121. A geostatistical fusion approach to treat vineyard sample and remote sensing data by deconvoluting kriging

G. Buttafuoco[1,*], A. Belmonte[2], S. Leggiero[3], G. Bigot[4] and A. Castrignanò[2]
[1]*National Research Council of Italy - ISAFOM, Rende (CS) 87036, Italy; * gabriele.buttafuoco@cnr.it*
[2]*National Research Council of Italy, Institute for Electromagnetic Sensing of the Environment, (CNR-IREA), 70126 Bari, Italy*
[3]*Freelance wine researcher, Brindisi 70100, Italy*
[4]*Perleuve S.r.l., Cormons (GO) 34071, Italy*

Abstract

Remote Sensing (RS) data offer a considerable source of information for implementing site-specific management in agriculture. The main limitation to their application lies in their support that may prove too coarse for applications that instead require very fine spatial resolutions. In this work, a multi-step method of data fusion using PLANET satellite data (3 m spatial resolution) is proposed in integration with measurements made on the canopy of the rows of a vineyard of very fine thickness (less than 1 m). The vineyard has an area of 5.3 ha and has rows facing south/south-west. An Italian black grape variety, Sangiovese, is cultivated in it to produce 'Chianti Classico'. The approach consists in downscaling the different radiometric variables to 1 m-scale by deconvoluting kriging and then in estimating the linear coregionalization model, considering only the directional variograms in the row direction of all (both sample and raster) data. Finally, multi-collocated cokriging was used to produce thematic maps and multi-collocated factor cokriging to partition the vineyard into homogeneous zones. The approach allowed for using satellite imagery at a resolution too coarse for the particular configuration of the vineyard under study and can be applied to any type of row crop.

Keywords: factor cokriging, multi-collocated cokriging, PLANET, support change

Introduction

In the current context, similarly to other crops, viticulture is asked to be both economically viable and environmentally sustainable. However, proper monitoring of plant condition is a prerequisite for cost-effective vine management. One of the key technologies in precision viticulture is the use of satellite remote sensing (RS) for monitoring or predicting of grape yield and quality. However, a critical issue in satellite monitoring is the moderate spatial resolution and general inability to create pure vine pixels. A pixel size greater than 1 m would result in mixed-feature pixels, containing both vines and the space between rows and cover crops (Matese *et al.*, 2015).
Therefore, a crucial problem in RS is to transform a coarse-resolution image into a finer-resolution one without introducing spurious information not contained in the original image. Moreover, in this process of disaggregating the pixel, called deconvolution or downscaling, it is important to take into account the spatial correlation that is always present when dealing with georeferenced data.
To solve this problem, geostatistics offers appropriate tools, such as downscaling with kriging and fusion of multi-source data with cokriging, which compared with traditional methods offers significant advantages as shown in Pardo-Igúzquiza *et al.* (2006, 2011).
The objective of this paper is to propose a geostatistical multi-step method of data fusion, using PLANET satellite data (3 m spatial resolution) in integration with measurements made on the canopy of the rows of a vineyard, for a delineation into homogeneous zones in the perspective of precision agriculture.

Materials and methods

Study area and measurements

The study site, covering 5.3 ha, was a vineyard planted with *Vitis vinifera L.* cv. Sangiovese to produce "Chianti Classico" wine (Figure 1).

The vineyard was established in 2018, with rows oriented at a 10° angle northward and managed using an espalier training system.

The soil has a loamy texture and before planting was submitted to redistribution of the surface horizon, in which topsoil from the upper slopes of the hill was relocated to lower areas. This facilitated a more uniform distribution of fertile topsoil across the field.

Canopy data

The viticultural attributes measured during the study included several key characteristics of the vine that provide important insights into the vineyard's health and grape quality. Among these, Exposed Leaf Surface Area (ELA, m^2) is the surface area of the canopy exposed to solar radiation. ELA was measured in field and allows to highlight significant differences between vineyards with the same total leaf area in the scope of assessing the canopy 'efficiency' based on type, spatial distribution, and density of leaf population (Argillier, 1989). ELA is also a critical indicator for evaluating vine vigour, which directly influences grape quality (Gil *et al.*, 2011). Its calculation followed the method proposed by Carbonneau (1995).

Figure 1. Study area and survey scheme on Google satellite map.

Other important attributes included canopy height and canopy width (cm), which provide valuable information about the overall structure of the plant. The average number of bunches per vine and the mean bunch weight calculated by sampling grape bunches directly from the vine, were also measured. They both provide insights about the vineyard's productivity. Bunches were weighed in field and the average weight was calculated from the bunches collected by each vine.
Finally, the mobile application Canopeo (Patrignani and Ochsner, 2015) was used to take images of plants and accurately identify areas covered by vegetation.
These attributes were assessed by harvesting three plants within the 1x1 m area.
The 4Grapes® application is available for any smartphone devices and has proven to be an essential tool for vineyard monitoring.
The sampling period extended over 22 days, from 4 to 26 September 2023, during the phenological stage of berry ripening and close to the harvest.

PLANET satellite data
All imagery in this study used PlanetScope's pre-processed surface reflectance (https://www.planet.com/products/satellite-monitoring/), focusing on the most recent generation "SuperDoves-PSB.SD" seven-bands: Band 2 Blue (465–515 nm), Band 3 Green I (513–549 nm), Band 4 Green (547–583 nm), Band 5 Yellow (600–620 nm), Band 6 Red (650–680 nm), Band 7 Red-Edge (697–713 nm), and Band 8 NIR (845–885 nm). The sensor collected data at 12-bit radiometric resolution with an approximate ground sampling distance of 3.7 m in September 2023 with less than 10% cloud cover, and complete coverage of study areas.

Geostatistical approach for data fusion
In image analysis using satellite images the available data are not on a point support but are convoluted by a weighting function on an extended area (pixel). Therefore, knowing the weighting function, estimate of the image at each pixel from the data known at coarser pixels, can be done with kriging. This approach has been developed in the frame of geostatistics (Chilès and Delfiner, 2012) and considers information on the spatial structure of the underlying phenomena. It essentially consists in the estimation of the unknown data by a local linear regression.
The practical implementation of deconvolution by kriging requires one to solve the cokriging system after performing a so-called structural analysis to estimate the underlying variogram γ. The procedure consists of different steps and is based on the comparison between experimental and model variogram of the raw coarser variable.
The multivariate data set of the case study included both sample variables (point vectorial variables) and gridded areal variables (raster variables). The auxiliary variables were the deconvoluted PLANET variables migrated to the locations of plant samples.
The multivariate spatial correlations were modelled using the linear model of coregionalization (LMC) (Wackernagel, 2003), which assumes all the n variables as produced by the same independent physical processes, acting on NS different spatial scales.
To fuse the sample variables with gridded variables for the final prediction of vine properties, the variant multi-collocated point cokriging was used (Rivoirard, 2001; Castrignanò *et al.*, 2009, 2012) when modelling sparse primary variables with secondary information much more finely sampled. Only the secondary data at the estimation location and at points where the primary variable was available were employed for interpolation.
To perform a delineation of the vineyard into homogeneous zones multi-collocated factor cokriging was applied (Castrignanò *et al.* 2009) that consists in a principal component analysis at each relevant spatial scale. The procedure provides synthetic indicators of the system of multivariate spatial correlations, called regionalized factors.

Results and Discussion

Figure 2a shows, as an example since the others appear very similar, the directional map along the vine rows of the PLANET red band after downscaling. It displays in considerable detail two large macro-areas characterised by a different reflectivity and thus absorption of the red band associated with the chlorophyll function. This would suggest a lower leaf abundance in the western part of the vineyard.

Table 1 reports the basic statistics of the variables included in the coregionalisation data set, after the downscaling process by kriging of the RS variables, in order to reduce all variables to the same support. Apart from band 2, which is the least represented of the PLANET auxiliary variables, the others do not show significant deviations from the normal distribution. It is a different situation for the sample variables because canopy width and estimated yield per vine show skewness >1 and canopy height −0.66, which presupposes the presence of outliers.

With regard to the correlations between the variables, it can be seen (Table 2) that all PLANET bands are strongly correlated with each other with the exception of the NIR band. This is due to the fact that the former give information on chlorophyll functionality while the latter on the structural characteristics of the plant. This is confirmed by the relatively high negative correlation of NIR band reflectivity with the variable Empty spaces in the canopy, which indicates a higher permeability of the canopy to NIR in the presence of empty spaces.

As for the variables measured on the plant, these are in general poorly correlated with each other and with the PLANET variables. Those most correlated are mean bunch weight and empty spaces in the canopy, the latter also correlated significantly with the PLANET bands.

Due to the presence of strongly skewed variables and because they are expressed in different units, it was preferred to transform them into standardised Gaussian variables using the Gaussian anamorphosis function (Wackernagel, 2003).

A directional model (LMC) along the direction of the rows (10°N) with an angular tolerance of 10° and with the slicing of 1m was adapted to the direct and cross directional experimental variograms of the gaussian transformed variables, in order to filter out any points belonging to the inter-row.

Table 1. Summary of statistics for the variables of the coregionalization data set.

Variable	Minimum	Median	Mean	Max	SD	Skewness	Kurtosis
Canopy height (cm)	79.00	143.00	138.46	171.00	19.28	−0.66	2.71
Canopy width (cm)	22.00	41.00	41.38	79.00	8.06	1.12	6.63
Average number of bunches per vine (count)	3.67	9.67	10.06	17.67	3.22	0.35	2.62
Mean bunch weight (g)	78.00	139.00	140.43	254.00	39.29	0.47	2.75
Estimated yield per vine (kg)	0.29	1.37	1.43	4.23	0.65	1.02	5.17
Exposed leaf area (m^2)	0.31	1.19	1.22	2.12	0.37	0.11	3.00
Empty spaces in the canopy	0.20	0.47	0.47	0.77	0.11	0.00	2.85
Band 2 Blue	0.039	0.051	0.051	0.078	0.006	0.91	5.15
Band 3 Green I	0.046	0.062	0.062	0.093	0.008	0.60	4.21
Band 4 Green	0.055	0.074	0.074	0.111	0.009	0.57	4.34
Band 5 Yellow	0.054	0.075	0.074	0.112	0.010	0.55	4.04
Band 6 Red	0.057	0.080	0.080	0.131	0.012	0.71	4.70
Band 7 Red-Edge	0.087	0.110	0.109	0.152	0.010	0.50	4.02
Band 8 NIR	0.191	0.211	0.210	0.239	0.010	0.19	2.46

Table 2. Correlation matrix between primary and auxiliary variables.

	Canopy height	Canopy width	Average number of bunches per vine	Mean bunch weight	Estimated yield per vine	Exposed leaf area	Empty spaces in the canopy	Band 2 Blue	Band 3 Green I	Band 4 Green	Band 5 Yellow	Band 6 Red	Band 7 Red-Edge	Band 8 NIR
Canopy height	1.00	0.41	0.38	0.30	0.42	0.60	0.24	0.06	0.09	0.08	0.08	0.08	0.08	0.13
Canopy width		1.00	0.36	0.58	0.59	0.73	0.67	-0.27	-0.26	-0.27	-0.26	-0.25	-0.26	-0.28
Average number of bunches per vine			1.00	0.19	0.77	0.40	0.32	-0.02	-0.03	-0.02	-0.02	-0.02	-0.01	-0.04
Mean bunch weight				1.00	0.69	0.66	0.67	-0.45	-0.46	-0.47	-0.46	-0.45	-0.46	-0.35
Estimated yield per vine					1.00	0.66	0.59	-0.29	-0.30	-0.29	-0.29	-0.28	-0.28	-0.15
Exposed leaf area						1.00	0.91	-0.46	-0.46	-0.46	-0.46	-0.45	-0.46	-0.39
Empty spaces in the canopy							1.00	-0.60	-0.60	-0.61	-0.61	-0.59	-0.60	-0.52
Band 2 Blue								1.00	0.99	0.98	0.98	0.99	0.98	0.73
Band 3 Green I									1.00	0.99	0.99	0.99	0.99	0.74
Band 4 Green										1.00	0.99	0.99	0.99	0.75
Band 5 Yellow											1.00	0.99	0.99	0.72
Band 6 Red												1.00	0.99	0.74
Band 7 Red-Edge													1.00	0.76
Band 8 NIR														1.00

LMC includes 3 spatial structures: nugget effect, short-range spherical model with a range of 8 m and longer-range spherical model with a range of 32 m. The results of cross validation testing to check the suitability of the LMC were satisfactory because the mean error was always close to zero and the standardized variance of the kriging error close to 1 and, however, within the interval of tolerance (Chilés and Delfiner, 2012).

The three spatial components account for a different proportion of the total spatial variability 24, 16 and 60 %, respectively. Although the largest proportion is attributable to the longer-range structured component, indicating a spatial association spanning some 20 rows, the micro- and short-range components comprise cumulatively about 40% of the total variance, resulting in low prediction precision. These results therefore emphasise the importance of appropriately planning sampling with a higher average density than that adopted in this study.

Figure 2b presents as an example the directional cokriging map of yield per vine obtained with an interpolation neighbourhood of the same orientation as the variograms. As can be seen, the map is characterised by considerable short-range variability and lacks large macro structures as observed for the PLANET bands (Figure 2a). However, consistent with this, it can be also noted that the western part of the vineyard tends to be less productive than the central and eastern parts.

In order to delineate the vineyard into homogenous areas for possible site-specific management, Table 3 reports the regionalised factor descriptions with eigenvalue greater than 1, omitting those concerning the nugget effect as they are mostly affected by error.

Focusing on the 32-m structure, the two regionalised factors retained cumulatively account for about 88% of the variability at this scale. On the former, the PLANET bands weigh the most and negatively, although to a lesser extent that of the NIR; therefore it might be interpreted as an index of radiometric absorption and thus proportional to leaf growth.

On the latter average number of bunches per vine weighs more and positively. It might then be interpreted as a production index.

Although map in Figure 3b might be used as a guide for differentiated management, it should be considered that it explains only about 25 % of the variability at this scale and furthermore it appears considerably fragmented. There may be various causes for this fragmentariness: essentially the considerable component of unstructured variability due for the most part to the coarse sampling scheme and the numerous anthropogenic operations to which the vineyard was subjected (soil displacement, and pruning).

Table 3. Decomposition of the coregionalization matrix in the regionalized factors for the studied variable.

Variable	Model		
	Spherical range=8 m	Spherical range=32 m	
	F1	F1	F2
Canopy height	0.131	0.058	-0.117
Canopy width	-0.014	-0.023	-0.075
Average number of bunches per vine	0.349	0.199	0.744
Mean bunch weight	-0.469	-0.030	0.178
Estimated yield per vine	0.033	0.117	0.547
Exposed leaf area	-0.178	0.103	0.082
Empty spaces in the canopy	-0.265	0.107	0.177
Band 2 Blue	0.281	-0.384	0.103
Band 3 Green I	0.297	-0.384	0.096
Band 4 Green	0.274	-0.372	0.083
Band 5 Yellow	0.268	-0.393	0.098
Band 6 Red	0.255	-0.405	0.123
Band 7 Red-Edge	0.293	-0.375	0.085
Band 8 NIR	0.270	-0.167	0.011
Eigenvalue	1.043	2.797	1.817
Var. Perc.	74.43	53.18	34.54

Only the regionalized factors (F1, F2) associated to an eigenvalue greater than 1 are reported. The percentage of explained variance (Var. Perc.) is also reported.

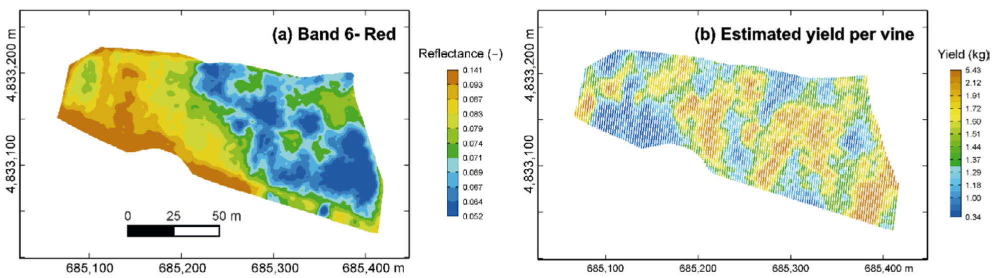

Figure 2. Maps of Planet reflectance of Band 6 Red data at 0.5 m obtained applying deconvolution by kriging (a) and of estimated yield per vine obtained using the data fusion of all variables (b).

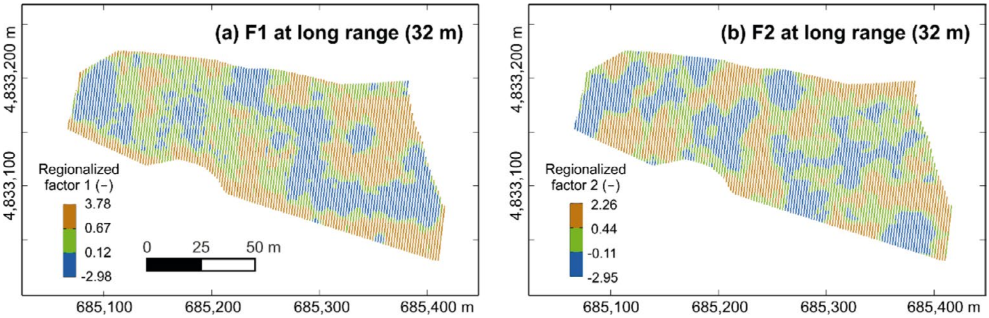

Figure 3. Maps of the first (a) and second (b) regionalized factors (F1 and F2) at longer range. As observed for the map in Figure 2, no clearly delimited homogenous areas can be seen.

The vineyard analysis, supported by thematic maps and satellite data, highlight the impact of land management interventions. Specifically, the movement of soil from the western (higher part of the vineyard) areas towards the eastern (lower part of the vineyard) areas significantly affected soil fertility distribution and then plant growth conditions. This was due to the probable removal to the west of the most pedologically evolved surface horizons and the use of sub-surface horizons with characteristics less suitable for vine cultivation.

The key advantages of remote sensing, particularly in the context of precision viticulture, include repeatability, as it is a non-destructive method for collecting data related to vigor. This approach serves as a valuable complement to traditional, costly field sampling techniques. Although the results require further investigation, evidence suggests that the proposed approach could effectively delineate different vineyard zones within precision viticulture, providing more accurate estimation of canopy influenced also by other factors, such as ground cover vegetation between the rows.

Conclusions

A multivariate geostatistical approach was proposed in order to achieve a cost-effective and efficacious field-delineation into homogenous areas. Although a clear partitioning of the vineyard was not achieved, the method is sufficiently flexible and adaptable to every experimental context. Certainly, the method can be improved by paying particular attention to data collection in the field. Understanding the variation in grape yield and quality prior to the vintage would allow differential management strategies to be adjusted to achieve the desired quantum-quality attributes.

References

Argillier, J.P. (1989). Interdépendance des facteurs de qualité (Interdependence of quality factors). Chambre d'agriculture de l'Hérault, Lattes.

Carbonneau, A. (1995). La surface foliaire exposée potentielle – guide pour sa mesure (Potential exposed leaf area – guide to its measurement). Le Progrès Agricole et Viticole, 9, 204–212.

Castrignanò, A., Costantini, E.A.C., Barbetti, R., & Sollitto, D. (2009). Accounting for Extensive Topographic and Pedologic Secondary Information to Improve Soil Mapping. Catena, 77, 28–38. https://doi.org/10.1016/j.catena.2008.12.004

Castrignanò, A., Wong, M.T.F., Stelluti, M., De Benedetto, D., & Sollitto, D. (2012). Use of EMI, gamma-ray emission and GPS height as multi-sensor data for soil characterisation. Geoderma, 175–176, 78–89. https://doi.org/10.1016/j.geoderma.2012.01.013.

Chilès, J.-P., & Delfiner, P. (2012). Geostatistics: modelling spatial uncertainty, 2nd edn. Wiley, Hoboken, NJ.

Gil, E., Llorens, J., Landers, A., Llop, J., & Giralt, L. (2011). Field validation of Dosaviña, a decision support system to determine the optimal volume rate for pesticide application in vineyards. European Journal of Agronomy, 35, 33–46. https://doi.org/10.1016/j.eja.2011.03.005

Matese, A., Di Gennaro, S.F., Genova, G., Orlandini, S., & Valentini, R. (2015). Intercomparison of UAV, aircraft, and satellite remote sensing platforms for precision viticulture. Remote Sensing, 7, 2971–2990. https://doi.org/10.3390/rs70302971

Pardo-Iguzquiza, E., Rodriguez-Galiano, V.F., Chico-Olmo, M., & Atkinson, P.M. (2011). Image fusion by spatially adaptive filtering using downscaling cokriging. ISPRS Journal of Photogrammetry and Remote Sensing, 66 (3), 337–346. https://doi.org/10.1016/j.isprsjprs.2011.01.001

Pardo-Igúzquiza, E., Chica-Olmo, M., & Atkinson, P.M. (2006). Downscaling cokriging for image sharpening. Remote Sensing of Environment, 102, 86–98, https://doi.org/10.1016/j.rse.2006.02.014

Patrignani, A., & Ochsner, T.E. (2015). Canopeo: A powerful new tool for measuring fractional green canopy cover. Agronomy Journal, 107, 2312–2320. https://doi.org/10.2134/agronj15.0150

Rivoirard, J. (2001). Which models for collocated cokriging? Mathematical Geology, 33, 117–131. https://doi.org/10.1023/A:1007530600152

Wackernagel, H. (2003). Multivariate Geostatistics: an introduction with applications. 3rd edition. Springer, Berlin.