Generation and quantitative evaluation of interpretable maps in precision agriculture

Patrice Loisel, Brigitte Charnomordic, Hazaël Jones, Bruno Tisseyre

Generation and quantitative evaluation of interpretable maps in precision agriculture

Patrice Loisel*[1,] Brigitte Charnomordic[1], Hazaël Jones[2], Bruno Tisseyre[2].

1. INRA, UMR MISTEA, Montpellier, France
2. UMR ITAP, IRSTEA/SupAgro, Montpellier, France
Corresponding author: Patrice.loisel@inra.fr

ABSTRACT

In this paper, an algorithm for the generation of interpretable maps yielding management zones in agriculture is presented. The key points are the definition of numerical zoning quality criteria that are used to guide a global optimization procedure based on quantile values. This makes it an alternative to classical zoning procedures based on local criteria, such as segmentation-based ones. The algorithm is applied to real yield data from precision agriculture.

Keywords: Management zones; Precision agriculture; Zoning; Optimization; Visualization; France

1. INTRODUCTION

Delineation of management zones is of paramount importance for many applications (Urretavizcaya et al., 2014) in precision agriculture. A management zone is a spatially contiguous area to which a particular treatment may be applied. A management class, which can be assigned a label, is the area over which a particular treatment may be applied. This may constitute more than one zone.

The automatic generation of interpretable maps (called zonings in the following) is a challenge. There are two main types of methods to generate zones from spatial data: classification based on a priori classes (Taylor et al., 2007) and segmentation algorithms (Pedroso et al., 2010), based on Union-Find procedures with initial clusters centered on seeds. However classification algorithms do not consider spatial constraints to define contiguous classes, while the sensitivity to the initial choice of seeds and to the order of merging operations are well known drawbacks of segmentation algorithms.

The aims of this paper are i) to propose a numerical criterion to evaluate zoning quality ii) to present an optimization-based approach to generate interpretable maps based on this criterion. The proposed approach avoids the definition of a priori classes, while assigning interpretable labels to the generated zones.

2. THE QUALITY CRITERION

The criterion is designed in order to make the resulting zoning easy to interpret. The aim of the quality criterion proposed hereafter is two-fold: i) quantify how zones are heterogeneous on the whole field under study, which is of importance for zone delineation algorithms aiming at maximizing inter-zone variability, ii) quantify how neighboring zones are similar, which is of importance considering most segmentation algorithms involve a fusion step aiming at merging similar contiguous zones. The novelty of the proposed approach is to consider these two aspects simultaneously. The key idea is to maximize the contrasts in all parts of the map. Consider n georeferenced data points $(s_i, F(s_i)), i = 1 \dots n$ where $s_i = (\mathrm{x}, \mathrm{y})$ represents the Cartesian coordinates. The attribute $F(s_i)$ is a one-dimensional numerical value. A zoning is composed of p zones, and denoted $Z = (z_I)_{I \in \{1,2 \dots p\}}$.

2.1 Heterogeneity between neighbouring zones z_I, z_J

Zones are neighbours if they share a common edge.

$M_{IJ} = \frac{\sum_{s_i \in z_I} \sum_{s_j \in z_J} \left(F(s_i) - F(s_j)\right)^2 A(s_i) A(s_j)}{A_I A_J}$, where $A(s_i)$ is the area of $P(s_i)$, A_I is the area of zone z_I.

M_{IJ} can be decomposed into three parts: $M_{IJ} = \sigma_I^2 + \sigma_J^2 + d_{IJ}^2$, σ_I^2 (resp. σ_J^2) being the variance of F in zone z_I (resp. z_J), d_{IJ} is the difference of mean values of F in zones z_I and z_J.

2.2 Heterogeneity within zone z_I

The same formulae are used to define M_{II} and decompose it, $M_{II} = 2\sigma_I^2$.

<u>Contrast indicator C_{IJ}:</u> ratio of the inter zone heterogeneity M_{IJ} and of the average intra zone heterogeneity $\frac{M_{II} + M_{JJ}}{2}$, for any two neighbouring zones z_I, z_J. $C_{IJ} = \frac{2M_{IJ}}{M_{II} + M_{JJ}} = 1 + \frac{d_{IJ}^2}{\sigma_I^2 + \sigma_J^2}$

2.3 Criterion

$Crit(\mathrm{Z}) = min_{z_I \in Z,\, z_J \in V(z_I)} C_{IJ}$, where $V(z_I)$ is the neighbourhood of z_I (composed of all neighbouring zones of z_I). The choice of the minimum value over all pairs of neighbouring zones avoids compensations between pairs, as it would be the case if average were used instead. It ensures good contrast among all zones. The higher the criterion value, the better the zoning quality.

3. OPTIMIZATION-BASED ZONING

The objective of this part is to propose a zoning algorithm aiming at demonstrating the interest of the criterion. The approach, shown in Figure 1, takes as starting point a zoning defined by a combination of contour lines. Contour lines based on attribute data quantiles delineate zones that are easy to label, all values inside the zones being either lower or higher than the cut point.

The goal is to find the best combination of contour lines. This is an advantage in comparison of region growing/merging algorithms, which do not guarantee a clear labelling. The algorithm consists of several embedded steps.

The correction procedure considers the area (size) of the zones resulting from the contour line definition. Assuming that small zones may correspond to unmanageable regions, the correction procedure consists in modifying the initial zoning in order to respect the size constraint, either by removing or enlarging small zones.

The procedure result consists in a subset of best zonings for a given number of labels.

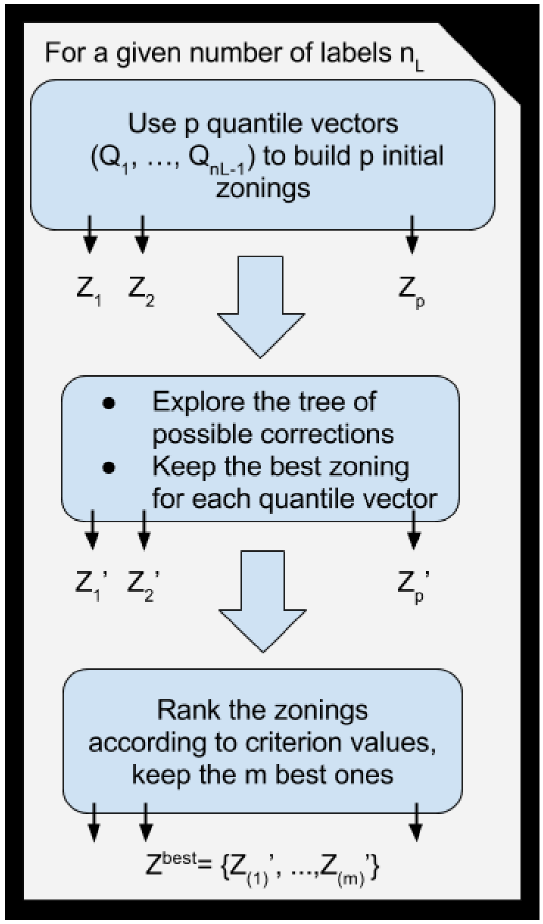

Figure 1. Zoning procedure.

4. CASE STUDY

The cereal field is located in Peterborough, England and was harvested with a CLAAS harvester with a 6m cutting bar. It contains (1000 points/ha) yield values automatically geolocated by the harvester.

The Yield variable ranges between 0.47 and 7.67. Figure 2 displays the best zoning obtained with 3 labels, corresponding to data quantile-associated probabilities (0.2, 0.7). The mean value within zone values is indicated on the map..

The optimal zoning leads to 6 well differentiated zones. The criterion value is 5.07.

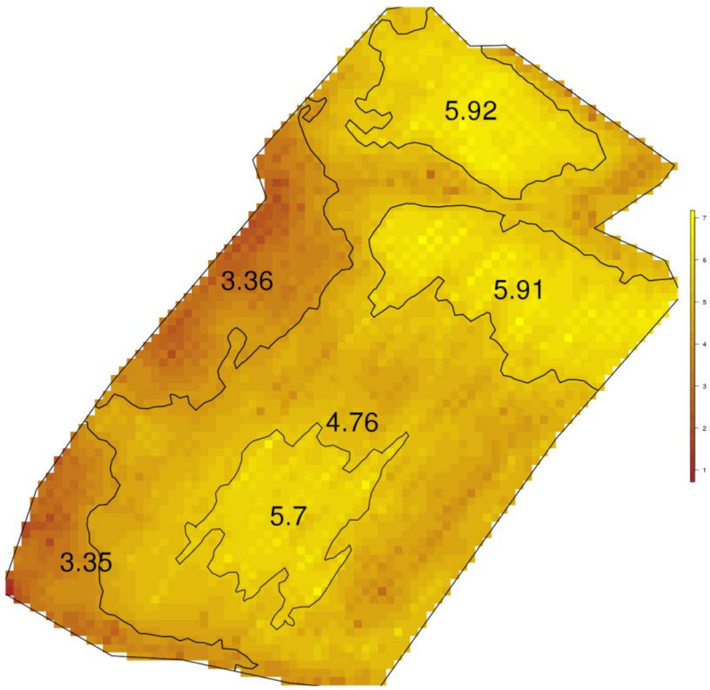

Figure 2. Best zoning of yield data field for 3 labels.

5. CONCLUSION

In this paper, we propose criteria to evaluate the quality of any zoning and to design efficient and interpretable within-field zonings. The approach allows fast and objective evaluation of various zonings, required when large quantities of data are to be analyzed.

6. REFERENCES

Pedroso, M, Taylor, J, Tisseyre, B, Charnomordic, B, & Guillaume, S 2010 ‘A segmentation algorithm for the delineation of agricultural management zones’ Computers and Electronics in Agriculture, vol. 70, no. 1, pp. 199-208.

Taylor, JA, McBratney, AB, & Whelan, BM 2007 ‘Establishing management classes for broadacre grain production’ Agronomy Journal, vol. 99, pp. 1366-1376.

Urretavizcaya, I, Santesteban, L G, Tisseyre, B, Guillaume, S, Miranda, C, & Royo, J B 2014 ‘Oenological significance of vineyard management zones delineated using early grape sampling’ Precision Agriculture, vol. 15, no. 1, pp. 111-129.