

Segmentation of land use maps for precision agriculture

Wilian F. Costa¹, Michel J.M. Bieleveld¹,
Raphael G. Chinchilla¹, Antonio M. Saraiva¹

¹Escola Politécnica da USP - Av. Prof. Luciano Gualberto, travessa 3 no 380
CEP 05508-010 - São Paulo - SP, Brazil

wilianfc@gmail.com

Abstract. *We present a multilayer image segmenter adapted to be used for Precision Agriculture (PA). PA depends strongly on the application of information technologies to divide and group geographic areas based on land use, soil data, metrological data and agricultural resources for planning and implementation of activities to increase output by using optimal strategies for each segment. We implemented a modified Baatz algorithm in the statistical language R and speed sensitive code was implemented in C++. The code will be made publicly available under the GNU Lesser Public License. We show the merit of our approach at the hand of a landscape and discuss the obtained segments generated by our tool.*

1. Introduction

PA allows for the definition of optimal agricultural management with the use of decision support systems. This leads to: effective management of resources to increase farm profitability, responding to crop inter and intra field variability, and the reduction of environmental impact [Zhang et al. 2002]. An essential step in managing land use, and keeping the complexity manageable, is to divide land into smaller homogeneous regions with similar production factors [Saraiva 2003]. Unsupervised segmentation algorithms are used to generate these divisions. Typically these algorithms find there use in the segmentation of images to automatically extract all objects and regions of interest for a certain task. In this work one such algorithm, Baatz [Baatz and Schape 2000], is implemented and applied to geospatial multi-scale and multi-spectral production data segmentation. The resulting segments can then be used as part of a larger decision system to divide the land into manageable areas with similar characteristics, such as soil, vegetation and climate factors, that otherwise would be difficult to separate visually [Saraiva 2003, Costa 2016].

2. Materials and Methods

The segmentation algorithm was implemented in R, a computer language and environment used for statistical computing and popular in the scientific community [R Core Team 2014]. The R environment runs on all popular platforms and is extensible through new software modules called packages. Many such packages are readily available that deal with geospatial data on the CRAN Repository¹. For example, the *Raster* package that provides functions to load and save Geographic Information System (GIS) maps and data.

¹<https://cran.r-project.org>

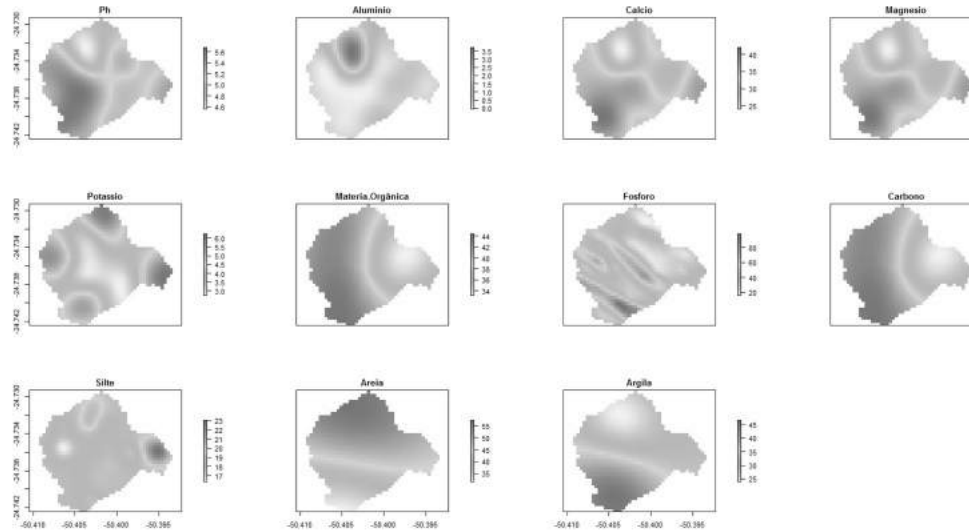


Figure 1. Input raster stack containing the environmental factors

The R language is an interpreted language and while many packages are solely written in R, many compute intensive packages also make use of C/C++ code. In our implementation we made use of the Rcpp API [Eddelbuettel and François 2011] to access, extend and modify seamlessly R objects using C++ for those function that require more time to execute than other parts of the algorithm.

3. Results

In Figure 1 typical land use data is shown and composed of eleven layers. Each map within the figure represents the distribution of a particular soil factor. The first layer represents the acidity of the soil, the following seven layer the concentration of several minerals, and the last three represent land composition information. These maps are generated by using agricultural data obtained and discussed in earlier work [Costa et al. 2013].

Figure 2 illustrates that our implementation separates the input data into meaningful segments, thus defining areas for optimal agricultural management. The figure shows the obtained segments in a two dimensional space defined by the two most dominant components obtained with Principle Component Analysis (PCA) after normalization. The loadings indicate a clear separation of the land in clay, sand and silt regions. We also see there is a region clearly defined with high potassium. The obtained segments are grouped by variables that without using a complex algorithm, such as the one implemented in this work, would be difficult to distinguish, for example the two clusters in the center top.

Figure 3 shows the result of the application of the algorithm on the input data and segmented according to the separation just described. The first map at the top left shows the segmented land area considering all the soil factor, whereas the remaining maps show how each layer by itself is segmented.

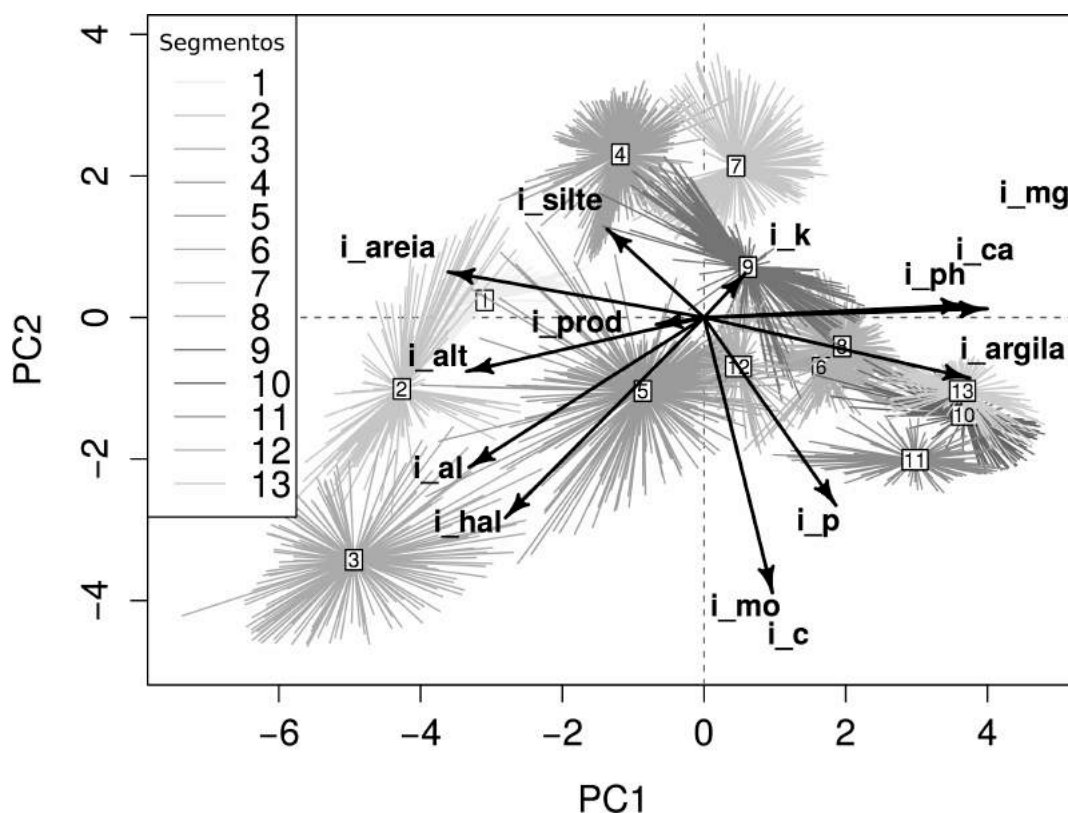


Figure 2. Correlations between the input data mapped onto first two principal components of PCA. The vectors represent loadings of the variables. The centroids of each group is identified and represents a segment.

4. Conclusion

Our adaptation and implementation of the Baatz algorithm made available through an R package is shown to be effective for the segmentation of geospatial data [Costa 2016]. Our test data has shown that the implementation is able to divide the virtual landscape defined by complex factors in regions with homogeneous characteristics which has its applicability in the effective management of agricultural land use.

Acknowledgement

The authors are grateful to CAPES and NAP-Biocomp - University of São Paulo. The funders had no role in study design, data analysis, decision to publish, or preparation of the manuscript.

References

- Baatz, M. and Schape, A. (2000). Multiresolution segmentation: An optimization approach for high quality multi-scale image segmentation. In J. Strobl, T. B. and Griesbner, G., editors, *Angewandte Geographische Informations-Verarbeitung*, pages 12–23. Wichmann Verlag, Karlsruhe, Germany, xii edition.
- Costa, W. F. (2016). *Segmentação multiresolução variográfica ótima*. PhD thesis, Universidade de São Paulo, São Paulo, Brazil.

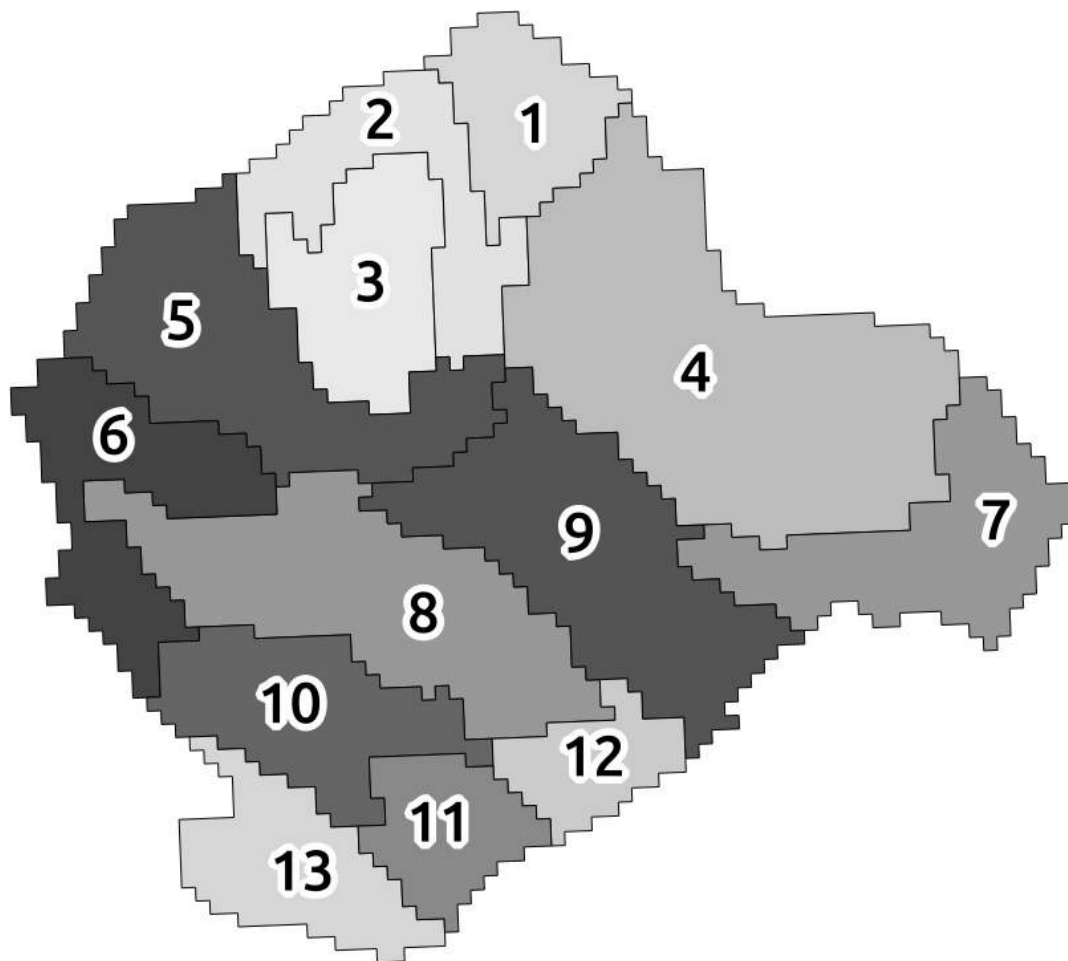


Figure 3. The result of the segmenter, the first image depicting the overall segmented result, the numbers indicating the centroids of Figure 2

- Costa, W. F., Saraiva, A. M., Veiga, A. K., Soares, F., and Quintanilha, J. A. (2013). Applying multiresolution segmentation algorithm to generate crop management zones based on interpolated layers.
- Eddelbuettel, D. and François, R. (2011). Rcpp: Seamless R and C++ integration. *Journal of Statistical Software*, 40(8):1–18.
- R Core Team (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Saraiva, A. M. (2003). Tecnologia da informação na agricultura de precisão e biodiversidade: Estudos e proposta de utilização de web services para desenvolvimento e integração de sistemas. Tese de livre docência, Universidade de São Paulo.
- Zhang, N., Wang, M., and Wang, N. (2002). Precision agriculture - a worldwide overview. *Computers and Electronics in Agriculture*, 36(2 - 3):113 – 132.