

Clustering of grape yield maps to delineate site-specific management zones

J. Arno^{1*}, J. A. Martínez-Casasnovas², M. Ribes-Dasi¹ and J. R. Rosell¹

¹ Departamento de Ingeniería Agroforestal

² Departamento de Medio Ambiente y Ciencias del Suelo. Universidad de Lleida.
C/ Rovira Roure, 191. 25198 Lleida. Spain

Abstract

Zonal management in vineyards requires the prior delineation of stable yield zones within the parcel. Among the different methodologies used for zone delineation, cluster analysis of yield data from several years is one of the possibilities cited in scientific literature. However, there exist reasonable doubts concerning the cluster algorithm to be used and the number of zones that have to be delineated within a field. In this paper two different cluster algorithms have been compared (*k*-means and fuzzy *c*-means) using the grape yield data corresponding to three successive years (2002, 2003 and 2004), for a 'Pinot Noir' vineyard parcel. Final choice of the most recommendable algorithm has been linked to obtaining a stable pattern of spatial yield distribution and to allowing for the delineation of compact and average sized areas. The general recommendation is to use reclassified maps of two clusters or yield classes (low yield zone and high yield zone) and, consequently, the site-specific vineyard management should be based on the prior delineation of just two different zones or sub-parcels. The two tested algorithms are good options for this purpose. However, the fuzzy *c*-means algorithm allows for a better zoning of the parcel, forming more compact areas and with more equilibrated zonal differences over time.

Additional key words: cluster analysis; fuzzy *c*-means algorithm; grape yield maps; *k*-means algorithm; precision viticulture; zonal management.

Resumen

Análisis cluster de mapas de vendimia para la delimitación de zonas de manejo diferencial

El manejo zonal en viña requiere la delimitación previa de zonas estables de cosecha dentro de la parcela. Entre los diferentes métodos de zonificación existentes, el análisis cluster de los datos de cosecha de diferentes años es una de las posibilidades que han sido citadas en la literatura científica. Sin embargo, existen dudas razonables en relación al algoritmo que debe ser utilizado y al número de zonas que deben delimitarse dentro de la parcela. En este artículo se comparan dos algoritmos diferentes de clasificación (algoritmo de *k*-medias y algoritmo difuso de *c*-medias), a partir de los datos de cosecha (mapas de cosecha) correspondientes a tres años sucesivos de vendimia (2002, 2003 y 2004) en una parcela de uva 'Pinot Noir'. La elección final del método (algoritmo) ha estado vinculada a la obtención de un patrón de cosecha estable en el tiempo y a la delimitación dentro de la parcela de zonas compactas y de cierto tamaño. La recomendación general es la utilización de mapas reclasificados en dos clases de cosecha (zona de cosecha baja y zona de cosecha alta) y, en consecuencia, el manejo sitio-específico en viña debería estar basado en la delimitación previa de únicamente dos zonas diferentes o subparcelas. Los dos algoritmos evaluados son una buena opción para este propósito. Con todo, el algoritmo difuso de *c*-medias permite una mejor zonificación de la parcela, formando áreas que, siendo más compactas en superficie, mantienen asimismo mejor equilibradas las diferencias zonales a lo largo del tiempo.

Palabras clave adicionales: algoritmo difuso de *c*-medias; algoritmo de *k*-medias; análisis cluster; manejo zonal; mapas de vendimia; viticultura de precisión.

* Corresponding author: jarno@eagrof.udl.es

Received: 25-11-10. Accepted: 30-06-11.

Abbreviations used: FPI (fuzziness performance index); H (classification entropy); MZA (management zone analyst); NCE (normalized classification entropy); PV (precision viticulture).

Introduction

Grape yield is variable within the same parcel. This spatial variability normally presents little changes from one campaign to the next and follows a clearly defined spatial distribution pattern, even when there have been significant differences in total grape yield between consecutive years (Bramley and Hamilton, 2004). Then, the delineation of differential management zones using historical series of grape yield maps offers winegrowers and winemakers an excellent opportunity to improve the site-specific management of vineyards.

Yield sensors and monitors supply reliable and geo-referenced values for the grape harvest (Arnó *et al.*, 2009). Once the maps have been constructed, various methodologies can be used for zonal delineation of a parcel using yield data from several years (Blackmore, 2000; Diker *et al.*, 2004; Bocchi and Castrignanò, 2007). However, cluster analysis, used on a series of yield maps, is one of the most recommended classification methodologies to allow zoning at parcel level (Lark and Stafford, 1997; Lark, 1998; Panneton *et al.*, 2001; Shatar and McBratney, 2001; Bramley and Hamilton, 2004; Ping and Dobermann, 2005; Taylor *et al.*, 2007). Through an iterative process, this procedure enables the clustering of values interpolated from yield maps into homogenous groups (classes). The number of groups can be predetermined, but the delineation of 2 to 5 classes is the general recommendation (Bramley and Hamilton, 2004). The aim is to zone the parcel taking into account the classes provided by cluster analysis.

Another possibility is to use, as variables for the cluster analysis, yield, soil data (*e.g.* the apparent soil electrical conductivity), remote sensing images and site characteristics together. In most cases, this possibility has proved to be an interesting way to identify within-field homogeneous and stable zones (Cupitt and Whelan, 2001; Boydell and McBratney, 2002; Taylor *et al.*, 2003; King *et al.*, 2005; Yan *et al.*, 2007; Guastaferro *et al.*, 2010). The delineation of management zones only based on soil fertility variables is also proposed by Ortega and Santibáñez (2007).

In a multivariate cluster analysis of grape yield maps, zones of differential crop management can vary within a parcel, to a greater or lesser degree, depending on whether the hard *k*-means algorithm or the fuzzy *c*-means algorithm is used. Although the particularities how these methods work for different applications are well known (*e.g.* Guastaferro *et al.*, 2010), the objective of the present paper is to determine which algorithm

is the most appropriate for zoning purposes in precision viticulture (PV). For that, both procedures were applied and compared using the grape yield data corresponding to three successive years. Final choice of the most recommendable algorithm was possible through verification of the temporal stability of the spatial yield distribution patterns. Notwithstanding the above, it should be remembered that, as stated by King *et al.* (2005), the use of yield maps should always be complemented by the owner's experience and field accumulated knowledge.

Material and methods

Data acquisition

The yield data used in the study correspond to the harvests of the years 2002, 2003 and 2004 of a vineyard parcel (P30, *Vitis vinifera* L. cv. Pinot Noir) located in Raimat (Lleida, Spain). This parcel has a total surface area of 5.0 ha and was planted in 1985 with a 2.1 m vine spacing and 3.2 m row spacing pattern. The vineyard is sprinkler-irrigated.

The monitoring equipment was, for each year of the study, the *Canlink 3000* yield monitor (Farmscan, Bentley, Western Australia). The exact harvesting dates were September 16 (2002), September 4-6 (2003) and August 24 (2004).

While the yield monitor was configured in 2002 to capture data each 4 s, in 2003 and 2004 this data capture interval was raised to 5 s. This enabled a reduction in the number of monitored observations to a value closer to the number of interpolations which were subsequently used to construct the yield maps.

Grape yield mapping

Local block kriging based on local variograms was used to map the grape harvest over a predefined 3 m grid surface. The interpolated yield values were, therefore, directly comparable as they referred to the same points or co-ordinates. Spatial interpolation was carried out using *VESPER*, version 1.6.3, geostatistical software (Minasny *et al.*, 2005).

While the yield maps for 2002 and 2003 were constructed from the original data supplied by the monitor, the 2004 map was obtained from interpolated yield values from the first map provided by the Codorníu

Company. Since the reference grid of the Codorníu map did not coincide exactly with the grid used in 2002 and 2003, the yield values for the desired locations had to be obtained from values already interpolated at other different locations.

Clustering and statistical analysis

A cluster analysis of the study parcel was carried out using the values interpolated from the three grape yield maps together (2002, 2003 and 2004). This procedure allowed, first, the construction of a reclassified yield map according to a predefined number of groups or yield classes. Clustering was initially performed with the k -means algorithm. In a second step, and following the methodology proposed by Bramley and Hamilton (2004), the temporal stability of grape yield pattern was tested using an analysis of variance. The separate ANOVA of yield for each year, in relation to the zones or clusters (analysis factor), should allow observation of any significant differences between the various yield zones, with a means separation which should also display the same tendency (ranking) for each of the years under study. The data analysis procedure is shown in Figure 1.

To compare the results of the k -means algorithm, the same procedure was repeated using a fuzzy logic clustering method: the fuzzy c -means algorithm. This algorithm has been widely used in the delineation of site-specific management zones in extensive crop-farming (a comparison of different algorithms can be

found in Guastaferro *et al.*, 2010), but no study has been published to date of its use in precision viticulture.

Unlike the k -means algorithm (or other deterministic clustering methods like ISODATA), in which each observation can only belong to a single group or cluster, the fuzzy c -means algorithm allows more than one group or cluster to be assigned to each individual, according to its characteristics and with different degrees of belonging. That is, fuzzy clustering methods assign to each individual a partial class membership value to each one of the classes or groups. The application of this type of classification in precision viticulture is particularly interesting. Given, on the one hand, the continuous nature of the spatial variation of the grape yield (since it is influenced in general by physical environment factors which are in themselves spatially continuous) and, on the other, the spatial imprecision of the information supplied by the yield monitors, using a yield classification method which takes into consideration the continuity of the different classes could result in a more representative (or less artificial) zoning of the parcel. In this particular case, the aim was to determine the degree of belonging of each individual (map pixel) to each of the yield classes which would result from zoning of the parcel. The class in which each pixel obtained the highest score (assignment) was the class which would finally appear on the reclassified grape yield map (“defuzzification” process). Cluster analysis using the k -means algorithm was carried out with SAS Enterprise Guide (Statistical Analysis System Institute Inc., Cary, NC, USA). Yield classifi-

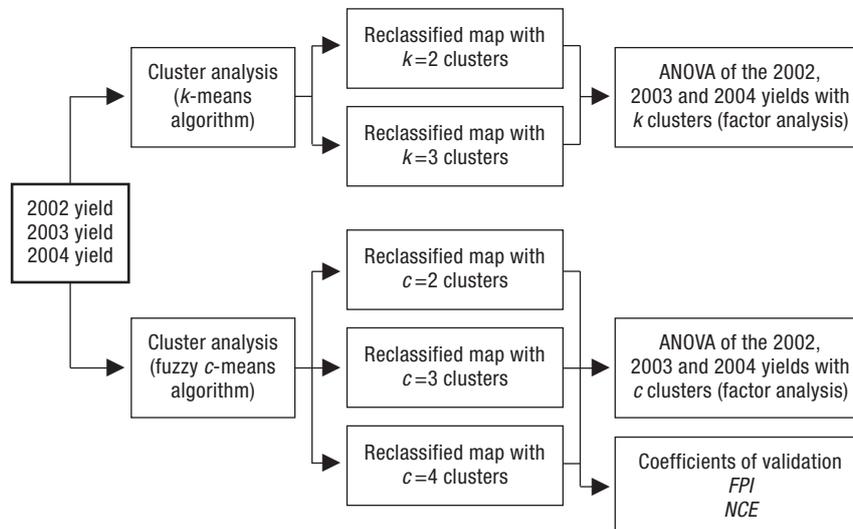


Figure 1. Methodology used to check the temporal stability of the yield maps (*FPI*: fuzziness performance index, *NCE*: normalized classification entropy).

cation using the fuzzy c -means algorithm was performed with the software *Management Zone Analyst (MZA)* version 1.0.1 (Fridgen *et al.*, 2004).

The optimal fuzzy classification of yield data (n observations) into yield groups (c clusters) is the classification which minimizes the objective function J_m (Fridgen *et al.*, 2004; Yan *et al.*, 2007):

$$J_m = \sum_{i=1}^n \sum_{j=1}^c (u_{ij})^m (d_{ij})^2 \quad [1]$$

where m is the fuzziness exponent [m is usually set in the range $1 < m \leq 2$ (Lark and Stafford, 1997); in this particular case m was set to 1.30], $(d_{ij})^2$ the squared distance between observations and centroids (clusters), and u_{ij} the membership value assigned to pixel i within the cluster j . The imposed conditions are shown in [2], and must be met for any $i = 1$ to n and for any $j = 1$ to c :

$$u_{ij} \in [0,1] \quad \forall i,j \quad ; \quad \sum_{j=1}^c u_{ij} = 1, \quad \forall i \quad [2]$$

As can be seen in [1], the fuzzy c -means algorithm introduces the use of a weighting exponent (fuzzy exponent), the function of which was to control the degree of overlap established between the groups or clusters. The algorithm also used the concept of distance between points to evaluate the similarity (proximity) between observations and the centroids of each group. In this respect, the distance used in expression [1] (Mahalanobis distance) was calculated through the matrix product:

$$d_{ij} = [(y_i - v_j)' S^{-1} (y_i - v_j)]^{1/2} \quad [3]$$

where y_i represents the vector of the i^{th} observation of the data, v_j the means vector (centroid) of the cluster j , and S the matrix of variances and covariances of the original data. In this way, Mahalanobis distance, as a measure of similarity, enabled a “bringing together” or “separation” of the points, taking into consideration the existing degree of correlation between the variables (yields of 2002, 2003 and 2004).

Validation of the zoning was performed through the calculation of two coefficients: the Fuzziness Performance Index (FPI) and the Normalized Classification Entropy (NCE).

The FPI measures the degree of separation (or overlap) between the partitioned groups of the original data matrix (Fridgen *et al.*, 2004). It was calculated from the following expression:

$$FPI = 1 - \frac{cF - 1}{c - 1} \quad [4]$$

where,

$$F = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^c (u_{ij})^2 \quad [5]$$

Values range between 0 and 1. Those values approaching 0 were indicative of distinct classes with little overlap between the groups, while values approaching 1 indicated non-distinct classes or, in other words, classes with a very high degree of overlap.

The NCE estimates the amount of disorganisation created by the fuzzy partition of the data matrix with a specific number of clusters or yield classes (Lark and Stafford, 1997). Firstly, the classification entropy (H) was calculated through the expression:

$$H = - \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^c u_{ij} \cdot \log_a (u_{ij}) \quad [6]$$

where the logarithmic base a could be any positive integer, with the values of H varying from 0 to $\log_a (c)$. From [6], the final NCE value was given by:

$$NCE = \frac{H}{[1 - (c/n)]} \quad [7]$$

The NCE were similar to those of entropy when c was relatively small in comparison with n . As before, the NCE values approaching 0 were indicative of a more appropriate classification (or with a higher degree of organisation). The optimum number of clusters was, therefore, the number which managed to minimize the two proposed coefficients.

Results

The 2002, 2003 and 2004 yield maps for parcel 30 (Pinot Noir cultivar) are shown in Figure 2. The mean yield gave different values for the three years under study (2002, 2003 and 2004). According to the coefficient of variation values, the 2002 grape yield (Table 1) displayed greater within-parcel variation than those of 2003 and 2004. However, its spatial distribution followed similar patterns in the three campaigns (Fig. 2), with no apparent influence from the differing total production for each of the years. A visual evaluation seemed to indicate a higher similarity between the 2003 and 2004 maps. It could be considered that the yields (maps) of the years 2003 and 2004 displayed more similarity with each other than with the 2002 map.

A cluster analysis was performed to verify the stability of the spatial variation of the yield, with the analysis variables being the respective 2002, 2003 and

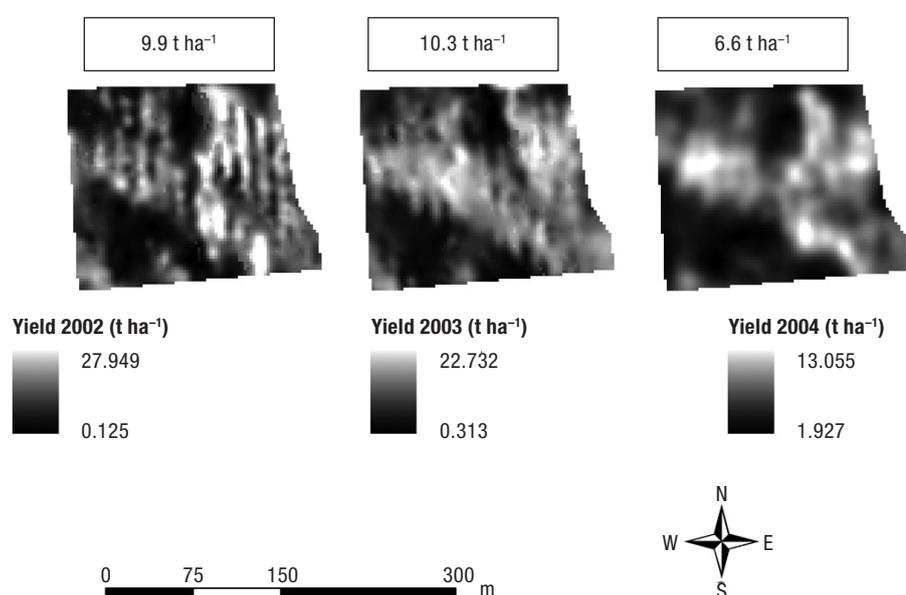


Figure 2. Yield maps of parcel 30 (Pinot Noir cultivar) for the years 2002, 2003 and 2004 (the average yields of each year are highlighted in boxes). The 2004 map was obtained from values interpolated from the yield map provided by the Codorniu Company.

2004 yields. Using the unsupervised k -means classification algorithm, a conventional (or hard) zoning was obtained in which each of the classified individuals (5323 points or pixel values) could only belong to one group or yield class. Figure 3 shows the two yield zoning maps: on one map two production zones can be distinguished (cluster 1 and cluster 2), and on the other three zones (clusters 1, 2 and 3). The fuzzy c -means algorithm was used as an alternative to this method with up to three different maps obtained depending on the number of groups used (two, three and four different zones) (Fig. 3). In this way, by considering the possibility that each point could belong to a higher or lower degree to more than one yield class, the resulting zoning reflected the possible inherent variability of the yield at a same location.

Once the different zonings of parcel 30 had been obtained (two with the k -means algorithm and three

with the fuzzy c -means algorithm, Fig. 3), an analysis of variance (ANOVA) was performed, for each zoning and for each of the years under study, in which the factor considered was the cluster or yield zone. The final results are shown in Table 2 (k -means algorithm) and Table 3 (fuzzy c -means algorithm). Finally, for the latter method, the values of the FPI and NCE indices are shown in Figure 4.

Discussion

In Table 2, the means separation offered a very clear interpretation. For the three years studied, the zones of the parcel classified as cluster 2 (Fig. 3) were the zones where the lowest yield was obtained. Likewise, the zones within the parcel which had been classified as cluster 1 were the zones of highest productive potential.

Table 1. Descriptive analysis of the values interpolated from the 2002, 2003 and 2004 yield maps in parcel 30 (Pinot Noir cultivar)

	2002	2003	2004
Number of interpolations	5,323	5,323	5,323
Mean (t ha ⁻¹)	9.905	10.303	6.645
Median (t ha ⁻¹)	9.863	10.753	6.594
Minimum (t ha ⁻¹)	0.125	0.313	1.928
Maximum (t ha ⁻¹)	27.949	22.732	13.055
Standard deviation (t ha ⁻¹)	5.511	4.604	2.679
Coefficient of variation, CV (%)	55.64	44.68	40.31

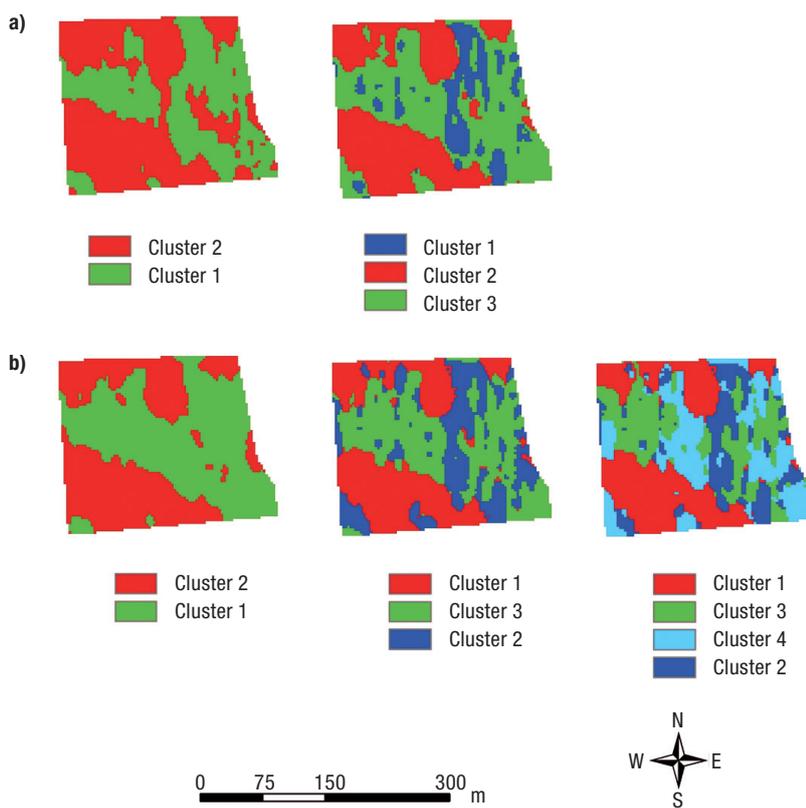


Figure 3. Zoning of parcel 30 (Pinot Noir cultivar) based on classification of the yields of 2002, 2003 and 2004 by means of cluster analysis: a) *k*-means and b) fuzzy *c*-means algorithms.

It was clear that the subdivision of the parcel into two clusters distinguished two zones (low and high yield), whose within-parcel distribution remained constant from one campaign to the next.

The results were almost identical when zoning was extended to three clusters or crop zones. Identification of a three-tiered spatial pattern of the yield (clusters 1, 2 and 3, Fig. 3) also displayed clear stability or continuity for the three years under study (2002, 2003 and 2004). Though no significant differences were observed between cluster 3 and cluster 1 for the 2003

harvest, the temporal stability of the yield pattern enabled identification of cluster 2 as the zone of lowest production, cluster 3 as a zone of intermediate yield and cluster 1 as the zone of highest production.

In short, it has been demonstrated that the grape yield was not only variable within the parcel, but that this spatial variability followed a particular distribution pattern which remained stable over time. This result is undoubtedly of enormous importance since, in principle, it justifies site-specific crop management (SSCM) of the vineyard based on within-parcel delineation of

Table 2. Zoning of parcel 30 (Pinot Noir cultivar) based on grouping of the yield for 2002, 2003 and 2004 using the unsupervised *k*-means algorithm

Production zone	Number of points	Yield 2002	Yield 2003	Yield 2004
Cluster 2	2,929	6.083 ^b	7.362 ^b	4.924 ^b
Cluster 1	2,394	14.581 ^a	13.901 ^a	8.751 ^a
Cluster 2	2,074	4.864 ^c	5.654 ^b	4.197 ^c
Cluster 3	2,417	11.186 ^b	13.262 ^a	7.784 ^b
Cluster 1	832	18.749 ^a	13.297 ^a	9.440 ^a

In the same column, means with the same letter are not significantly different for a 95% confidence level.

Table 3. Zoning of parcel 30 (Pinot Noir cultivar) based on grouping of the 2002, 2003 and 2004 yields through the unsupervised fuzzy *c*-means algorithm

Production zone	Number of points	Yield 2002	Yield 2003	Yield 2004
Cluster 2	2,387	6.286 ^b	6.014 ^b	4.307 ^b
Cluster 1	2,936	12.847 ^a	13.790 ^a	8.547 ^a
Cluster 1	1,869	4.633 ^c	5.356 ^c	4.091 ^c
Cluster 3	2,023	10.416 ^b	14.113 ^a	8.295 ^a
Cluster 2	1,431	16.068 ^a	11.379 ^b	7.650 ^b
Cluster 1	1,719	4.544 ^d	5.108 ^d	3.921 ^d
Cluster 3	1,264	9.909 ^c	13.152 ^b	9.180 ^a
Cluster 4	1,370	11.509 ^b	13.623 ^a	6.427 ^c
Cluster 2	970	17.134 ^a	11.109 ^c	8.479 ^b

In the same column, means with the same letter are not significantly different for a 95% confidence level.

zones of different productive potential. The high value of the coefficient of variation (Table 1) also reinforces this possibility.

The use of the fuzzy *c*-means algorithm for the zoning of parcel 30 gave similar results (Fig. 3 and Table 3). However, the peculiarity of the method allowed some aspects to be highlighted which had gone unnoticed when the *k*-means algorithm had been used as partition method.

The main conclusion from the analysis of variance (Table 3) was that the persistence over time of the spatial distribution pattern of the yield was only evident when the parcel was divided into two clusters of different yield. In the other cases (3 and 4 clusters), only cluster 1 (when 3 and 4 zones were delineated) corresponded to the area of less production for each of the years under study. However, the ranking of the remain-

ing clusters (though significantly different) displayed no consistency for the series of years analysed. This result, which coincides with some of the tests conducted by Bramley and Hamilton (2004), implies it would be advisable to subdivide parcel 30 (Pinot Noir cultivar) into just two zones or sub-parcels (low yield zone of cluster 2 and high yield zone of cluster 1), given that this zoning is the only one which guarantees a yield spatial variation pattern which is stable over time.

As mentioned above, the yield maps for 2003 and 2004 displayed higher similarity to each other (Fig. 2) than they seemed to display to the 2002 map. Confirmation of this was found in the fact that the correlation (measured as Pearson's linear correlation coefficient) between the 2003 and 2004 maps ($\rho = 0.770$) was higher than that found between the 2002 and 2003 maps ($\rho = 0.624$), and between the 2002 and 2004 maps ($\rho = 0.691$). This would explain the behaviour between clusters 2 and 3 in the zoning which followed three classes of yield (Table 3). In 2002, the highest production potential zone corresponded to cluster 2, while the spatial variation pattern of the yield inverted clusters 3 and 2 and remained stable from 2003 to 2004. However, the most interesting pattern occurred when the parcel was subdivided into 4 clusters (Table 3). Again, there was a zone (now classified as cluster 1) which was stable over time and for which the lowest yields were obtained. However, the remaining clusters were ranked differently for each of the years under study. The most logical interpretation of this leads to the same conclusion as before, namely that the subdivision of parcel 30 (Pinot Noir cultivar) into just two zones of different productive potential seems to be the option which obtains better stabilisation over time of the reclassified grape yield map.

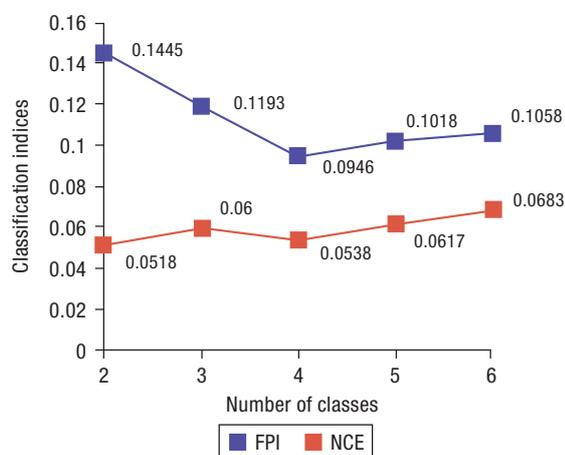


Figure 4. Validation of the zoning of parcel 30 (Pinot Noir cultivar) based on the evaluation of the *FPI* (fuzziness performance index) and *NCE* (normalized classification entropy) indices.

Normally the *FPI* and *NCE* values should fall as the number of classes increases, and it is considered that the optimum number of clusters (or zones of different yield) is the number which is able to minimize the respective values of these two indices. While minimization of the *FPI* was useful because the overlap between groups decreased, minimization of the *NCE* was interesting because the degree of organisation increased as a consequence of data partitioning. In our case (Fig. 4), the number of yield classes recommended by the *FPI* criterion (4 clusters) differed from the optimum number of classes obtained through the *NCE* criterion (2 clusters). So, the initial expectation of obtaining a matching result for both indices was not met [Boydell and McBratney (2002) in cotton, and Ping and Dobermann (2005) in corn, obtained similar results].

Since there was no agreement between the *FPI* and *NCE* indices as to the optimum number of clusters (Fig. 4), one option was to choose the least number of classes (Lark and Stafford, 1997), which in this case was two. It is also true that the *FPI* values for two and four clusters were not so much different, as the index can vary from 0 to 1 (in both cases, two and four clusters, the overlap between groups was small). The same interpretation is valid for the *NCE* coefficient, or even more because the values were very similar besides of the number of classes (Fig. 4). Likewise, the choice of four classes, in view of the results shown in Table 3 (ANOVA), would not seem to be a good solution. Zoning of the parcel with two sub-parcels is possibly the most suitable option. The simplification thereby achieved is clear and favours subsequent field management. In addition, this result coincides with the recommendation made by Taylor *et al.* (2003), in the sense of selecting a small number of zones even though the zoning may not resolve certain differences (or the differences remain unseen).

It is clear that the use of reclassified grape yield maps (constructed on a historical series of yield) requires, first of all, temporal stability of the zones delineated within the parcel. So, faced with the decision as to which classification method to use, the most interesting zoning would be the one which fulfilled the temporal stability criterion and which was able to subdivide the parcel into two zones, whose differential yield remained as uniform and consistent as possible over successive campaigns. If we compare the ratios between the high yield zone (cluster 1) and the low yield zone (cluster 2) using first the zones delineated through the *k*-means algorithm (Table 2) (2.4:1 for

2002, 1.89:1 for 2003 and 1.78:1 for 2004), and then the zones delineated through the *c*-means algorithm (Table 3) (2.04:1 for 2002, 2.3:1 for 2003 and 1.98:1 for 2004), we can see that they produce different results, but the values for the *c*-means classification method are closer to each other. This indicates that the fuzzy *c*-means method was able to delineate two zones (clusters 1 and 2) in the study parcel which, in addition to showing significantly different yield levels, also displayed a productive ratio which remained appreciably constant over the years. So, in agreement with Lark (1998), classification using the continuous *c*-means algorithm achieved, in comparison with the *k*-means algorithm, a more coherent zoning of the parcel, forming more compact regions and with a spatial distribution which should facilitate the work of site-specific management.

The present research has shown that cluster analysis is an appropriate method for obtaining classified grape yield maps, and interesting application in precision viticulture. The spatial distribution pattern of the grape yield remains appreciably constant from one campaign to the next, with classified maps in two clusters or yield classes (low yield zone and high yield zone) displaying higher temporal stability. Consequently, the application of possible site-specific management in the vineyard should be based on the prior delineation of just two types of different zones or sub-parcels.

Compared with the *k*-means algorithm, the fuzzy *c*-means algorithm is better at equilibrating possible zonal differences over time, allowing for greater consistency in site-specific management. Then, fuzzy clustering should be rather used when delineation of management zones is conducted in winegrape production systems.

Acknowledgements

The authors would like to thank the *Ministerio de Ciencia e Innovación* and FEDER (*Fondo Europeo de Desarrollo Regional*) for funding part of this research through the *OPTIDOSA* project (AGL2007-66093-C04-03/AGR). Special thanks are also given to the Agricultural Division of Codorníu for the data and technical support.

References

ARNÓ J., MARTÍNEZ-CASASNOVAS J.A., RIBES-DASI M., ROSELL J.R., 2009. Review. Precision viticulture.

- Research topics, challenges and opportunities in site-specific vineyard management. *Span J Agric Res* 7(4), 779-790.
- BLACKMORE S., 2000. The interpretation of trends from multiple yield maps. *Comput Electron Agric* 26, 37-51.
- BOCCHI S., CASTRIGNANÒ A., 2007. Identification of different potential production areas for corn in Italy through multitemporal yield map analysis. *Field Crop Res* 102, 185-197.
- BOYDELL B., McBRATNEY A.B., 2002. Identifying potential within-field management zones from cotton-yield estimates. *Precis Agric* 3, 9-23.
- BRAMLEY R.G.V., HAMILTON R.P., 2004. Understanding variability in winegrape production systems. I. Within vineyard variation in yield over several vintages. *Aust J Grape Wine Res* 10, 32-45.
- CUPITT J., WHELAN B.M., 2001. Determining potential within-field crop management zones. *Proc III ECPA-European Conference on Precision Agriculture*, Montpellier, France, June 18-21. pp. 7-12.
- DIKER K., HEERMANN D.F., BRODAHL M.K., 2004. Frequency analysis of yield for delineating yield response zones. *Precis Agric* 5, 435-444.
- FRIDGEN J.J., KITCHEN N.R., SUDDUTH K.A., DRUMMOND S.T., WIEBOLD W.J., FRAISSE C.W., 2004. Management zone analyst (MZA): software for subfield management zone delineation. *Agron J* 96, 100-108.
- GUASTAFERRO F., CASTRIGNANÒ A., DE BENEDETTO D., SOLLITTO D., TROCCOLI A., CAFARELLI B., 2010. A comparison of different algorithms for the delineation of management zones. *Precis Agric* 11, 600-620.
- KING J.A., DAMPNEY P.M.R., LARK R.M., WHEELER H.C., BRADLEY R.I., MAYR T.R., 2005. Mapping potential crop management zones within fields: use of yield-map series and patterns of soil physical properties identified by electromagnetic induction sensing. *Precis Agric* 6, 167-181.
- LARK R.M., 1998. Forming spatially coherent regions by classification of multivariate data: an example from the analysis of maps of crop yield. *Int J Geogr Inf Sci* 12(1), 83-98.
- LARK R.M., STAFFORD J.V., 1997. Classification as a first step in the interpretation of temporal and spatial variation of crop yield. *Ann Appl Biol* 130, 111-121.
- MINASNY B., McBRATNEY A.B., WHELAN B.M., 2005. VESPER version 1.62. Australian Centre for Precision Agriculture, McMillan Building A05, The University of Sydney, NSW 2006. Available in <http://www.usyd.edu.au/su/agric/acpa> [15 November, 2010].
- ORTEGA R.A., SANTIBÁÑEZ O.A., 2007. Determination of management zones in corn (*Zea mays* L.) based on soil fertility. *Comput Electron Agric* 58, 49-59.
- PANNETON B., BROUILLARD M., PIEKUTOWSKI T., 2001. Integration of yield data from several years into a single map. *Proc III ECPA-European Conference on Precision Agriculture*, Montpellier, France, June 18-21. pp. 73-78.
- PING J.L., DOBERMANN A., 2005. Processing of yield map data. *Precis Agric* 6, 193-212.
- SHATAR T.M., McBRATNEY A.B., 2001. Subdividing a field into contiguous management zones using a k-zone algorithm. *Proc III ECPA-European Conference on Precision Agriculture*, Montpellier, France, June 18-21. pp. 115-120.
- TAYLOR J.C., WOOD G.A., EARL R., GODWIN R.J., 2003. Soil factors and their influence on within-field crop variability, part II: spatial analysis and determination of management zones. *Biosyst Eng* 84 (4), 441-453.
- TAYLOR J.A., McBRATNEY A.B., WHELAN B.M., 2007. Establishing management classes for broadacre agricultural production. *Agron J* 99(5), 1366-1376.
- YAN L., ZHOU S., FENG L., HONG-YI L., 2007. Delineation of site-specific management zones using fuzzy clustering analysis in a coastal saline land. *Comput Electron Agric* 56, 174-186.