2013 Conference
June 23-27, 2013
TORINO, ITALY
EFITA WCCA CIGR
Sustainable Agriculture through ICT innovation

Applying Multiresolution Segmentation Algorithm to Generate Crop Management Zones based on Interpolated Layers

Wilian França Costa [1], Antonio Mauro Saraiva [1], Allan Koch Veiga [1], Fabiana Soares Santana [3] José Alberto Quintanilha [2]

[1]Department of Computer & Electrical Engineering, Escola Politécnica Universidade de São Paulo. Av. Prof. Luciano Gualberto, travessa 3, 158 – Cid. Universitária São Paulo - SP - CEP:05508-010, Brazil

[2] Department of Transportation Engineering, Escola Politécnica Universidade de São Paulo. Av. Prof. Almeida Prado, travessa 2, 83 – Cid. Universitária São Paulo - SP - CEP:05508-900, Brazil

[3] Center of Mathematics, Computing and Cognition, Universidade Federal do ABC Rua Santa Adélia, 166. Bairro Bangu, Santo André SP- CEP 09.210-170, Brazil

wilianfc@gmail.com, fabiana.santana@gmail.com, allankv@usp.br, saraiva@usp.br, jaquinta@usp.br

ABSTRACT

Agricultural Crop Field Modelling is a technique that uses georeferenced data points of soil information analysis or spectral remote sensing imagery raster layers to obtain models to: identify spatial differences in the regions; maximize the production, minimize the nutrient reposition; and find the desired characteristics in terrain based on class of production selected. Concepts related to Object Based Image Analysis were adopted to generate a meaningful Management Zone delineation aim the construction of models based on a multi source georeferenced data fusion, building a concrete relationship among different input of information layers employed using the Multiresolution region-based segmentation algorithm.

Keywords: Obia, crop management zone, segmentation, multiresolution, brazil

1. INTRODUCTION

Agricultural Crop Field Modelling is a technique that uses georeferenced data points of soil information analysis or spectral remote sensing imagery raster layers to obtain models to: identify spatial differences in the regions; maximize the production, minimize the nutrient reposition; and find the desired characteristics in terrain based on class of production selected (Jones et al., 2010). Concepts related to Object Based

Image Analysis - OBIA (Blaschke, 2010) were adopted to generate a meaningful Management Zone - MZ delineation.

2. MATERIALS AND METHODS

OBIA is a sub-discipline of Geographic Information Sciences (Blaschke, 2010) developed to partitioning remote sensing imagery into meaningful image-objects through spatial, spectral and temporal scale (Hay and Castilla, 2006). This requires a efficient image segmentation, attribution, classification, query and link to individual objects in space and time. The strength of OBIA is the ability to deal with multi-scale data, enabling your application to environmental monitoring, modelling and management. OBIA main advantages are (Hay and Castilla, 2006): reduction of computational load due to dimension reduction, since the basic units are image-objects; a more comprehensible way to analyze the region since it is partitioned into humans conceptually organized objects (e.g. more productivity regions); possibility to apply complex statistical techniques (e.g.: non-parametric); possibility to explore interesting relations through OBIA features (shape, texture and object context); easier integration with vector GIS and spatial databases. The concepts of OBIA are related to multi-spectral images but the meaning of geospatial objects can be applied with minor adaptation to be used with geoprocessed interpolated layers maps. These layers are usually generated using some interpolation algorithm (Pebesma et al., 2011), such as Kriging, Spline or others for data from georeferenced sample points that have chemical or physical crop field information. The objective of geospatial multi-layer generated data type is to establish an n-dimensional space where each dimension represents one attribute for generated geo-objects. Virtually, we can use any type of georeferenced information that can be quantified and have some quantity of observations sampled over the place under analysis.

2.1 Segmentation

The basic activity for object oriented modeling is successful map segmentation (i. e. the process to produce a set of non-overlapping segments/polygons that is expected to have relatively homogeneous and semantically significant groups of pixels (Blaschke, 2010). The main objective is the construction of models based on a multi source geo-referenced data fusion, building a concrete relationship among different layers employed (Baatz et al., 2000).
Considering crop productivity MZ as a set of homogeneous conditions to support the desired culture behavior, the identification and segmentation of this areas is a key factor support further inferences and the profiler analysis step. To achieve this objective, the Multiresolution region-based segmentation algorithm (Baatz et al., 2000) was chosen because it deals with multiple information layers, seeking to aggregate regions by minimizing their spatial and spectral heterogeneity (in our case, interpolated surface values).

This algorithm is based in a iterative process of local optimization that minimizes the mean heterogeneity inside each segment. The heterogeneity definition adopted has a component spatial and a spectral (set of values on each pixel position) and is defined for the equation 1 where f represents the fusion factor among segments (objects), h_{spc}, the calculated spectral factor, h_{form}, the object form factor and the weight factor w_{spc} defines the relative importance between those factors.

$$f = w_{spc}h_{spc} + (1 - w_{spc})h_{form} \tag{1}$$

The spectral component h_{spc} is defined as the associated value to each pixel that belongs to one segment. This measurement is proportional to weighted mean of the standard deviation of each layer or dimension. It is defined in equation 2 where s_1 is the selected segment, s_2 is the neighbour under analysis and s_3 is the union resulting segment; c is the band (for remote sensing) or map interpolated layer index; w_c is the weight defined for each band/layer; $\sigma_c^{s_i}$ is the standard deviation on band/layer c; and n is the number of pixels for each s_i, for i={1, 2, 3} (Happ et al., 2009).

$$h_{esp} = \sum_c w_c(n_{s_3}\sigma_c^{s_3}(n_{s_1}\sigma_c^{s_1} - n_{s_2}\sigma_c^{s_2})) \tag{2}$$

The spatial component h_{form} (equation 3) is composed by two form attributes: compactness and smoothness, where the relative importance between those factors is adjusted by w_{comp} weight value.

$$h_{form} = w_{comp}h_{comp} + (1 - w_{comp})h_{smo} \tag{3}$$

Compactness degree is defined as a ratio between segment perimeter and the squared root of your area in pixels units times the number of object pixels. Smoothness is defined as the ratio between segment perimeter and your minimum bounding box.
Initially each segment (object) is defined as a single image pixel. The segment grow as they are merged with their neighbours always trying to minimize the incorporated heterogeneity and only if the minimum fusion factor f is less than an value e defined as a squared for a scale parameter. The process continues until anyone segment could not grows more. More details about the Multiresolution algorithm can be seen in (Baatz et al., 2000) and (Happ et al., 2009). One of the most employed implementations is found in eCognition software [http://www.ecognition.com].

2.2 Classification

To assign meanings to obtained objects, is necessary to analyze each one internal characteristics and apply some technique to classify and separate classes aiming the identification of important aspects that can give us some ”glue” about what is happening, and support to further decisions about the best way to deal with each MZ selected. Considering the data-mining and machine learning point of view, we can use supervised (e.g.:Principal Component Analysis, Support Vector Machines), and unsupervised (e.g.: Kmeans, Maximum Likelihood) algorithms to classify the data geo-objects (Navulur, 2006). For this work, we mixed both techniques in a semi-supervised approach. In our case we select objects according the mean productivity obtained from harvest machine data.

The experiments were conducted applying the Interimage [www.lvc.ele.puc-rio.br/projects/interimage] version of Multiresolution. As the implementation of Multiresolution Baatz algorithm found on Interimage software just extract objects based in remote sensing image types (multi-band image formats), we implemented a set of R scripts using some of existing packages (e.g.: rgdal, sp) to convert ESRI shapefiles with crop field point information to interpolated raster images that was inserted into a multi-band tiff image.

The profiler applied in this study uses non-parametric statistics based on the Akaike information criterion to measure the relative goodness of fit for samples histogram to some previous defined distribution (normal and weibull). With this technique, the maximum productivity probability and region profile can be indicated, helping the farmer to get best decisions about fertilizers and crop MZ selection.

2.3 Model Fitting and Inferences

For each object attribute, we can fit a statistical distribution (Gaussian, Weibull) based on your data histogram. These facts enable us to choose the best distribution that can represent our class of objects data. The distribution fitted for each layer can be used to infer probability of a desired behavior to happen. The model selection for distribution can be done thought tree steps: 1. Choose a set of models to be fitted; 2. Search for best parameter adjustment for each model using Maximum Likelihood Estimation - MLE; 3. Selection for best model on set using the Aikane Information Criterium Index - AIC. How lower is the AIC value for model, better is the distribution fitting.

3. RESULTS

To exemplify the proposed approach, testes were performed using data fromba private farm crop field production of wheat (planted on set. 2008) on the State of Paraná, Brazil. A total of 67 samples were collected on mar. 2009 after the harvest over an area of 1.6Km x 1.2Km (approx.) to estimate the levels of pH, H+AL($mmol_c$/dm3), Al($mmol_c$/dm3), Ca($mmol_c$/dm3), Mg($mmol_c$/dm3), K($mmol_c$/dm3), Base Saturation(%), CTC(mmol$_c$/dm3), P(mg/dm3), C(g/dm3), Organic Matter(g/dm3), V(%), Silt(%), Sand(%) and Clay(%). The distribution of sampling points is shown in Fig. 1.

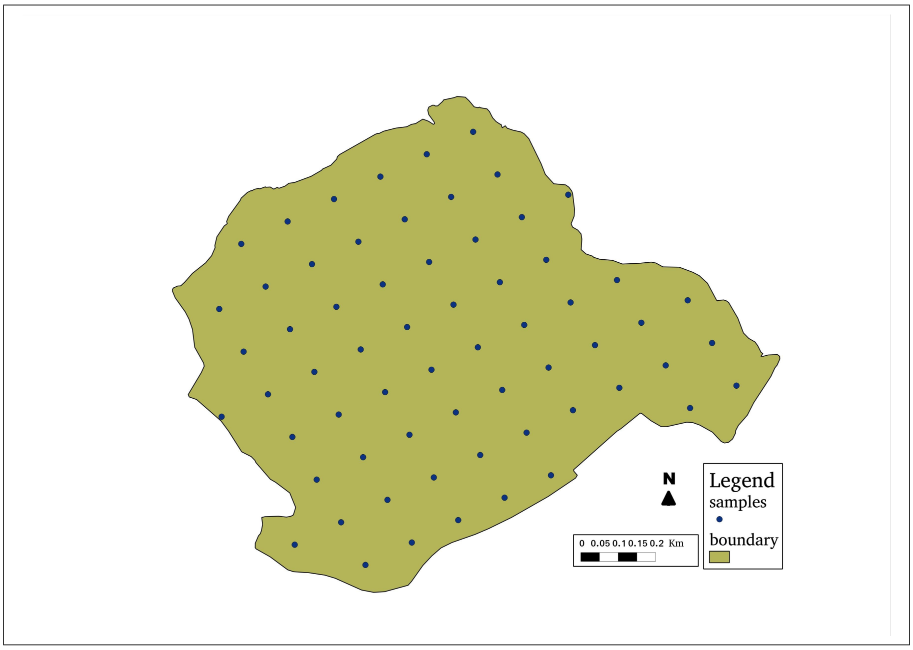


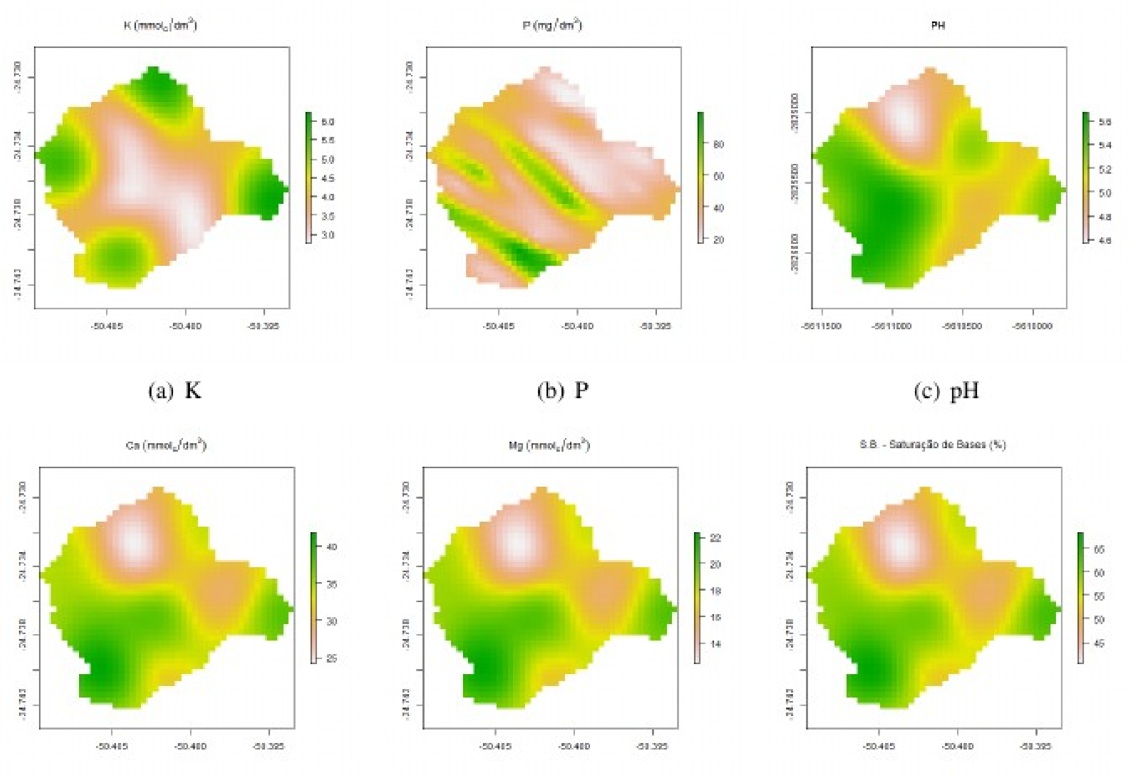
Figura 1: Sample points

For the segmentation step, the values of pH, Ca, Mg, K, BS, CTC, P, Silt, Sandy and Clay were interpolated using Kriging algorithm from R Intamap packages [www.intamap.org] (Fig. 2). We can clearly see in the Fig.2 that layers Ca, Mg and base Saturation have high level of correlation and exposing a spatial pattern. For this aproach this is an important behavior that is captured by Multiresolution to generate a meaningful zone area based on spatial homogeneity related to information provided.
After that, we perform a data normalization of values to [0, 1] interval. The resulting of normalization were used in Multiresolution algorithm as input. The generated segments are shown in Fig. 3 in which the where colored according productivity mean values (red: low, yellow:medium, green: high).
In the next step, the generated geo-objects is evaluated in respect of their attributes and a MLE test was performed for each layer, through a negative log-likelihood function that returns a probability value, representing the similarity among geo-objects with respect to best productivities(or another hypothesis under evaluation).

4. CONCLUSIONS

This is a work in progress project and the interference still under development and will be published as soon the authors finish the proposal of method to evaluate the effectiveness of geo-objecs and model related with crop field productivity and maps.

C0214
Wilian F., Costa; Antonio M. Saraiva; Allan K. Veiga, Fabiana S. Santana; José A. Quintanilha. “Applying Multiresolution Segmentation Algorithm to Generate Crop Management Zones based on Interpolated Layers”. EFITA-WCCA-CIGR Conference “Sustainable Agriculture through ICT Innovation”, Turin, Italy, 24-27 June 2013.



(a) K (b) P (c) pH

(d) Ca (e) Mg (f) Base Saturation

Figure 2: Generated Maps of some of interpolated layers using Kriging method from Intamap R package

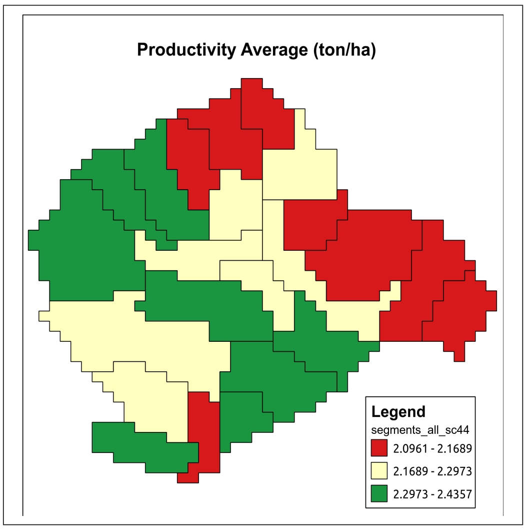


Figure 3: Generated segments

C0214
Wilian F., Costa; Antonio M. Saraiva; Allan K. Veiga, Fabiana S. Santana; José A. Quintanilha. "Applying Multiresolution Segmentation Algorithm to Generate Crop Management Zones based on Interpolated Layers". EFITA-WCCA-CIGR Conference "Sustainable Agriculture through ICT Innovation", Turin, Italy, 24-27 June 2013.

The experiments were conducted applying the following open source softwares: Interimage [www.lvc.ele.puc-rio.br/projects/interimage/], PostGIS, R [cran.r-project.org] and 52North OGC-WPS [52north.org/wps] implementation.

5. AKNOWLEDGEMENTS

The autors is grateful to Financiadora de Estudos e Projetos do Ministério da Ciência e Tecnologia, FINEP/MCT, for funding the PROSENSAP, de n° 01.08.0566.00, in which this research is part and also too the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES.

6. REFERENCES

Baatz, M., A. Schäpe and J Strobl. 2000. Multiresolution Segmentation: an optimization approach for high quality multi-scale image segmentation. In: *Angewandte Geographische Informationsverarbeitung XII. Beiträge zum AGIT-Symposium Salzburg 2000*, Karlsruhe, Herbert Wichmann Verlag. 12–23.

Blaschke, T. 2010. Object based image analysis for remote sensing. *IS-PRS Journal of Photogrammetry and Remote Sensing* 65 (1): 2–16.

Happ, P. N., R. da S. Ferreira, C. Bentes, G. A. O. da Costa, and R. Q. Feitosa. 2009. Segmentação multiresolução:uma abordagem paralela para segmentação de imagens de alta resolução em arquiteturas de múltiplos núcleos. In: Anais *XIV Simpósio Brasileiro de Sensoriamento Remoto*. INPE, Natal, Brasil,6935–6942.

Hay, G. J., and G. Castilla. 2006. Object-based image analysis: Strengths, weaknesses, opportunities and threats (SWOT). In *Proceedings of 1st international conference on object-based image analysis (OBIA 2006)*. Salzburg University, Austria.

Jones, G., C. Gee, S. Villette, and F. Truchetet. 2010. A crop field modeling to simulate agronomic images. In *2010 4th International Symposium on Communications, Control and Signal Processing (ISCCSP)*, 1–4. doi:10.1109/ISCCSP.2010.5463310.

Navulur, Kumar. 2006. *Multispectral Image Analysis Using the Object-Oriented Paradigm*.CRC Press.

Pebesma, E., D. Cornford, G. Dubois, G. Heuvelink, D. Hristopulos, J. Pilz, U. Stöhlker, G. Morin, and J. O. Skrien. 2011. INTAMAP: The design and implementation of an interoperable automated interpolation web service. Computers & Geosciences 37 (3):343–352.
