Spectral-Based Discrimination of *Vitis vinifera* and *Vitis labrusca* Using Contact SpectroradiometryTechniques

Tainá Fragoso[1], José Eduardo Costa[1,2], Jorge Ducati[1]

[1] State Center for Research in Meteorology and Remote Sensing - Federal University of Rio Grande do Sul, Brazil (CEPSRM/UFRGS)
[2] Physics Institute - Federal University of Rio Grande do Sul, Brazil (IF/UFRGS)

Keywords: Contact Spectroradiometry, Spectral Classification, Grapevine Varieties Discrimination.

1. Introduction

Brazil holds a rising position in the global viticulture landscape, both in terms of production and varietal diversity, hosting *Vitis vinifera* and *Vitis labrusca* vineyards within the same geographical regions. This configuration is partly the result of historical processes of territorial occupation, marked by successive waves of migration and the local adaptation of grape varieties to diverse edaphoclimatic conditions (Protas et al., 2024). This varietal richness is an important and unique characteristic of Brazilian grape production, but it also poses challenges, particularly for vineyard monitoring, traceability, origin certification, and the standardization of wine quality.

Viticulture in Brazil is undergoing a gradual transition from a labrusca-based production to a vinifera-based one. Even though *Vitis labrusca* vines are still the majority of vineyards, either for *in-natura* grape consumption or for table wines, efforts over the last decades have aimed to develop a production of fine wines based on an increasing area of *Vitis vinifera* grapes (Protas et al., 2002). This effort has recently led to initiatives focused on wine quality specifications where, based on geographical, environmental, and cultural descriptors, certain precisely defined regions are being granted origin certifications, such as Denomination of Origin (*Denominação de Origem*) or Indication of Provenance (*Indicação de Procedência*) (Camargo et al., 2011).

While in some Brazilian regions with no previous labrusca vineyards, mainly in Central Brazil or in the Northeast, these origin certifications are applied directly to new *Vitis vinifera* projects, in southern Brazil the transition occurs in regions with a well-established tradition of *Vitis labrusca* production by European settlers since the late 19th century. Therefore, new *Vitis vinifera* vines coexist in the same geographical space with *Vitis labrusca* (Protas et al., 2002).

As legal regulations for quality and origin certification include requirements for production volume and yield per area, knowledge of the area of grapes and their location is necessary. This information is, in principle, available from the *Cadastro Vitícola* (Embrapa Uva e Vinho, 2025), an inventory maintained by Embrapa (the Brazilian Agricultural Research Corporation), whose main source comes from the producers themselves. However, an independent way of knowing where and how much *Vitis vinifera* grapes are planted would be desirable, and remote and proximal sensing provide tools potentially useful for these purposes.

The identification and mapping of the spatial distribution of grapevine varieties are critical steps for production planning, implementation of public policies, and consolidation of territorial identities. In this context, remote and proximal sensing technologies offer promising, non-invasive, and systematic alternatives for crop detection and discrimination (Mazzia et al., 2020). Among these, contact spectroradiometry stands out, as it captures subtle variations in leaf reflectance and reveals distinct spectral signatures among grapevine varieties (Lacar et al., 2001).

The present study investigates the use of contact spectroradiometry for the discrimination of grapevine varieties in vineyards from Rio Grande do Sul, Brazil, with a particular focus on distinguishing between *Vitis vinifera* and *Vitis labrusca* cultivars. One hybrid variety and a rootstock were included as auxiliary classes to assess model performance, rather than as central targets of the analysis. The main objective is to identify spectral regions with high discriminative power by analyzing leaf reflectance and to evaluate the effectiveness of analytical approaches such as spectral ratio, derivatives and Linear Discriminant Analysis (LDA). By characterizing spectral differences among varieties, this work aims to support the development of classification models that enhance vineyard monitoring, to assist in origin certification with application in public policies aimed at territorial management of viticultural production.

2. Material and methods

2.1 Data Collection

For the selection of the study site we looked for an already-existing situation of the simultaneous presence, within the same area (meaning less than a few hectares), of both vinifera and labrusca vines. The smallness of area would ensure that environmental differences (as soil and microclimate) would be minimized. Moreover, we looked for an as large as possible diversity of grape varieties, to have a better statistical sample representing both groups. We were able to locate a producer that satisfies these criteria, displaying an expressive diversity of grape varieties in a relatively small space (less than three hectares).

Spectral data were collected in December 2024 from vineyards located in the municipality of Mariana Pimentel, Rio Grande do Sul, Brazil. Measurements were carried out using a handheld FieldSpec® 3 ASD spectroradiometer, operating across the 350–2500 nanometers (nm) spectral range. The instrument was equipped with a Contact Probe, which contains an integrated light source, and a Leaf Clip. System calibration was performed using a Spectralon® white reference panel.

Five plants per variety were sampled, and one leaf was collected from each plant — except for Cabernet Sauvignon, from which two leaves per plant were measured, in view of future studies also incorporating transmittance spectra — totaling 60 leaves.

Leaves were harvested from the mid-canopy region and measured immediately after collection. For each leaf, five reflectance spectra were acquired at two distinct positions on the adaxial surface (upper and lower regions), avoiding major veins. This procedure was designed to capture intra-leaf variability and ensure more representative spectral profiles. In the case of the Paulsen 1103 rootstock, only one measurement per leaf was taken due to its morphological uniformity.

In summary, 50 spectra were obtained for each variety (100 for Cabernet Sauvignon) and 25 for the rootstock, totaling 575 spectra. One spectrum presented a measurement error and was identified as an outlier; therefore, it was excluded from the dataset, resulting in 574 adaxial surface reflectance spectra covering the following varieties: *Barbera*, *Cabernet Sauvignon*, *White Garganega*, *Merlot*, *Sangiovese*, *Syrah*, *Isabella*, *White Niagara*, *Rose Niagara*, *Goethe*, and Paulsen 1103 (Table 1).

Identifier	Variety
ba	Barbera
cs	Cabernet Sauvignon
ga	White Garganega
go	Goethe
isa	Isabella
me	Merlot
nb	White Niagara
nr	Rose Niagara
sa	Sangiovese (Brunello clone)
sy	Syrah
enx_pe	Paulsen 1103 rootstock

Table 1. Analyzed grapevine varieties.

2.2 Processing and Analysis

The use of hyperspectral data presents great potential for biophysical characterization and for detecting subtle variations in the spectral signatures of leaves. However, this requires a deeper understanding of the quality of the spectral responses associated with different types of vegetation and environmental conditions (Lacar et al., 2001; Pithan et al., 2021). As highlighted by Schroeder (2000), the performance of a detection system is often evaluated using the signal-to-noise ratio (SNR), with higher SNR values allowing more accurate identification of subtle differences between leaf spectra.

The signal-to-noise ratio (SNR) varies as a function of wavelength, and is defined as the ratio between the mean reflectance and the associated standard deviation. The SNR was estimated at wavelengths close to the values recommended by the spectroradiometer manufacturer: 620 nm, 1150 nm, and 2200 nm, based on reflectance measurements obtained from the Spectralon® white reference panel.

The data analysis methodology applied in this study is based on previous works by Pithan et al. (2021), Tsai and Philpot (1998), and Blackburn (2006), and was developed with the aim of identifying wavelengths with high discriminative power for grapevine variety differentiation. The workflow involves the calculation of mean reflectance spectra, spectral ratios, first derivatives, standard deviations, and LDA combined with a Forward Stepwise variable selection procedure.

The preprocessing steps included detector-level leveling correction, spectral smoothing, and normalization by the area under the spectral curve. The ultraviolet range (350–399 nm) was excluded due to its high noise level.

After preprocessing, mean reflectance spectra were calculated for each variety. Figure 1 shows the mean spectral reflectance curves of the grapevine varieties analyzed in this study.

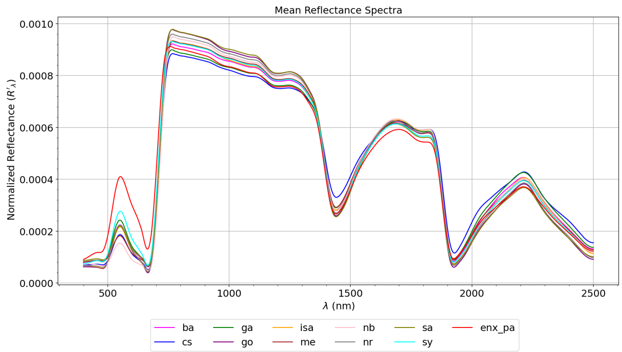

Figure 1. Mean spectral reflectance curves of the samples.

Next, the ratio point-to-point between each different pair of mean spectra were generated. Ratios between spectra, as used by Pithan et al. (2021), can reveal subtle spectral differences. The first derivatives of these ratio curves are shown in Figure 2, highlighting regions where slope changes are most pronounced, which suggest indicates discriminative spectral features.

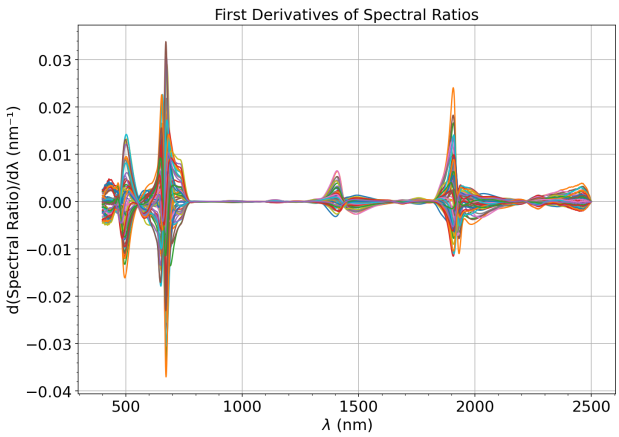

Figure 2. First derivative curves of the spectral ratios, highlighting the regions with the most significant differences between variety combinations.

The variability of the derivatives was expressed by their standard deviation, allowing the identification of spectral regions with the highest inter-varietal variability. Based on this analysis, a preliminary selection of wavelengths was carried out by applying a minimum distance constraint between selected bands to avoid redundancy.

To select the final set of wavelengths, the Forward Stepwise algorithm was combined with Linear Discriminant Analysis (LDA). At each step, LDA evaluates whether a candidate band improves class separation by maximizing the ratio between inter-class and intra-class variance. Wavelengths that significantly increase separability are retained, while redundant or non-informative ones are excluded. Linear Discriminant Analysis (LDA) leverages the selected wavelengths by combining them into linear discriminant functions that maximize between-group variance while minimizing within-group variance. In practice, each of the final selected wavelengths contributes with a weighted coefficient, reflecting its discriminative power, to build a reduced set of canonical axes in which varietal separation is optimized. The accuracy and robustness of the resulting LDA models were then evaluated through 5-fold cross-validation.

3. Results

The results demonstrate that spectral analysis using contact spectroradiometry enables efficient discrimination of the grapevine varieties analyzed. The assessment of data quality through the SNR yielded values of 445, 2642, and 544 for 620 nm, 1150 nm, and 2200 nm, respectively, high enough to ensure the reliability of reflectance measurements.

The spectral ratio method, combined with the calculation of first derivatives and their standard deviations, allowed the identification of spectral regions with the highest variability among varieties. These regions are mainly located in the visible spectrum—particularly in the green and red-edge regions—and in the shortwave infrared (SWIR). The spectral ratio process generated 55 pairwise comparisons between varieties, highlighting wavelengths where the ratios deviated most from 1, indicating greater discriminative potential.

A total of 21 wavelengths were preselected, distributed across the VNIR (400–1050 nm), SWIR1 (1051–1800 nm), and SWIR2 (1801–2500 nm) regions, maintaining a minimum spacing of 25 nm between bands to avoid redundancy. These bandwidths are in accordance with the optimum spectral bandwidths for agricultural applications, as indicated by Ray et al. (2010). The preselection was based on derivative variability.

The application of Forward Stepwise selection combined with LDA and 5-fold cross-validation resulted in the identification of the ten most discriminative wavelengths: 428, 497, 649, 674, 699, 724, 1408, 1515, 1540, and 1858 nm. Model performance analysis showed that the progressive addition of bands improves accuracy up to a saturation point—around 10 to 12 bands—beyond which the inclusion of additional bands does not yield significant improvements and may increase the risk of overfitting. Figure 3 shows how the average accuracy varies with the increase in the number of wavelengths used in the LDA model.

The selected wavelengths fall within specific spectral regions that represent known leaf characteristics. The bands at 428, 497, 649, 674, and 699 nm are located in the visible region, which is dominated by pigment absorption. According to Hennessy et al. (2020), chlorophylls a and b exert the strongest influence on absorption in this region, followed by carotenoids and anthocyanins. The band at 724 nm, in the red-edge region of the near-infrared, is associated with internal leaf structure, chlorophyll concentration, and water content. The bands at 1408, 1515, 1540, and 1858 nm lie within the shortwave infrared, primarily related to water absorption, but also associated with lignin, cellulose, and tannin content. The set of selected wavelengths is consistent with feature selection reported in previous studies, both at the leaf and canopy scales.

The final LDA model achieved an average classification accuracy of approximately 93% (70% training / 30% testing). The confusion matrix shown in Figure 4 illustrates the model's performance in discriminating between varieties. Each row corresponds to the true variety of the sample, while each column

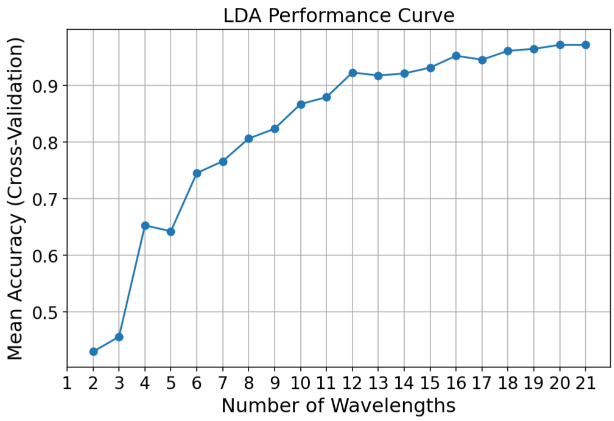

Figure 3. Performance curve showing how the average accuracy (obtained through cross-validation) varies with the increase in the number of wavelengths used in the LDA model.

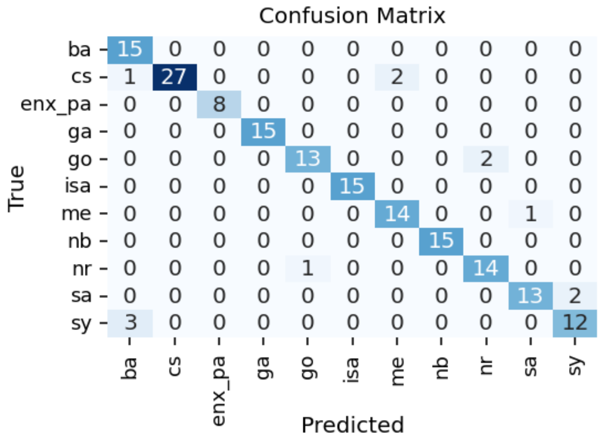

Figure 4. Confusion matrix of the LDA model based on selected wavelengths, illustrating the classification performance among the analyzed groups.

represents the predicted variety by the model. The values indicate how many samples were classified into each combination of true versus predicted classes.

The projection of the samples into the 2D discriminant space, shown in Figure 5, revealed three well-defined clusters: (A) *Vitis vinifera* cultivars, (B) hybrids and American species (*Vitis labrusca*), and (C) the Paulsen 1103 rootstock, confirming the effectiveness of the applied methods.

In the projection of Figure 5, it can be observed that the varieties classified as labruscas share the space with the hybrid Goethe, a pattern particularly evident for the Isabella variety. Although generally classified as a labrusca, Isabella is described in different sources as a hybrid between V. labrusca and Meslier Petit (VIVC, 2025). Furthermore, the White Niagara and Rose Niagara varieties present a genetic proportion of V. vinifera through their ancestor Concord, as demonstrated in studies such as (Huber et al., 2016). These findings indicate that future investigations into the origin of these varieties may be relevant, especially since this factor appears to be captured by the discrimination models.

Tests applying different minimum-spacing criteria for band se-

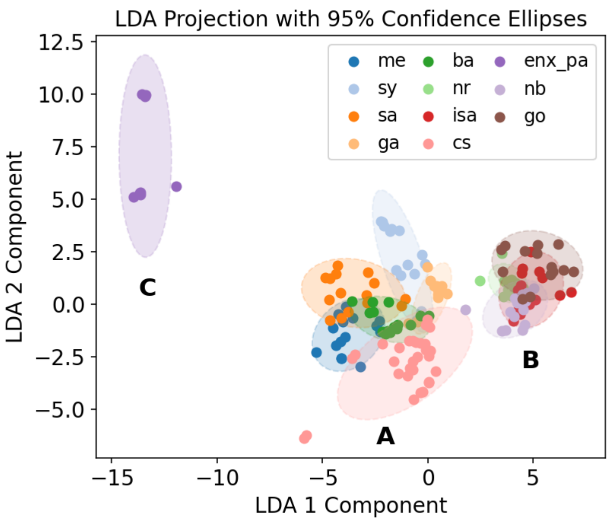

Figure 5. Projection of the samples onto the LDA model space using the selected wavelengths, highlighting three distinct groups: (A) *Vitis vinifera* cultivars, (B) hybrids and American species (*Vitis labrusca*), and (C) the Paulsen 1103 rootstock.

lection resulted in alternative wavelength sets, yet all were concentrated in the same spectral regions. This consistency indicates that varietal discrimination is predominantly driven by specific spectral features, which reflect the biochemical and morphological characteristics of the leaves. These features are associated with pigment absorption bands, water content, and regions related to the internal leaf structure. Therefore, the method maintains its efficiency even with slight variations in the selected wavelengths.

The results confirm the feasibility of using spectral analysis as an auxiliary tool for varietal mapping, traceability, and origin certification in viticulture. Nevertheless, some limitations of the study may introduce biases and should be addressed in future work with broader datasets. These include the geographical scope, since the data were collected in a single municipality with specific soil and terrain characteristics; the sample scope, as only one field campaign was conducted (December 2024), without accounting for seasonal variability; and methodological aspects, as the measurements were based on foliar contact spectra, which may require adaptations for application with proximal or remote sensors.

4. Conclusion

The results suggest that certain wavelengths play an instrumental role in the spectral separation between the two studied groups. As several regions in southern Brazil are undergoing processes of wine origin certification, this knowledge can support the fine-tuning of aerial observing and monitoring devices, which are increasingly available in both local and international markets.

Furthermore, the findings indicate that contact spectroradiometry, combined with spectral analysis methods and multivariate statistics, represents a promising approach for the discrimination of grapevine varieties. The most relevant spectral regions for this differentiation are concentrated in the visible and near-infrared (VNIR) and shortwave infrared (SWIR) ranges, reflecting the biochemical composition and structural characteristics of the leaves.

Overall, the results demonstrate that systematic variations exist in the spectral signatures of the analyzed varieties, and that these variations can be effectively exploited through spectral analysis combined with statistical methods, enabling the identification of spectral regions with the highest discriminative power between varietal groups.

5. Acknowledgements

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

References

Blackburn, G. A., 2006. Hyperspectral remote sensing of plant pigments. *Journal of Experimental Botany*, 58(4), 855-867. https://doi.org/10.1093/jxb/erl123.

Camargo, U. A., Tonietto, J., Hoffmann, A., 2011. Progressos na viticultura brasileira. *Revista Brasileira de Fruticultura*, 33(spe1). https://www.oalib.com/research/1017199?utm$_{s}ource = chatgpt.com.Acessoem: 16set.$2025.

Embrapa Uva e Vinho, 2025. Cadastro vitícola nacional. Acesso em: 16 set. 2025.

Hennessy, A., Clarke, K., Lewis, M., 2020. Hyperspectral Classification of Plants: A Review of Waveband Selection Generalisability. *Remote Sensing*, 12(1). https://www.mdpi.com/2072-4292/12/1/113.

Huber, F., Röckel, F., Schwander, F., Maul, E., Eibach, R., Cousins, P., Töpfer, R., 2016. A view into American grapevine history: Vitis vinifera cv. 'Sémillon' is an ancestor of 'Catawba' and 'Concord'. *VITIS - Journal of Grapevine Research*, 55(2), 53–56. https://ojs.openagrar.de/index.php/VITIS/article/view/6422.

Lacar, F., Lewis, M., Grierson, I., 2001. Use of hyperspectral reflectance for discrimination between grape varieties. *IGARSS 2001. Scanning the Present and Resolving the Future. Proceedings. IEEE 2001 International Geoscience and Remote Sensing Symposium (Cat. No.01CH37217)*, 6, 2878–2880 vol.6.

Mazzia, V., Comba, L., Khaliq, A., Chiaberge, M., Gay, P., 2020. UAV and Machine Learning Based Refinement of a Satellite-Driven Vegetation Index for Precision Agriculture. *Sensors*, 20(9). https://www.mdpi.com/1424-8220/20/9/2530.

Pithan, P. A., Ducati, J. R., Garrido, L. R., Arruda, D. C., Thum, A. B., , Hoff, R., 2021. Spectral characterization of fungal diseases downy mildew, powdery mildew, black-foot and Petri disease on Vitis vinifera leaves. *International Journal of Remote Sensing*, 42(15), 5680–5697.

Protas, J. F. d. S., Camargo, U. A., Mello, L. M. R. d., 2002. A viticultura brasileira: realidade e perspectivas. *Simpósio Mineiro de Viticultura e Enologia, 1., 2002, Andradas, MG. Anais... Viticultura e Enologia: atualizando conceitos*, Caldas: EPAMIG, 17–32. Acesso em: 16 set. 2025.

Protas, J. F. d. S., Lazzarotto, J. J., Machado, C. A. E., 2024. Panorama da vitivinicultura brasileira em 2022. Comunicado Técnico 233, Embrapa Uva e Vinho, Bento Gonçalves, RS, Brasil.

Ray, S., Jain, N., Miglani, A., Singh, J., Singh, A. K., Panigrahy, S., Parihar, J., 2010. Defining optimum spectral narrow bands and bandwidths for agricultural applications. *Current science*, 93, 1365-1369.

Schroeder, D. J., 2000. *Astronomical Optics*. 2nd edn, Academic Press, San Diego, CA. Department of Physics and Astronomy, Beloit College.

Tsai, F., Philpot, W., 1998. Derivative Analysis of Hyperspectral Data. *Remote Sensing of Environment*, 66(1), 41-51.

VIVC, 2025. Isabel — vitis international variety catalogue. https://www.vivc.de/index.php?r=passport/viewid=5560. Accessed: 2025-09-21.