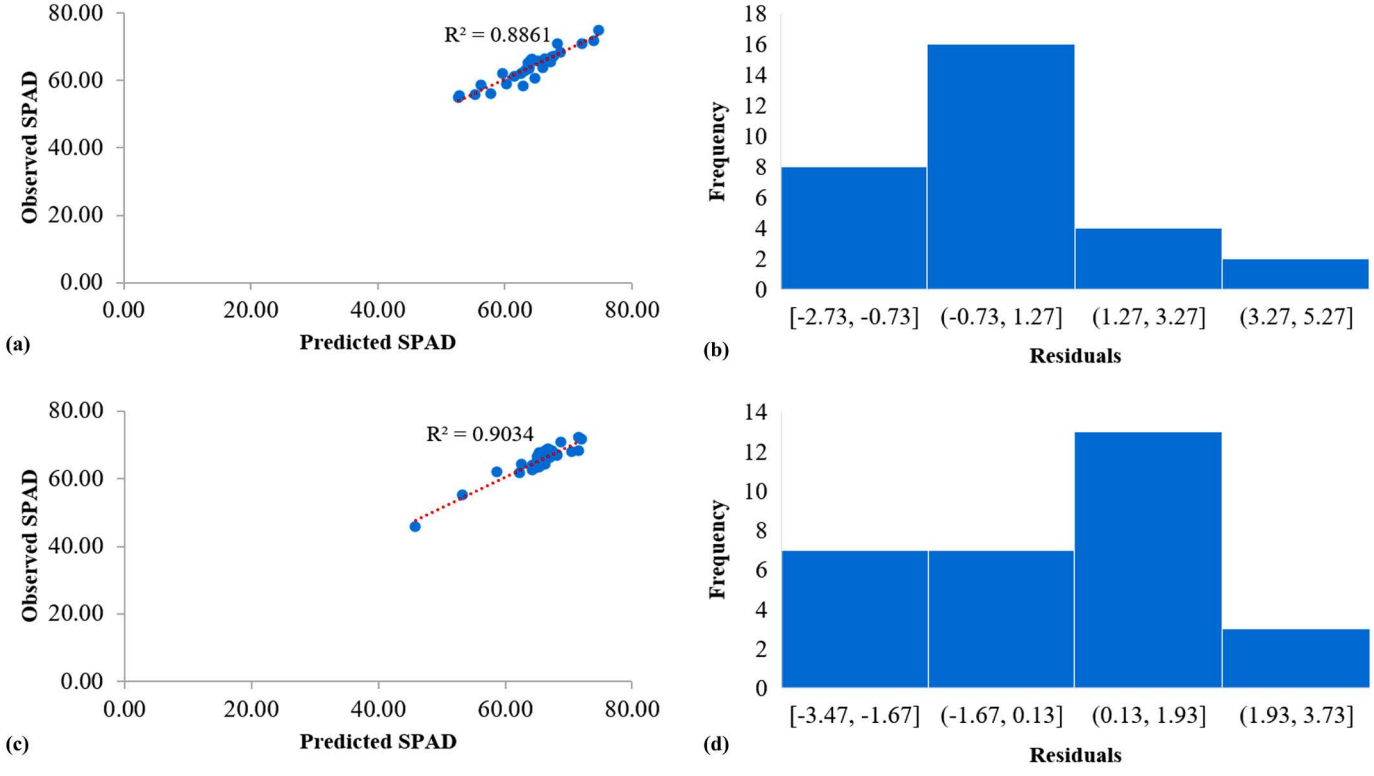


Figure 6: Observed versus predicted scatter plots and residual distributions for the best performing SPAD estimation PLSR models. (a) Scatter plot for Satsuma mandarin; (b) Residual distribution for Satsuma mandarin; (c) Scatter plot for Valencia orange and (d) Residual distribution for Valencia orange.

species. For Satsuma mandarins, the explained variance ranged from 87.23% to 88.62%, while for Valencia oranges it increased from 86.48% to 90.34% with a higher number of latent factors. Across all models, the RE band, LCI, and TGI consistently exhibited high variable importance (VIP > 0.8), confirming their dominant contribution to SPAD estimation. In contrast, the blue band and the GLI showed relatively lower influence, and several variables fell below Wold's criterion depending on species and model structure. The retention of all variables above the threshold in Valencia orange Model-2 suggests a more complex spectral–chlorophyll relationship compared to Satsuma mandarins. Although individual vegetation indices showed strong correlations with SPAD values, the PLSR approach allowed the integration of multiple spectral variables into a single predictive framework, potentially providing a more comprehensive representation of chlorophyll-related spectral responses. Differences between species were attributed to physiological and structural variations, such as leaf thickness, pigment concentration, and internal leaf architecture, which influence spectral reflectance behavior. The higher number of latent factors required for Valencia oranges likely reflects greater heterogeneity in leaf structure and pigment distribution. Leaf sampling was conducted during the fruit ripening stage for both species, ensuring physiologically comparable conditions despite differences in calendar dates. Previous studies indicate that chlorophyll content remains relatively stable during this period, supporting the reliability of phenology-based sampling for SPAD modeling. Increasing the number of latent factors resulted in moderate performance improvements, particularly for Valencia oranges, suggesting that additional spectral information contributes incrementally to species with more complex spectral responses. The strong sensitivity of SPAD values to the red-edge region is consistent with its established role in chlorophyll estimation within the 670–760 nm range (Tanaka et al., 2015; Shi et al., 2023; Tang et al., 2022). The red-edge region represents the transition between strong chlorophyll absorption in the red wavelengths and high reflectance in the near-infrared region, making it particularly sensitive to variations in leaf chlorophyll concentration. This physiological relationship explains why the red-edge band consistently emerged as one of the most influential predictors in the PLSR models. Similarly, the strong performance of the LCI can be attributed to its ability to combine information from chlorophyll-sensitive spectral regions while minimizing the influence of background effects and illumination variability. Because LCI incorporates spectral information from the red-edge and near-infrared regions, it has been widely recognized as a reliable indicator of leaf chlorophyll status and photosynthetic activity. Among the RGB-based indices, TGI exhibited particularly high importance values. The TGI is designed to estimate chlorophyll-related changes by describing the geometric relationship between reflectance in the green, red, and blue bands. Since chlorophyll strongly absorbs red and blue radiation while reflecting more green light, TGI is sensitive to variations in leaf greenness and pigment concentration. Its strong contribution in this study highlights the potential of low-cost RGB sensors as practical alternatives for chlorophyll-related assessments in citrus orchards. Beyond the expected dominance of the red-edge region, the consistent contribution of RGB-based indices, especially TGI and VARI, further highlights the potential of low-cost RGB sensors for chlorophyll-related assessments (Garza et al., 2020; Yuan et al., 2022). On a broader scale, the non-destructive estimation of vegetation parameters has increasingly benefited from advanced, multi-source remote sensing methodologies. While proximal hyperspectral sensing offers high precision for physiological and biochemical traits like chlorophyll at the leaf scale, broader structural and biophysical parameters, such as canopy height and aboveground biomass, are effectively retrieved by integrating multi-source aerospace data, including UAV-LiDAR, Synthetic Aperture Radar (SAR), and optical stereo satellite imagery (Fu et al., 2025b). Contextualizing leaf-level spectral estimations within these multi-scale and multi-source frameworks highlights the expanding and complementary capabilities of remote sensing technologies in comprehensive vegetation monitoring. Overall, these findings demonstrate that multivariate evaluation of spectral bands and vegetation indices provides a robust and transferable framework for SPAD estimation across different citrus species and sensing approaches. Although the LOO approach enabled efficient use of the available observations and is commonly applied in studies with limited sample sizes, it may provide slightly optimistic estimates of model performance when compared with independent validation datasets. Therefore, the predictive accuracies reported in this study should be interpreted within the context of the available dataset. Future research should evaluate the robustness and transferability of the developed models using independent datasets collected from different orchards, growing seasons, citrus cultivars, and environmental conditions.

Conclusions

This study successfully developed robust PLSR models (Model-2) to accurately estimate leaf SPAD values in Satsuma mandarin and Valencia orange cultivars, achieving R^2 values of 0.89 and 0.90, respectively. However, important limitations must be acknowledged, including the exclusive reliance on a single-season dataset and limited sample diversity. These factors may restrict the direct transferability of the models to other citrus varieties and growing environments. Future studies should incorporate multi-seasonal and geographically diverse datasets to enhance model robustness and generalizability.

Author contribution

Conceptual idea: Çoşlu, M; Sönmez, N.K.S; Methodology design: Çoşlu, M; Sönmez, N.K.S; Data collection: Çoşlu, M; Sönmez, N.K.S; Data analysis and interpretation: Çoşlu, M; Writing and editing: Çoşlu, M; Sönmez, N.K.S

Acknowledgements

This study utilized a subset of data collected during the author's doctoral dissertation; however, the analyses and interpretations were conducted independently.

Data availability statement

Data available upon request to authors

References

Abdi, H. (2003). Partial least squares regression (PLS-regression). In M. Lewis-Beck., A. Bryman., & T. Futing. *Encyclopedia for research methods for the social sciences*. Thousand Oaks, CA: Sage. (pp.792-795).

Abdi, H. (2010). Partial least squares regression and projection on latent structure regression (PLS Regression). *WIREs Computational Statistics*, *2*:97-106.

Abdulridha, J., Batuman, O., & Ampatzidis, Y. (2019). UAV-based remote sensing technique to detect citrus canker disease utilizing hyperspectral imaging and machine learning. *Remote Sensing, 11*(11):1373.

Ampatzidis, Y. et al. (2019). Citrus rootstock evaluation utilizing UAV-based remote sensing and artificial intelligence. *Computers and Electronics in Agriculture*, *164*:104900.

Arnon, D. I. (1949). Copper enzymes in isolated chloroplasts. Polyphenoloxidase in *Beta vulgaris*. *Plant Physiology*, *24*(1):1-15.

Aydoğdu, M., & Yıldız, H. (2024). Use of hyperspectral data for chlorophyll estimation based on leaf area index (LAI) in wheat. *Turkish Journal of Remote Sensing*, *6*(2):97-111.

Battude, M. et al. (2016). Estimating maize biomass and yield over large areas using high spatial and temporal resolution Sentinel–2 like remote sensing data. *Remote Sensing of Environment, 184*:668-681.

Birth, G. S., & McVey, G. (1968). Measuring the color of growing turf with a reflectance spectroradiometer. *Agronomy Journal*, *60*(6):640-643.

Croft, H. et al. (2017). Leaf chlorophyll content as a proxy for leaf photosynthetic capacity. *Global Change Biology*, *23*(9):3513-3524.

Datt, B. et al. (2003). Preprocessing eo-1 hyperion hyperspectral data to support the application of agricultural indexes. *IEEE Transactions on Geoscience and Remote Sensing*, *41*(6):1246-1259.

Dedeoglu, M. (2020). Estimation of critical nitrogen contents in peach orchards using visible-near infrared spectral mixture analysis. *Journal of Near Infrared Spectroscopy*, *28*(5-6):315-327.

Ding, S. et al. (2023). Citrus canopy SPAD prediction under bordeaux solution coverage based on texture- and spectral-information fusion. *Agriculture*, *13*(9):1701.

Du, M., & Noguchi, N. (2017). Monitoring of wheat growth status and mapping of wheat yield's within-field spatial variations using color images acquired from UAV-camera system. *Remote Sensing*, *9*(3):289.

Feng, H. et al. (2022). Comparison of winter wheat yield estimation based on near-surface hyperspectral and UAV hyperspectral remote sensing data. *Remote Sensing*, *14*(17):4158.

Ferreira, M. P. et al. (2013). Analyzing the spectral variability of tropical tree species using hyperspectral feature selection and leaf optical modeling. *Journal of Applied Remote Sensing*, *7*(1):073502.

Fu, B. et al. (2024). Synergistic retrieval of mangrove vital functional traits using field hyperspectral and satellite data. *International Journal of Applied Earth Observation and Geoinformation*, *131*:103963.

Fu, B. et al. (2025a). A novel mechanism-guided retrieval framework for mangrove chlorophyll content considering leaf area index and spectral resolution based on active learning hybrid model. *International Journal of Applied Earth Observation and Geoinformation*, *144*:104926.

Fu, B. et al. (2025b). Estimation of mangrove heights and aboveground biomass using UAV-LiDAR, Sentinel-1 and ZY-3 stereo images. *Ecological Informatics*, *88*:103160.

Garza, B. N. et al. (2020). Quantifying citrus tree health using true color UAV images. *Remote Sensing*, *12*(1):170.

Gitelson, A. A. et al. (2002). Vegetation and soil lines in visible spectral space: A concept and technique for remote estimation of vegetation fraction. *International Journal of Remote Sensing, 23*(13):2537-2562.

Gitelson, A. A. et al. (2003). Remote estimation of leaf area index and green leaf biomass in maize canopies. *Geophysical Research Letters*, *30*(5):1248.

Guermazi, E. et al. (2023). Combining remote sensing, SPAD readings, and laboratory analysis for monitoring olive groves and olive oil quality. *Precision Agriculture*, *25*:65-82.

Hunt, E. R. et al. (2013). A visible band index for remote sensing leaf chlorophyll content at the canopy scale. *International Journal of Applied Earth Observation and Geoinformation*, *21*:103-112.

Im, J., & Jensen, J. R. (2008). Hyperspectral remote sensing of vegetation. *Geography Compass*, 2(6):1943-1961.

Jin, Y. et al. (2025). Combining hyperspectral imaging with ensemble learning for estimating rapeseed chlorophyll content under different waterlogging durations. *Plants*, *14*(24):3713.

Kamarianakis, Z., & Panagiotakis, S. (2023). Design and implementation of a low-cost chlorophyll content meter. *Sensors*, *23*(5):2699.

Konica Minolta (2022). Non-destructive measurement of plant chlorophyll to optimise fertiliser usage. Available in: <https://www.konicaminolta.eu/eu-en/hardware/measuring-instruments/colour-measurement/chlorophyll-meter/spad-502plus>. (accessed on 19 November 2022).

Li, C. et al. (2023). Hyperspectral imaging with machine learning approaches for assessing soluble solids content of tribute citrus. *Foods*, *12*(2):247.

Liang, L. et al. (2018). Estimation of leaf nitrogen content in wheat using new hyperspectral indices and a random forest regression algorithm. *Remote Sensing*, *10*(12):1940.

Ling, Q. et al. (2011). Use of a SPAD-502 meter to measure leaf chlorophyll concentration in *Arabidopsis thaliana*. *Photosynthesis research*, *107*(2):209-214.

Louhaichi, M. et al. (2001). Spatially located platform and aerial photography for documentation of grazing impacts on wheat. *Geocarto International*, *16*(1):65-70.

Ma, W. et al. (2024). UAV multispectral remote sensing for the estimation of SPAD values at various growth stages of maize under different irrigation levels. *Computers and Electronics in Agriculture*, *227*:109566.

Meyer, G. E., & Neto, J. C. (2008). Verification of color vegetation indices for automated crop imaging applications. *Computers and Electronics in Agriculture*, *63*(2):282-293.

Myers, L., & Sirois, M. J. (2006). Spearman correlation coefficients, differences between. *Encyclopedia of Statistical Sciences*, *12*.

Navalgund, R. R., Jayaraman, V., & Roy, P. S. (2007). Remote sensing applications: An overview. *Current Science*, *93*(12):1747-1766.

Nengsih, T. et al. (2019). Determining the number of components in PLS regression on incomplete data set. *Statistical Applications in Genetics and Molecular Biology*, *18*(6):20180059.

Osco, L. P. et al. (2020). A machine learning framework to predict nutrient content in Valencia-Orange Leaf hyperspectral measurements. *Remote Sensing*, *12*(6):906.

Porcar-Castell, A. et al. (2014). Linking chlorophyll a fluorescence to photosynthesis for remote sensing applications: Mechanisms and challenges. *Journal of Experimental Botany*, *65*(15):4065-4095.

Porra, R. J., Thompson, W. A., & Kriedemann, P. E. (1989). Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls a and b extracted with four different solvents: Verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, *975*(3):384-394.

Qi, X. et al. (2026). Estimation of citrus leaf relative water content using CWT combined with chlorophyll-sensitive bands. *Sensors*, *26*(2):467.

Qiao, L. et al. (2020). Dynamic influence elimination and chlorophyll content diagnosis of maize using UAV spectral imagery. *Remote Sensing*, *12*(16):2650.

Qiao, L. et al. (2022). UAV-based chlorophyll content estimation by evaluating vegetation index responses under different crop coverages. *Computers and Electronics in Agriculture*, *196*:106775.

Qin, J. et al. (2009). Detection of citrus canker using hyperspectral reflectance imaging with spectral information divergence. *Journal of Food Engineering*, *93*(2):183-191.

Rouse, J. W. et al. (1974). Monitoring vegetation systems in the Great Plains with ERTS. *NASA Special Publication*, *351*:309-317.

Shi, H. et al. (2023). Estimation of chlorophyll content in soybean crop at different growth stages based on optimal spectral index. *Agronomy*, *13*(3):663.

Sönmez, N. K. et al. (2026). Monitoring bread wheat plants with multi-layer remote sensing approach. *Pakistan Journal of Agricultural Sciences*, *63*(1):213-228.

Tanaka, S. et al. (2015). Spectral index for quantifying leaf area index of winter wheat by field hyperspectral measurements: A case study in Gifu Prefecture, central Japan. *Remote Sensing*, *7*(5):5329-5346.

Tang, Z. et al. (2022). Estimation of leaf area index and above-ground biomass of winter wheat based on optimal spectral index. *Agronomy*, *12*(7):1729.

Tao, H. et al. (2022). Proximal and remote sensing in plant phenomics: 20 years of progress, challenges, and perspectives. *Plant Communications*, *3*(6):100344.

Terentev, A. et al. (2022). Current state of hyperspectral remote sensing for early plant disease detection: A review. *Sensors*, *22*(3):757.

Toth, C., & Jóźków, G. (2016). Remote sensing platforms and sensors: A survey. *ISPRS Journal of Photogrammetry and Remote Sensing*, *115*:22-36.

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, *8*(2):127-150.

Wang, K. J. et al. (2018). Rapid detection of chlorophyll content and distribution in citrus orchards based on low-altitude remote sensing and biosensors. *International Journal of Agricultural & Biological Engineering*, *11*:164-169.

Wang, Q. et al. (2023). UAV hyperspectral data combined with machine learning for winter wheat canopy SPAD values estimation. *Remote Sensing*, *15*(19):4658.

Wang, J. et al. (2025). Estimation of canopy chlorophyll content of apple trees based on UAV multispectral remote sensing ımages. *Agriculture*, *15*(12):1308.

Wellburn, A. R. (1994). The spectral determination of chlorophylls a and b, as well as total carotenoids, using various solvents with spectrophotometers of different resolution. *Journal of Plant Physiology*, *144*(3):307-313.

Weng, H. et al. (2018). Hyperspectral reflectance imaging combined with carbohydrate metabolism analysis for diagnosis of citrus Huanglongbing in different seasons and cultivars. *Sensors and Actuators B: Chemical*, *275*:50-60.

Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, *58*(2):109-130.

Wu, Q. et al. (2023). Estimation of relative chlorophyll content in spring wheat based on multi-temporal UAV remote sensing. *Agronomy*, *13*(1):211.

Yin, Q. et al. (2023). Estimation of winter wheat SPAD values based on UAV multispectral remote sensing. *Remote Sensing*, *15*(14):3595.

Yuan, X. et al. (2023). Hyperspectral estimation of SPAD value of cotton leaves under verticillium wilt stress based on GWO–ELM. *Agriculture*, *13*(9):1779.

Yuan, Y. et al. (2022). Performance comparison of RGB and multispectral vegetation indices based on machine learning for estimating *Hopea hainanensis* SPAD values under different shade conditions. *Frontiers in Plant Science*, *13*:928953.

Zhang, Y. et al. (2015). Predicting apple sugar content based on spectral characteristics of apple tree leaf in different phenological phases. *Computers and Electronics in Agriculture*, *112*:20-27.

Zhang, R. et al. (2022). Evaluation of the methods for estimating leaf chlorophyll content with SPAD chlorophyll meters. *Remote Sensing*, *14*(20):5144.

Zhang, L. et al. (2024). Hyperspectral characteristics and SPAD estimation of wheat leaves under CO2 Microleakage Stress. *Sensors*, *24*(15):4776.