EGU General Assembly 2020
May 2020

Uncertainties associated with the delineation of management zones in precision agriculture

Tomás R. Tenreiro[a], Margarita García-Vila[b], Jose A. Gómez[a], Elías Fereres[a,b]

[a]Institute for Sustainable Agriculture (CSIC), Córdoba, Spain

[b]Department of Agronomy, University of Córdoba, Córdoba, Spain

Abstract

The characterization of spatial variations in soil properties and crop performance within precision agriculture, and particularly the delineation of management zones (MZ) and sampling schemes, are complex assignments currently far from being resolved. Considerable advances have been achieved regarding the analysis of spatial data, but less attention has been devoted to assess the temporal asymmetry associated with variable $crop \times year$ interactions. In this case-study of a 9 ha field located in Spain, we captured interactions between both spatial and temporal variations for two contrasting seasons of remotely sensed crop data (NDVI) combined with several geomorphological properties (i.e., elevation, slope orientation, soil apparent electrical conductivity - ECa, %Clay, %Sand, pH). We developed an algorithm combining Principal Component Analysis (PCA) and clustering k-means and succeeded to delineate four MZ's with a satisfactory fragmentation degree, each one associated with a different $Elevation \times ECa \times NDVI$ combination. Simulated yield maps were generated using NDVI maps correlated to ground cover to establish initial conditions in simulation settings with a crop model. Yield maps were spatially correlated but fitted into variograms with irregular spatial structure. Both CV and spatial patterns did not show consistency from year to year. The results indicate that MZ's temporal instability is an important issue for site-specific management as agronomic implications varied greatly with $crop \times year$ setting. We observed differences, not only regarding NDVI patterns but also in yield response to the combination of $Elevation \times ECa$ (and Texture) depending on the seasonal rainfall. A reduction of 14% of the 'Goodness of Variance Fit' was observed for simulated yield from the first to the second $crop \times year$, highlighting the difficulties in the delineation of MZ's with persistent confidence. The interpretation of $MZ \times Yield$ associations was not straight forward from the metrics selected here as it also depended on agronomic knowledge. We believe that precision agriculture will benefit greatly from improved protocols for MZ delineation and sampling schemes. However, the uncertainty associated with temporal asymmetry of yield clustering and MZ's interpretation reveals that 'automated digital agricultural systems' are still far from reality.[1]

1 Introduction

The characterization of spatial variation within precision agriculture, and particularly the delineation of management zones (MZ), is a complex assignment and it has been a point of discussion for many (Crawford et al., 1997; Moral et al., 2010; Milne et al., 2012; Schemberger et al., 2017). One of the most common approaches used for MZ delineation is the recognition of homogeneous groups in spatial data using cluster analysis. However, this is far from being straightforward because the outcomes can be strongly affected by biased considerations and the methods used to select explanatory variables have not been standardized in literature (Leroux and Tisseyre, 2019). In this context, the delineation of MZ and sampling schemes is not a simple task, mostly when spatial patterns of crop performance are strongly affected by $crop \times year$ interactions.

Our objective was to conduct a clustering analysis on a $crop \times year$ contrasting dataset, typically observed in annual cropping rotation schemes, to measure spatial variation and to discuss the uncertainties associated with unstable temporal-spatial correlations. The algorithm code (R-studio V3.6) is entirely available at https://github.com/RoquetteTenreiro/MZ/blob/master/script.md, as well as input and output data.

2 Materials & Methods

The experimental field, located in Córdoba, southern Spain (Mediterranean climate), is characterised by an expansive clay vertisol, 1.2-1.6 m depth and decent water storage capacity (i.e. > 0.14 cm^3 cm^{-3}). An area of 9.5 ha was delineated from a flow direction raster obtained with the SAGA - Wang Liu algorithm (Wang and Liu, 2006), from a DEM with 5 m spatial resolution collected with LiDAR and available at CNIG (http://centrodedescargas.cnig.es/CentroDescargas/index.jsp).

In order to capture interactions between spatial and temporal variability (Mulla and Schepers, 1997), two contrasting seasons of crop data were combined (i.e. winter wheat in 2017/18 and rapeseed in 2018/19), the first characterised by a satisfactory rainfall supply (>600 mm), the second by water shortage (<400 mm with 80% falling in early autumn). NDVI's of late phenological vegetative and flowering stages (Cattani et al., 2017; Scudiero et al., 2014), were estimated from atmospherically corrected satellite data, with cloud cover <4%, downloaded from https://apps.sentinel-hub.com/.

An electromagnetic induction sensor was used to measure soil ECa (dS m^{-1}) at 35 and 85 cm depth and with a spatial resolution of 1x15 m (Johnson et al., 2001), performed before sowing date and four days after a rainfall event of approximately 10 mm (McCutcheon et al., 2006). Soil samples (%Clay, %Sand, pH) were collected at 35 cm depth in ten point-sites according to the ECa pattern.

Within-field spatial co-variation was accessed with Principal Component Analysis (PCA), correlated principal components were selected (i.e. Elevation, deep ECa, and NDVI at specific dates), and a k-means clustering was applied by setting an optimal amount of three clusters according to Rousseeuw (1987) and Tibshirani et al. (2001). Means and standard deviations were estimated for each variable within each MZ as well as the 'Goodness of Variance Fit' (GVF).

Moran's I method and Z-scores (Bivand and Wong, 2018) were estimated for yields simulated with AquaCrop V6.1. The simulations were conducted under a standard crop parameterization for a typical clay soil (single horizon, 1.4 m depth), using weather data obtained from https://www.juntadeandalucia.es/agriculturaypesca/ifapa/ria/servlet/FrontController?action=Static&url=coordenadas.jsp&c_provincia=14&c_estacion=6 and considering different initial conditions according to the relations between NDVI and canopy cover reported by Goodwin et al. (2018), Han et al. (2017) and Luo et al. (2015). An analysis of spatial autocorrelation was also conducted to measure simulated yield spatial dependence.

3 Results & Discussion

The algorithm delineated three MZ's, each one associated to a different $Elevation \times ECa \times NDVI$ combination. The GVF values were higher than 50% for all principal components and, despite being slightly lower than values reported in other studies (Peeters et al., 2015; Scudiero et al., 2018), most variables had GVF comprehended in acceptable ranges.

According to Anselin (1995); Bivand and Wong (2018) and Kalogirou (2019), the results obtained from Moran's I method (I_{index} = 0.91 and Z_{score} = 72.3) were satisfactory too. Simulated yields were spatially correlated but fitted into variograms with irregular spatial structure according to Leroux and Tisseyre (2019). Both yield CV and spatial patterns did not show temporal stability.

The euclidean space represented in the PCA plot indicated that deep ECa was better correlated to NDVI than superficial ECa, which is in line with the principle that rainfed crops in our conditions, depend largely on the water availability at deeper soil layers during critical periods of stress (i.e. flowering and fruit development). The correlation was positive in 2019 but negative in 2018. These contrasting correlation signals can be explained by differences in accumulated rainfall between both years. Under a wet spring, the winter wheat crop has responded better on higher elevations (and lower ECa) due to runoff, and worse in lower zones due to water saturation. The opposite relation was observed for 2019 under water shortage. Such dynamics are consistent with those reported by Kravchenko and Bullock (2000). The water shortage in 2019 suggested a higher degree of crop spatial heterogeneity, but results are inconclusive as differences can be associated to the crop species.

We highlight that MZ's temporal instability is an important issue for site-specific management as agronomic implications varied greatly with the $crop \times year$ setting. Under the same combinations of $Elevation \times Texture \times ECa$, both NDVI and crop yield patterns differed depending on seasonal rainfall. A 14% reduction of crop yield GVF was observed, from the first to the second $crop \times year$ setting, which indicates that site-specific management should not be conducted with persistent confidence in the absence of continuous evaluation and re-delineation of MZ's. Further research requires additional years of observations, necessary to dissociate geophysical and environmental effects from crop related ones, which might allow us to better interpret the temporal instability of crop spatial patterns.

The spatial interpretation of results was not fully automated from the metrics proposed because it also depended on agronomic knowledge. We believe that progress must be achieved through development of crop simulation models accessing within field spatial heterogeneity. The success of precision agriculture will benefit from improved protocols for MZ delineation and sampling schemes and further coordinated research is needed at variable spatial scales, empowering stronger synergism between researchers, farmers and sensing engineers.

References

Anselin, L. (1995). Local indicators of spatial association—LISA. *Geogr. Anal.*, 27(2):93–115.

Bivand, R. S. and Wong, D. W. S. (2018). Comparing implementations of global and local indicators of spatial association. *Test*, 27(3):716–748.

Cattani, C. E. V., Garcia, M. R., Mercante, E., Johann, J. A., Correa, M. M., and Oldoni, L. V. (2017). Spectral-temporal characterization of wheat cultivars through NDVI obtained by terrestrial sensors. *Rev. Bras. Eng. Agric. Ambient./Braz. J. Agric. Environ. Eng.*, 21(11):769–773.

Crawford, C. A. G., Bullock, D. G., Pierce, F. J., Stroup, W. W., Hergert, G. W., and Eskridge, K. M. (1997). Experimental design issues and statistical evaluation techniques for Site-Specific management. In *The State of Site-Specific Management for Agriculture*, pages 301–335. American Society of Agronomy, Crop Science Society of America, Soil Science Society of America, Madison, WI.

Goodwin, A. W., Lindsey, L. E., Harrison, S. K., and Paul, P. A. (2018). Estimating wheat yield with normalized difference vegetation index and fractional green canopy cover. *Crop, Forage & Turfgrass Management*, 4.

Han, J., Wei, C., Chen, Y., Liu, W., Song, P., Zhang, D., Wang, A., Song, X., Wang, X., and Huang, J. (2017). Mapping Above-Ground biomass of winter oilseed rape using high spatial resolution satellite data at parcel scale under waterlogging conditions. *Remote Sensing*, 9(3):238.

Johnson, C. K., Doran, J. W., Duke, H. R., Wienhold, B. J., Eskridge, K. M., and Shanahan, J. F. (2001). Field-scale electrical conductivity mapping for delineating soil condition. *Soil Sci. Soc. Am. J.*, 65(6):1829–1837.

Kalogirou, S. (2019). Spatial autocorrelation.

Kravchenko, A. N. and Bullock, D. G. (2000). Correlation of corn and soybean grain yield with topography and soil properties. *Agron. J.*, 92:75–83.

Leroux, C. and Tisseyre, B. (2019). How to measure and report within-field variability: a review of common indicators and their sensitivity. *Precis. Agric.*, 20(3):562–590.

Luo, X., Ma, C., Yue, Y., Hu, K., Li, Y., Duan, Z., Wu, M., Tu, J., Shen, J., Yi, B., and Fu, T. (2015). Unravelling the complex trait of harvest index in rapeseed (brassica napus l.) with association mapping. *BMC Genomics*, 16:379.

McCutcheon, M. C., Farahani, H. J., Stednick, J. D., Buchleiter, G. W., and Green, T. R. (2006). Effect of soil water on apparent soil electrical conductivity and texture relationships in a dryland field. *Biosystems Eng.*, 94(1):19–32.

Milne, A. E., Webster, R., Ginsburg, D., and Kindred, D. (2012). Spatial multivariate classification of an arable field into compact management zones based on past crop yields. *Comput. Electron. Agric.*, 80:17–30.

Moral, F. J., Terrón, J. M., and Silva, J. R. M. d. (2010). Delineation of management zones using mobile measurements of soil apparent electrical conductivity and multivariate geostatistical techniques. *Soil Tillage Res.*, 106(2):335–343.

Mulla, D. J. and Schepers, J. S. (1997). Key processes and properties for Site-Specific soil and crop management. In *The State of Site-Specific Management for Agriculture*, pages 1–18. American Society of Agronomy, Crop Science Society of America, Soil Science Society of America, Madison, WI.

Peeters, A., Zude, M., Käthner, J., Ünlü, M., Kanber, R., Hetzroni, A., Gebbers, R., and Ben-Gal, A. (2015). Getis–Ord's hot-and cold-spot statistics as a basis for multivariate spatial clustering of orchard tree data. *Comput. Electron. Agric.*, 111:140–150.

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.*, 20:53–65.

Schemberger, E. E., Fontana, F. S., Johann, J. A., and Souza, E. G. D. (2017). Data mining for the assessment of management areas in precision agriculture. *Engenharia Agrícola*, 37(1):185–193.

Scudiero, E., Teatini, P., Corwin, D. L., Ferro, N. D., Simonetti, G., and Morari, F. (2014). Spatiotemporal response of maize yield to edaphic and meteorological conditions in a saline farmland. *Agron. J.*, 106:2163–2174.

Scudiero, E., Teatini, P., Manoli, G., Braga, F., Skaggs, T. H., and Morari, F. (2018). Workflow to establish Time-Specific zones in precision agriculture by spatiotemporal integration of plant and soil sensing data. *Agronomy*, 8(11):253.

Tibshirani, R., Walther, G., and Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *J. R. Stat. Soc. Series B Stat. Methodol.*, 63(2):411–423.

Wang, L. and Liu, H. (2006). An efficient method for identifying and filling surface depressions in digital elevation models for hydrologic analysis and modelling. *Int. J. Geogr. Inf. Sci.*, 20(2):193–213.

A Appendix

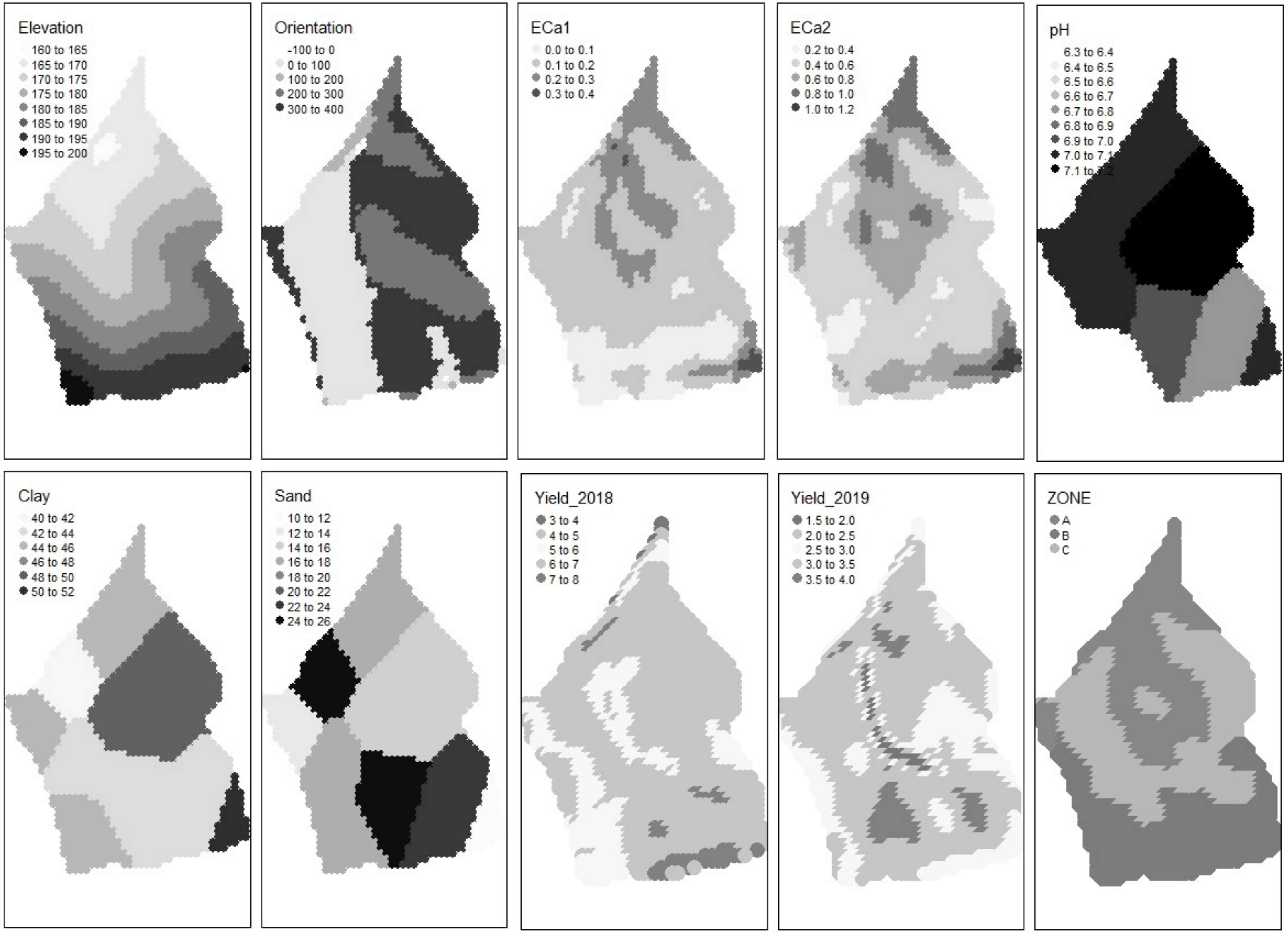

Figure 1: Visual outcomes: Maps of geomorphological properties (Elevation, Orientation, superficial ECa(1) and deep ECa(2), respectively expressed in m, degrees and dS/m), maps of pH and texture (clay and sand content, both expressed in %), simulated yield maps (expressed in Mg/ha), and the map of delineated Management Zones (MZ).