A HYBRID NON-DESTRUCTIVE APPROACH COMBINING IMAGE PROCESSING AND SPECTRAL FEATURE SELECTION FOR GRAPEVINE LEAF WATER CONTENT ESTIMATION

Bruno Ricardo Silva Costa[1], Ednaldo José Ferreira[2], Henrique Oldoni[3], Lúcio André de Castro Jorge[2], Luis Henrique Bassoi[2]

[1] Luiz de Queiroz College of Agriculture - University of São Paulo, Department of Biosystems Engineering, Precision Agriculture Laboratory – Piracicaba, São Paulo State, Brazil

[2] Brazilian Agricultural Research Corporation – Embrapa Instrumentation – São Carlos, São Paulo State, Brazil

[3] Brazilian Agricultural Research Corporation – Embrapa Digital Agriculture – São Carlos, São Paulo State, Brazil

A paper from the Proceedings of the
17th International Conference on Precision Agriculture and 11th Brazilian Congress on Precision and Digital Agriculture
13-15 July 2026
Porto Alegre, Brazil

Abstract.

Reliable and continuous estimation of leaf water content (LWC) is essential for viticulture, as it enables assessment of spatiotemporal variability in vine water demand and supports improved irrigation management efficiency within Precision Agriculture (PA) practices. Although the gravimetric method based on fresh weight (FW) and dry weight (DW) measurements provides accurate LWC estimates, it is time-consuming, destructive, and exhibits limited scalability for large sample sizes. In contrast, leaf area (LA) represents a non-destructive proxy for LWC that can be measured using digital technologies. However, the predictive capability of LA is strongly influenced by grapevine cultivar, given the ampelography diversity, and plant water status dynamics. Furthermore, short-term variations in LWC may occur with minimal changes in LA in fully expanded leaves, limiting the sensitivity of this predictor. Hyperspectral reflectance data offer complementary information on intraspecific variability related to both cultivar-specific leaf characteristics and LWC while preserving the non-destructive advantages of optical sensing. Nevertheless, the high dimensionality of hyperspectral data often leads to model overfitting and reduced generalization capability, emphasizing the need for effective feature selection methods prior to model fitting. This study aimed to develop an integrated approach combining digital image processing for LA estimation and feature selection to identify wavelengths related to water absorption in reflectance spectra and cultivar-specific ampelographic features, enabling a more robust and linear non-destructive model for LWC estimation in grapevines. LWC was

The authors are solely responsible for the content of this paper, which is not a refereed publication. Citation of this work should state that it is from the Proceedings of the 17th International Conference on Precision Agriculture and 11th Brazilian Congress on Precision and Digital Agriculture. EXAMPLE: Last Name, A. B. & Coauthor, C. D. (2026). Title of paper. In Proceedings of the 17th International Conference on Precision Agriculture and 11th Brazilian Congress on Precision and Digital Agriculture (unpaginated, online). Monticello, IL: International Society of Precision Agriculture and Brazilian Congress on Precision and Digital Agriculture.

calculated as the difference between FW and DW. LA was estimated from leaf images acquired using a conventional smartphone camera and processed through a segmentation pipeline based on thresholding and binarization. Leaf reflectance spectra ranging from 350–2500 nm at 1 nm resolution were collected using a field spectroradiometer. Digital LA measurements and corresponding spectral data were subjected to a wrapper-based feature selection method employing a greedy hill-climbing metaheuristic with backtracking in a backward search. Simple and multiple linear regression models were developed to predict LWC using LA alone and LA combined with reflectance at preselected wavelengths, respectively. The methodology was evaluated using pooled datasets obtained from samples collected at harvest from four vineyards over two consecutive seasons, including the cultivars Cabernet Franc and Sauvignon. Model performance was assessed using a holdout validation strategy, with 70% of the data allocated to training (n = 580) and 30% to testing (n = 248). LA alone provided strong predictive capability, explaining approximately 68% of the variability in LWC ($R^2 \approx 0.68$) in the test dataset. However, incorporation of selected spectral features, predominantly from the infrared region, improved model generalization, increasing the R^2 for the test dataset to 0.72 and reducing the relative absolute error from 58.4 to 40.7%. Paired Wilcoxon signed-rank tests applied to the absolute errors associated with the test dataset confirmed that the multiple linear regression model significantly outperformed the LA-only model (p-value < 0.001). These results demonstrate that integration of digital LA measurements with selected spectral features provides a more reliable and cultivar-robust approach for non-destructive LWC estimation, supporting its potential application for guiding water management in precision viticulture.

Keywords.

Vitis vinifera L., digital image, canopy reflectance, combined method.

Introduction

Providing an adequate water supply trough irrigation is essential for viticulture, once vine growth, yield and berry composition is influenced by water availability. Therefore, grapegrowers rely on accurate assessment of grapevine water status to meet these demands through irrigation management strategies, following Precision Agriculture (PA) principles and technologies (González-Fernández et al., 2015, Bahat et al, 2024), i.e., intending to meet the exact plant water requirements at a given moment of the growing season.

Traditional approaches have been practiced to evaluate grapevine water status, particularly leaf water content (LWC), including destructive laboratory procedures. In this sense, the gravimetric determination of LWC through fresh and dry biomass weighting is a reference method. However, despite its accuracy, this procedure requires labor consuming field campaigns for leaf sampling, and laboratory infrastructure for proper leaf oven drying and weighting. It makes this traditional method unsuitable for continuous monitoring of LWC at field scale to support operational decision-making process regarding the estimation of the irrigation depth and time.

In this context, current advances in optical sensing technologies have created opportunities for non-destructive estimation of leaf traits that could be assumed as proxies to estimate LWC. Particularly, digital imaging stands out due to its low cost, operational simplicity, and widespread availability through devices such as smartphones (Guerrero et al., 2012; Müller-Linow et al., 2019). Thus, we hypothesize that these easily available digital images can be processed to estimate leaf morphological traits, particularly leaf area (LA), which is related to leaf biomass accumulation and, consequently, to its water storage capacity. Despite of this correlation, LA alone could be limited to generically estimate LWC in viticulture due to the variations in cultivars ampelographic characteristics and plant water status dynamics. Furthermore, temporary variations in LWC can occur with negligible corresponding changes in LA, especially in mature and full expanded leaves, reducing the sensitivity of morphological predictors to predict water status by themselves.

Concurrently, optical technologies related to hyperspectral reflectance sensing could captures subtle variations in leaf composition and structure across multiple wavelengths, and then could be assumed as additional sources of information to estimate LWC with morphological leaf traits. Reflectance features in the near-infrared (NIR) and shortwave infrared (SWIR) regions are particularly sensitive to water absorption processes and have been widely associated with plant water status (Rapaport et al., 2015; Tardáguila et al, 2017; Laroche-Pinel et al., 2024). Hence, we also hypothesize that these technologies in association with the previous digital images can be useful for LWC estimations, specifically through regression modeling. However, when dealing with such high dimensional data provided by hyperspectral sensing, the large number of evaluated wavelengths often leads to redundancy or overfitting of the proposed models. In this sense, feature selection techniques are thus important for identifying significative wavelengths and improving model generalization intending reliable LWC estimations (Giovenzana et al., 2018).

Therefore, the aim of this study was to develop and evaluate the predictive performance of a composite non-destructive methodology combining LA estimation through digital image processing and hyperspectral feature selection of leaf reflectance data to improve regression-based estimation of LWC across different grapevine cultivars.

Material and methods

Study site and dataset description

The study was conducted using pooled datasets obtained from fully-developed and healthy green leaves collected at harvest in four commercial vineyards over two consecutive growing seasons, which included samples from the grapevine cultivars Cabernet Franc and Sauvignon (*Vitis vinifera* L.). A total of 828 leaf samples were available, with 580 observations allocated to model training and 248 observations reserved for independent testing. The datasets comprising morphological information of the leaves (i.e., fresh and dry weight and LA) and reflectance data

Leaf water status and area estimation

Leaf water content (LWC, in g leaf^{-1}) was calculated using the gravimetric difference between fresh weight (FW) and dry weight (DW), which served as the reference variable for model development. Fresh leaf samples were weighed immediately after collection to determine FW and subsequently oven dried until constant weight to determine DW. The FW−DW difference represented the water content retained in each leaf sample.

Leaf area was estimated using images acquired with a conventional smartphone camera. Images were processed through a segmentation workflow consisting of thresholding and binarization procedures designed to isolate the leaf blade from the background. The number of pixels corresponding to the leaf area was then converted into a quantitative leaf area measurement, generating a non-destructive predictor variable suitable for large-scale data collection. This image processing was performed using ImageJ (Wayne Rasband National Institute of Health, Bethesda, Maryland, USA) version 1.52 and the LeafArea R package (Katabuchi, 2015).

Leaf hyperspectral data acquisition and feature selection

Reflectance spectra were collected using the portable spectroradiometer FieldSpec 3 (Analytical Spectral Devices, Inc.; Boulder, Colorado, USA), covering wavelengths from 350 to 2500 nm with 1 nm spectral resolution. Thus, this spectral range encompasses visible, NIR, and SWIR regions.

We employed a wrapper-based feature selection procedure prior to model fitting due to the large number of highly correlated features in the hyperspectral reflectance data. This process used a greedy hill-climbing search with backtracking for a backward feature elimination. The aim of this data processing was to identify wavelengths that maximized the LWC predictive performance of the models while minimizing redundancy and reducing their complexity.

Regression model fitting and predictive performance evaluation

Two regression approaches were evaluated. The first one consisted of a simple linear regression (SLR) model using LA as the only predictor. The second consisted of a multiple linear regression (MLR) model combining LA with the selected spectral features.

We then used the training dataset for models fitting, and the independent test dataset to evaluate their predictive generalization for LWC estimation. Performance metrics included coefficient of determination (R^2) and error measurements were calculated. Additionally, paired Wilcoxon signed-rank tests were used to compare prediction errors between the two modeling approaches.

Results and discussion

Table 1 summarizes the results regarding the model fitting performance and figure presented the correspondence between observed and predicted values of LWC by each model. The SLR model demonstrated that LA alone already contains important information regarding grapevine LWC. Variations in LA explained approximately 74% of the variability observed in leaf water in the training dataset ($R^2 \approx 0.74$), confirming the robustness of this trait as a predictor for leaf water status that can be obtained through non-destructive digital image acquisition and processing. The results also demonstrated the generalization capability of the model, according to the correlation of 0.8249 in the test dataset, corresponding to approximately 68% of the explained variability in LWC ($R^2 \approx 0.68$). These findings demonstrated that larger leaves generally contain more water, reflecting the expected relationship between leaf size and biomass accumulation.

Prior to the MLR model fitting including the hyperspectral reflectance data, the wrapper-based feature selection identified a set of 36 wavelengths concentrated primarily in the NIR and SWIR regions. Several selected wavelengths were located near known water-sensitive spectral domains, particularly between approximately 1500 and 1900 nm. These regions are commonly associated with absorption features related to water molecules and internal leaf structure, suggesting that the proposed feature selection method was able to capture physically meaningful information related to leaf water status.

Despite the performance of the SLR model, the MLR model integrating LA and the selected spectral features improved LWC predictive performance. The MLR model explained approximately 88% of the variability observed in the training dataset ($R^2 = 0.8814$), indicating that most of the variation in leaf water content was captured by the combined predictors. Independent validation also revealed an improved predictive performance when applying the model to the test dataset, with the correlation coefficient increasing from 0.8249 in the SLR model to 0.8635 in the MLR model. Likewise, in this test dataset, the RAE decreased from 51.38% to 40.68%, while the RRSE decreased from 56.91% to 51.75%, demonstrating improved model generalization and predictive accuracy. Additionally, paired Wilcoxon signed-rank tests applied to the absolute prediction errors indicated significant differences between the models for the training ($p < 0.001$) dataset further supporting the superior prediction performance of the MLR model.

The advantage of the hybrid approach was also confirmed through paired Wilcoxon signed-rank tests, which indicated significantly lower prediction errors for the MLR model. This result demonstrates that spectral information contributes with leaf morphology for LWC predictions. These findings demonstrated that while LA itself represents structural features of leaves that makes it a primary predictor of LWC, hyperspectral reflectance add important information to increase its prediction capacity, once it captures physiological and biochemical variations associated with LWC dynamics.

Table 1. Statistics to evaluate the simple linear regression (SLR) and multiple linear regression (MLR) models fitting performance concerning grapevine leaf water content (LWC) prediction.

Model	Dataset	n	R^2	R^2_{adj}	r	MAE	RMSE	RAE (%)	RRSE (%)
SLR	Training	580	0.7425	0.7421	0.8617	0.2000	0.3000	44.5596	50.7441
	Test	248	-	-	0.8249	0.2000	0.3000	51.3813	56.9136
MLR	Training	580	0.8814	0.8736	0.9388	0.1000	0.2000	30.3476	34.4352
	Test	248	-	-	0.8635	0.2000	0.3000	40.6835	51.7548

SLR: simple linear regression; MLR: multiple linear regression; R^2: coefficient of determination; R^2adj: adjusted coefficient of determination; r: Pearson coefficient of correlation; MAE: mean absolute error; RMSE: root mean squared error; RAE: relative absolute error; RRSE: root relative squared error.

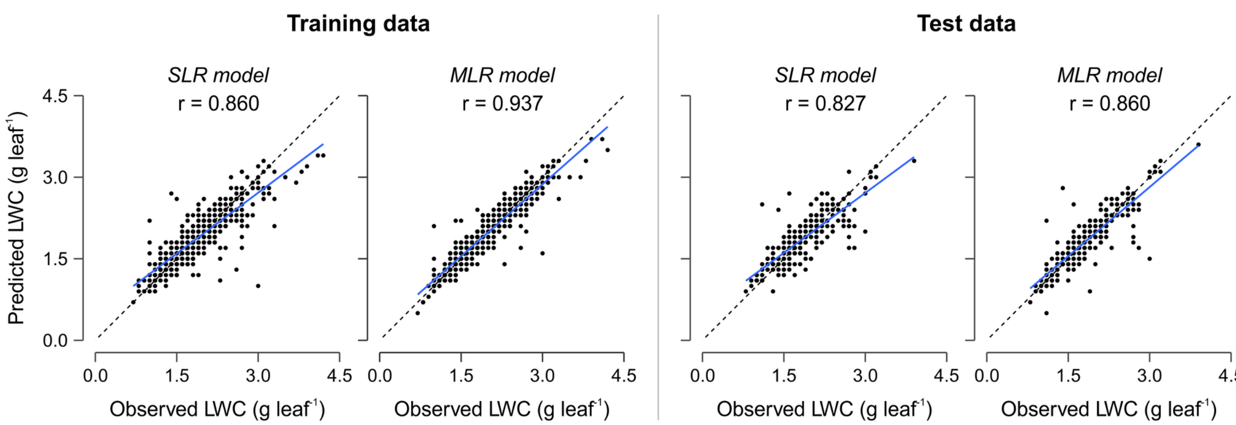

Fig 1. Observed x predicted grapevine leaf water content (LWC) for the simple linear regression (SLR) and multiple linear regression (MLR) models using training and test datasets.

Conclusion

Grapevine leaf area estimated via digital image processing provide a reliable measure of leaf water content and its variability. Incorporating reflectance at specific wavelengths (≈1500–1900 nm), which are sensitive to leaf water content and structure, reduce prediction errors and improved the estimation of grapevine leaf water content. These results support the approach´s potential application for guiding water management in precision viticulture.

Acknowledgments

The authors thank the Brazilian Agricultural Research Corporation (Embrapa Instrumentation) for laboratory infrastructure and the Guaspari Winery, Espírito Santo do Pinhal, state of São Paulo, Brazil, for providing access to the study sites and leaf sampling campaign.

References

González-Fernández A. B., Rodríguez-Pérez J. R., Marcelo V., & Valenciano J. B. (2015).

Using field spectrometry and a plant probe accessory to determine leaf water content in commercial vineyards. Agricultural Water Management, 156, 43–50, doi: https://doi.org/10.1016/j.agwat.2015.03.024.

Bahat I., Netzer, Y., Grünzweig J. M., Naor A., & Alchanatis V. (2024). How do spatial scale and seasonal factors affect thermal-based water status estimation and precision irrigation decisions in vineyards? Precision Agriculture, 25, 1477–1501. https://doi.org/10.1007/s11119-024-10120-5.

Rincón Guerrero N., Olarte Quintero M. A., & Pérez Naranjo J. C. (2012). Determinación del área foliar en fotografías tomadas con una cámara web, un teléfono celular o una cámara semiprofesional. Revista Facultad Nacional de Agronomía Medellín, 65(1), 6399–6405.

Müller-Linow M., Wilhelm J., & Briese C. et al. (2019). Plant Screen Mobile: an open-source mobile device app for plant trait analysis. Plant Methods, 15, 2, doi: https://doi.org/10.1186/s13007-019-0386-z.

Rapaport T., Hochberg U., Shoshany M., Karnieli A., & Rachmilevitch, S. (2015). Combining leaf physiology, hyperspectral imaging and partial least squares-regression (PLS-R) for grapevine water status assessment. ISPRS Journal of Photogrammetry and Remote Sensing, 109, 88–97. https://doi.org/10.1016/j.isprsjprs.2015.09.003.

Tardáguila, J., Fernández-Novales, J., Gutiérrez, S., & Diago, M. P. (2017). Non-destructive assessment of grapevine water status in the field using a portable NIR spectrophotometer. Journal of the Science of Food and Agriculture, 97(11), 3772–3780, doi: https://doi.org/10.1002/jsfa.8241.

Laroche-Pinel E., Vasquez K., & Brillante L. (2024). Assessing grapevine water status in a variably irrigated vineyard with NIR/SWIR hyperspectral imaging from UAV. Precision Agriculture, 25, 2356–2374, doi: https://doi.org/10.1007/s11119-024-10170-9.

Giovenzana V., Beghi R., Parisi S., Brancadoro L., & Guidetti R. (2018). Potential effectiveness of visible and near infrared spectroscopy coupled with wavelength selection for real time grapevine leaf water status measurement. Journal of the Science of Food and Agriculture, 98(5), 1935–1943, doi: https://doi.org/10.1002/jsfa.8676.

Katabuchi M. (2015) LeafArea: an R package for rapid digital image analysis of leaf area. Ecological Research, 30, 1073–1077, doi: https://doi.org/10.1007/s11284-015-1307-x.