



HAL
open science

Large-scale monitoring of California vineyards with Sentinel-2 time-series

Baptiste Oger, Mahyar Aboutalebi, Brent Sams, Nick Dokoozlian, Maria Mar Alsina

► To cite this version:

Baptiste Oger, Mahyar Aboutalebi, Brent Sams, Nick Dokoozlian, Maria Mar Alsina. Large-scale monitoring of California vineyards with Sentinel-2 time-series. European Conference on Precision Agriculture (ECPA 2025), Jun 2025, Barcelona, Spain. pp.775-782, <10.1163/9789004725232_102>. <hal-05283492>

HAL Id: hal-05283492

<https://hal.science/hal-05283492v1>

Submitted on 10 Oct 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC-ND 4.0 - Attribution - Non-commercial use - No Derivative Works - International License

101. Large-scale monitoring of California vineyards with Sentinel-2 time-series

Baptiste Oger¹, Mahyar Aboutalebi¹, Brent Sams¹, Nick Dokoozlian¹ and Maria Mar Alsina^{1,2,*}

¹E&J Gallo Winery, Winegrowing research, Modesto, CA, USA

²IRTA Mas Badia, 17134 La Tallada d'Empordà, Girona, Spain; * mariadelmar.alsina@irta.cat

Abstract

Remote sensing has become a valuable tool for enhancing agricultural efficiency. However, its application in vineyards is challenging due to the high variability of viticultural systems coexisting. Using Sentinel-2 time series data from 1,014 vineyards across California, this study analysed the influence of the region and vineyard characteristics on seasonal growth patterns and introduced a clustering approach to categorize vineyards into six typologies of expected growth behaviour. The results show that region, followed by vineyard age, row spacing, and trellis type, are key factors influencing clustering. The proposed clustering method showcases the potential of remote sensing to support large-scale, data-driven decision-making in viticulture.

Keywords: clustering, NDII, NDVI, remote sensing, trellis

Introduction

Over the past two decades, remote sensing has proven to be a valuable technique to monitor vegetation and its responses to biotic and abiotic stresses, and continues to be regarded as a powerful tool to help improve agriculture efficiency (Karthikeyan *et al.*, 2020). Vineyards present specific challenges to remote sensing techniques because of their architecture, where rows of clumped vegetation supported by a trellis system alternate with inter-row spaces consisting of bare soil or managed vegetation cover (Giovos *et al.*, 2021). The type of training system used, along with some other factors such as the big range of ages of vineyards in production, results in numerous “vineyard types” with unique characteristics.

Today, the ESA Sentinel-2 satellite platform provides multispectral images with a spatial resolution of 10–20 m per pixel and a revisit time of 4–5 days, enabling detailed time-series analyses. This temporal richness allows for monitoring throughout the growing season rather than relying on single snapshots. For instance, complete NDVI time series can reveal insights into spatial variability within vineyard fields, reflecting differences in management practices or vegetative development (Devaux *et al.*, 2019). Despite this potential, linking these spectral patterns to relevant agronomic parameters remains complex, underscoring the need for refined approaches to interpret and leverage such data effectively. As a result, practical applications of remote sensing for large-scale vineyard monitoring remain underdeveloped.

This study addresses these challenges using Sentinel-2 NDVI and NDII time-series data from 1014 vineyards across California. It proposes a straightforward method to harness the information contained in large remote sensing datasets for vineyard management at scale. First, it presents a simple representation of the time series, illustrating its complexity and how site characteristics influence seasonal patterns and maximum values. Then, it introduces a clustering approach to define typologies that represent expected vine growth behaviours for different vineyard types. These typologies provide a reference framework, allowing real-time Sentinel-2 data to be compared with known patterns to facilitate early stress detection and timely interventions.

Material and methods

Study vineyards

This study collected and analysed data from 1014 vineyard fields across California, covering 11 308 ha. Each “field” is a continuous surface planted to vines of the same variety and on the same trellis system. Information on grape variety, trellising, vine age, and vine and row spacing was available for all the fields, and, in some cases, weed and cover crop management were also known. The study focused on 50 grape varieties, with ten primary ones (Cabernet Sauvignon, Pinot Noir, Merlot, Syrah, Zinfandel, Petite Sirah, Malbec, Pinot Gris, Chardonnay and Sauvignon Blanc) accounting for 84% of the vineyard fields.

The vineyards analysed in this study were grouped into five classes, presented in Figure 1, based on the existing training systems:

1. **Quadrilateral cordon:** Horizontally split canopies where two bilateral cordons run in parallel at a height of 135–150 cm and spaced 30–60 cm. The shoots are left to freely sprawl from the cordons.
2. **Bilateral cordon:** Canopies trained on a single bilateral cordon or vertically divided (two bilateral cordons running in parallel spaced 45 cm on average) and the shoots are left to freely sprawl. The first cordon is situated 137–168 cm above ground, and, when the second cordon exists, is situated between 137 and 168 cm.
3. **Head trained:** Head-pruned vines organized in rows with shoots held by one catch wire forming a continuous stripe of canopy. Trunk height ranges from 76 to 114 cm, with the wire positioned at 127–165 cm and spaced 51–76 cm.
4. **Quadrilateral cordon (SP):** Horizontally split canopy where two bilateral cordons run in parallel 30–60 cm apart, at a height of 135–150 cm and the shoots are held by one catch wire at a height of 117–142 cm. The single catch wire is held by a metal structure consisting of a metal crossbar (T-Stack structure), two metal bars in a V shape, or a single bar shaped as a lyre (Elk-Horn) mounted on the trellis poles.
5. **VSP (Vertical shoot positioning):** Canopies trained into a single cordon at a height ranging from 53 to 122 cm, with the shoots held vertically by two or more levels of catch wires, separated by less than 64 cm. The fully developed canopy looks like a “wall” of vegetation whose maximum height ranges from 142 to 218 cm.

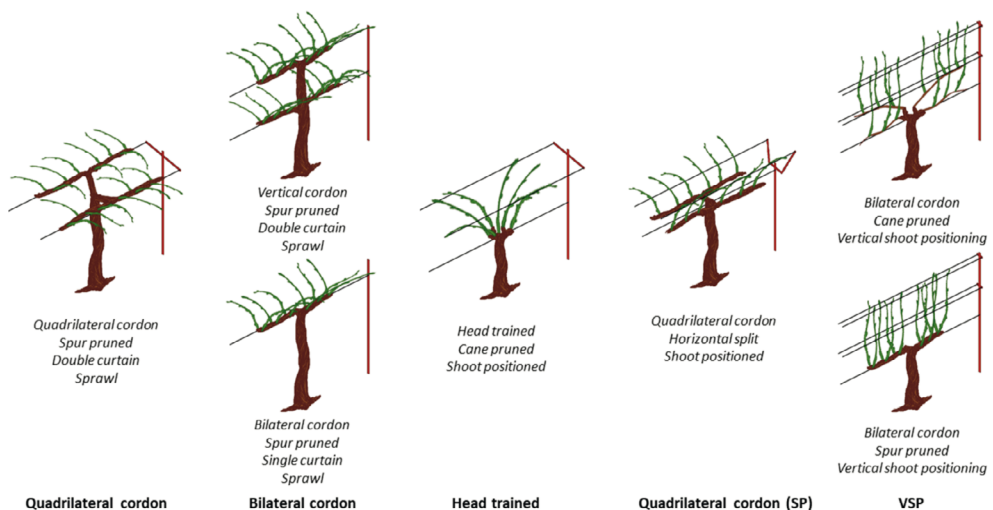


Figure 1. Typologies of the trellis system.

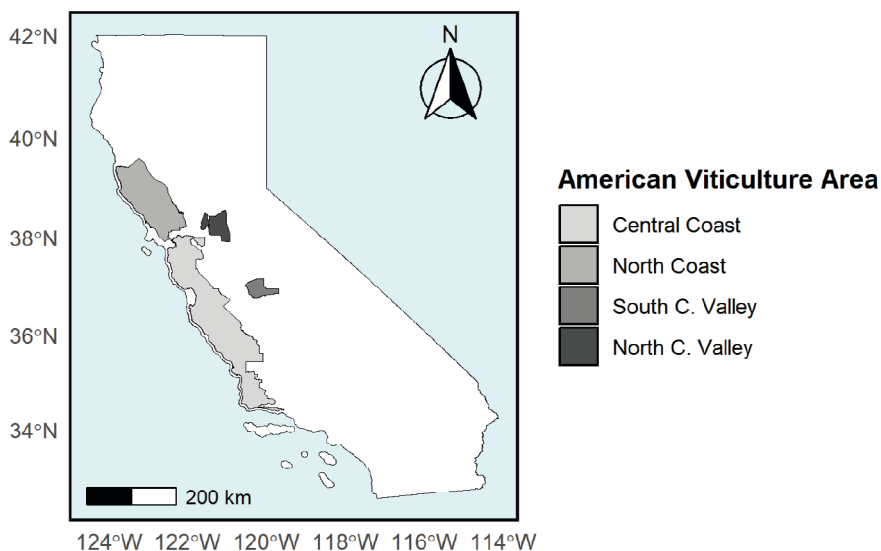


Figure 2. Map of the California American Viticulture Areas considered in this study.

Vineyards were grouped into four regions depending on their location: North Coast, Central Coast, Northern Central Valley, and Southern Central valley. The North Coast and Central Coast regions are based on American Viticultural Areas (AVAs). The Northern Central Valley corresponds to the Delta basin, mostly encompassing the Lodi and Clarksburg AVAs, while the Southern Central Valley corresponds to the San Joaquin and Tulare basins, around the Madera AVA. The geographic distribution of these regions is shown in Figure 2, while Table 1 summarizes the available fields within each region.

All vineyards in this study were irrigated, with most featuring inter-row cover crops that dried out by summer. Row distance ranged from 1.5 m to 3.6 m. Clear differences in production goals emerged between California’s central valley and coastal regions. The Central Valley prioritized maximizing yield, supported by a warm climate, heavy irrigation, and bigger canopies. Coastal regions, focusing on fruit quality, commonly employed Vertical shoot positioning (VSP) or other Quadrilateral cordon where shoots are positioned. Coastal areas experienced a Mediterranean climate with mild winters and moderated summer temperatures due to fog and ocean breezes, while the inland valleys had hotter summers and greater temperature extremes.

Table 1. Information about available data in each AVA Region.

American Viticulture Area	Main varieties	Main trellis type	Number of fields	Total area (ha)
North Coast	Cabernet sauvignon, Chardonnay, Pinot noir	VSP, Quadrilateral cordon (SP), Head trained	3259	3145
Central Coast	Pinot noir, Chardonnay, Cabernet sauvignon	VSP, Quadrilateral cordon (SP), Quadrilateral cordon	1266	1466
Northern Central Valley	Cabernet sauvignon, Pinot noir, Sauvignon Blanc	Bilateral cordon, Quadrilateral cordon, Quadrilateral cordon (SP)	696	2528
Southern Central Valley	Chardonnay, Merlot, Riesling	Bilateral cordon, Quadrilateral cordon, Quadrilateral cordon (SP)	814	4169

Remote sensing data

Sentinel-2 data were collected over a five-year period, from 2019 to 2023, using Google Earth Engine with a cloud filtering algorithm (Hollstein *et al.*, 2016). For each field and Sentinel-2 image, average NDVI and NDII were calculated from the near-infrared (NIR), red (RED), and shortwave infrared (SWIR) bands, as shown in Eqs. (1) and (2).

$$\text{NDVI} = (\text{NIR} - \text{RED}) / (\text{NIR} + \text{RED}) \quad (1)$$

$$\text{NDII} = (\text{NIR} - \text{SWIR}) / (\text{NIR} + \text{SWIR}) \quad (2)$$

Only pixels inside the field polygon and a minimum of 10 m away from the edge were considered to avoid pixels containing a mix of vineyard and non-vineyard reflectance. For the vineyards that were uprooted or planted during the five-year period, only the years when the vineyard was productive were considered. Additionally, vineyards less than three years old were excluded, except where specified otherwise for illustrative purposes. The final dataset includes a total of 4521 time-series. Additional filtering was applied to eliminate cloudy data and isolate points with unusually low or high NDVI values. Images from fields showing a change of 0.1 or more in NDVI compared to the previous and next values, as well as the linear interpolation between them, were removed.

Data processing

The classification of fields was performed using the k-means method based on the median values of NDVI and NDII observed over a six-week period, from early June to mid-July (week 23 to week 28). A type III ANOVA was undertaken to analyse the effect of each vineyard characteristic on the cluster assigned to each field using the car package (Fox *et al.*, 2012).

Results and discussion

Figure 3 highlights the diversity of NDVI time series and their dependence on multiple factors. Most trellis types, the Central Valley, and vineyards with wider row spacing exhibit significant variability in their NDVI time series. For these cases, it is challenging to explain and interpret NDVI values using just one or two parameters. The only notable exception occurs in fields with Vertical shoot positioning (VSP) trellises, where all NDVI time series values after the canopy is fully developed are consistently similar within a margin of ± 0.1 . This uniformity is attributed to the VSP training, where the canopies form a dense “wall” of vegetation along the vine row. This structure makes it difficult to detect variations using satellite imagery, which captures images from an overhead perspective.

This effect is also observed in the Central Valley and North Valley regions, as well as in vineyards with row spacing narrower than 2.4 m, since these groups are dominated by VSP canopies. While in absence of other vegetation in the scene, the early season growth most probably would still be observed using Sentinel-2 data, but most of these VSP vineyards in the North Coast have dense and tall cover crops early in the season, making it nearly impossible to monitor vine canopy development until early summer. The presence of these cover crops explains the declining NDVI observed during the first half of the year. By June and July, these cover crops dry out, minimizing their impact on the spectral response perceived by satellites.

Figure 4 presents the NDVI and NDII value ranges for each cluster generated by the k-means algorithm. As the clustering was performed using data from the peak canopy period (June to July), the clusters are most distinguishable during this timeframe. NDII values show clear separation between clusters, whereas NDVI differences, particularly between Clusters 1 and 2, are less pronounced. This reduced distinction in NDVI is primarily due to the dominance of fields with VSP trellises

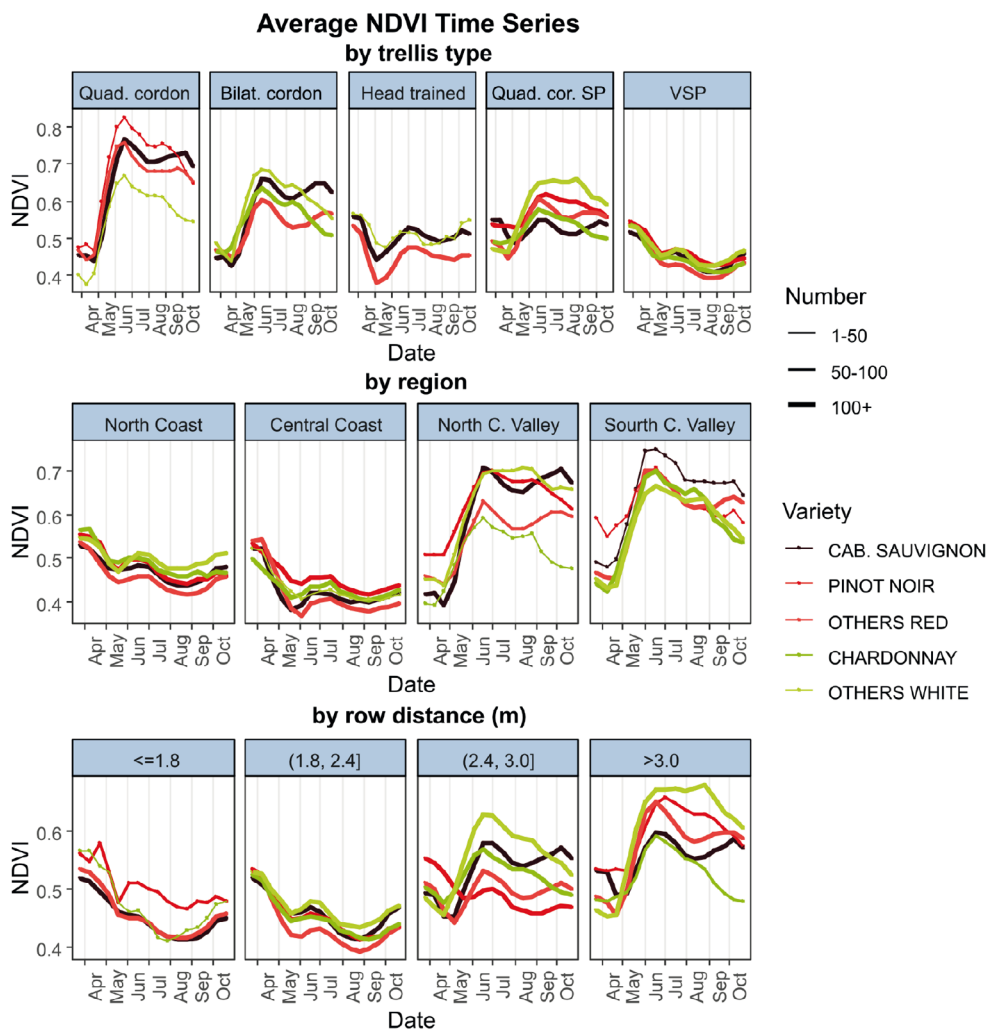


Figure 3. Average NDVI Times series based on trellis type, region, row distance and varieties. The line density represents the number of time series averaged.

in these clusters. The uniform canopy structure of VSP-managed fields makes it more challenging to differentiate them based only on NDVI values derived from satellite imagery. However, adding the SWIR bands to inform about vegetation water content, may allow to distinguish groups with different plant water content, probably resulting from different irrigation management.

Table 2 summarizes the ANOVA results analysing the influence of various factors on the cluster assignments of NDVI time series. The factors considered include trellis type, region, grape variety, year, vineyard age, and row spacing. Due to the large dataset size, all variables significantly impact the clustering results, with p -values below $2e-16$, except for the year, which has a slightly higher p -value of $7.9e-13$. This reduced effect of the year is attributed to the standardization applied during pre-processing. Without standardization, the year variable also exhibits a p -value below $2e-16$ (result not shown).

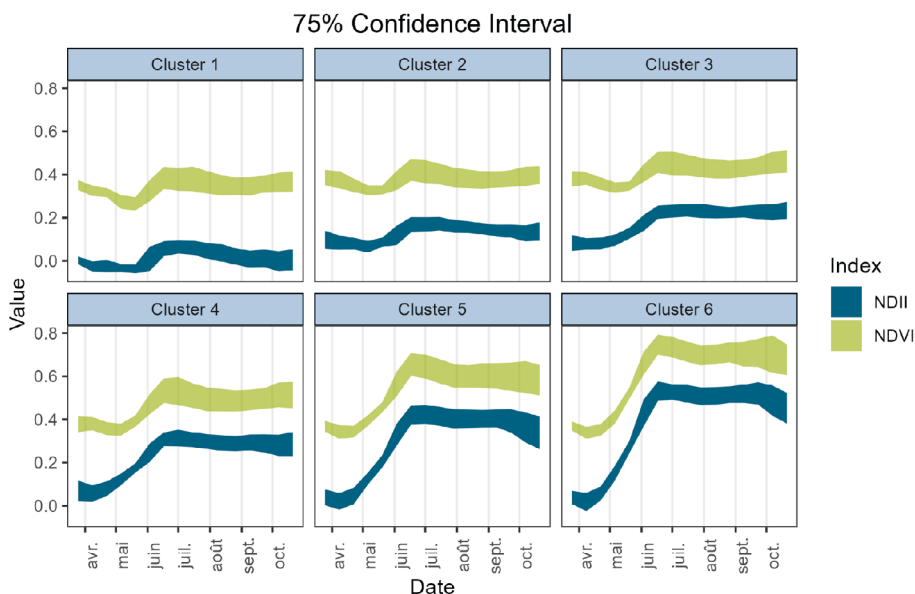


Figure 4. Results of k-means clustering. For each cluster, ribbons represent the range encompassing 75% of NDVI (green) and NDII (blue) values.

The ANOVA highlights the primary factors affecting clustering, with the region having the strongest influence, followed by vineyard age, row spacing, trellis type, and, to a lesser extent, grape variety and year. The importance of region as a factor can be readily explained by differences in climate and production goals between regions. Coastal areas are more focused on producing premium vintages, while the Central Valley regions of California aim for high yields. For vineyard age, the effect may stem from the accumulation of dead or missing vines over time. Regarding row spacing, larger distances logically increase the proportion of bare soil within mixed soil/vegetation pixels, impacting the spectral indices. Trellis type, which strongly influences the spatial arrangement of vine canopies, also plays a significant role in clustering. However, its effect might be underestimated due to its correlation with both region and row spacing. For instance, coastal regions, which emphasize fruit quality, commonly employ trellis systems with positioned shoots, reflecting their reliance on climate and vintage variability.

Table 2. ANOVA results of factors explaining the cluster assignments.

Factor	Sum Sq	df	F-value	p-value
(Intercept)	312	1	449	<2.2e-16
Trellis type	94	4	34	<2.2e-16
Region	721	3	346	<2.2e-16
Variety	106	11	14	<2.2e-16
Year	44	4	16	7.9e-13
Age	165	4	60	<2.2e-16
Row space	123	3	59	<2.2e-16
Residuals	2684			

Additional factors not represented here – such as the presence of weeds, tilled soil, cover crops, or soil type – should be investigated to provide a more comprehensive understanding. However, accessing exhaustive large-scale datasets that include such variables remains a significant challenge.

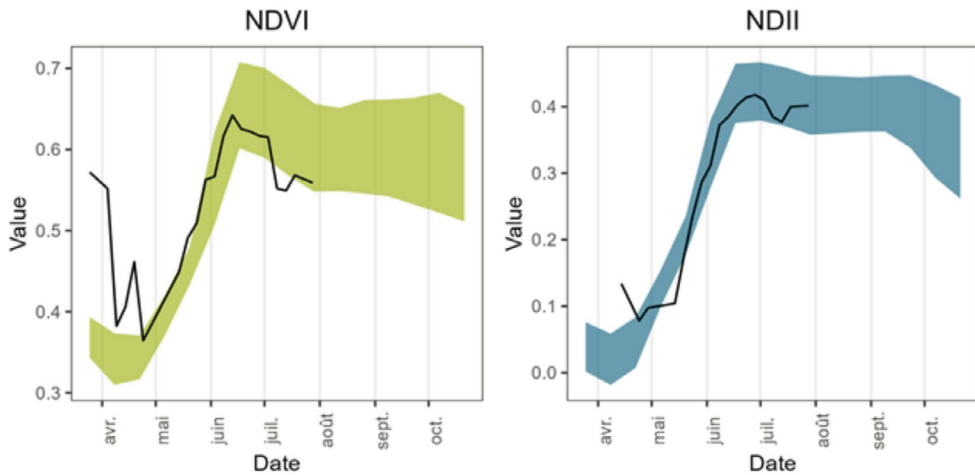


Figure 5. Real-time vegetation index time series for a Petite Sirah field in Northern Central Valley, managed with a Bilateral cordon. The ribbon represents the range encompassing 75% of the time series values within the same cluster (Cluster 6).

Figure 5 illustrates a potential use of clustering for vineyard management. With this type of interface, real-time NDII and NDVI values can be compared to existing data, such as those from previous seasons, collected from similar fields within the same cluster. Deviations or unusually high values in one or more indices may signal stress or abnormal events in the vineyard.

This system, combined with an alert mechanism for growers, could enable large-scale vineyard monitoring. Additionally, the clustering method, when paired with the appropriate indices and data, could support the prescription of management strategies based solely on remote sensing data, eliminating the need to collect vineyard-specific characteristics, which are often difficult to obtain. In this example, NDVI represents photosynthetic activity in the field, while NDII gives information about plant water content. The choice of indices is critical, as it could help guide recommendations provided to viticulturists, enhancing decision-making and precision in vineyard management.

Conclusion

This study underscores both the challenges and opportunities in using remote sensing to monitor vine stress. It highlights how Sentinel-2 time-series of NDVI and NDII are shaped by a variety of factors, often exhibiting significant variability. The clustering approach demonstrated here showcases the potential of remote sensing technologies, emphasizing how historical data can serve as a robust foundation for large-scale vineyard management with minimal reliance on ground data.

Methods such as those presented in this study could enable the effective use of remote sensing data, even in cases where not all factors influencing the spectral response of vines are fully understood. Ultimately, this work provides a foundation for developing scalable, data-driven tools to support viticulturists in balancing productivity and quality in diverse environmental and management contexts.

References

- Devaux, N., Crestey, T., Leroux, C., & Tisseyre, B. (2019). Potential of Sentinel-2 satellite images to monitor vine fields grown at a territorial scale. *OENO ONE*, 53(1), 51–58. <https://doi.org/10.20870/oeno-one.2019.53.1.2293>

- Fox, J., Weisberg, S., Adler, D., Bates, D., Baud-Bovy, G., Ellison, S., & Graves, S. (2012). Package 'car'. R Foundation for Statistical Computing, Vienna.
- Giovas, R., Tassopoulos, D., Kalivas, D., Lougkos, N., & Priovolou, A. (2021). Remote sensing vegetation indices in viticulture: a critical review. *Agriculture*, 11(5), 457. <https://doi.org/10.3390/agriculture11050457>
- Hollstein, A., Segl, K., Guanter, L., Brell, M., & Enesco, M. (2016). Ready-to-use methods for the detection of clouds, cirrus, snow, shadow, water and clear sky pixels in Sentinel-2 MSI images. *Remote Sensing*, 8(8), 666. <https://doi.org/10.3390/rs8080666>
- Karthikeyan, L., Chawla, L., & Mishra, A.K. (2020). A review of remote sensing applications in agriculture for food security: Crop growth and yield, irrigation, and crop losses. *Journal of Hydrology*, 586, 124905. <https://doi.org/10.1016/j.jhydrol.2020.124905>