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

A Reference Process for Management Zone Delineation

Raul Teruel dos Santos[1,3], José Paulo Molin[2], Antonio Mauro Saraiva[3]
[1]Universidade Federal de Mato Grosso – Instituto de Computação
[2]Universidade de São Paulo – Escola Superior de Agricultura "Luiz de Queiroz"
[3]Universidade de São Paulo – Escola Politécnica
raul@ic.ufmt.br, jpmolin@usp.br, saraiva@usp.br

ABSTRACT

A reference model is proposed for management zone delineation. The results showed that although there are differences in the steps used by each author, some are recurrent. Yet, steps not considered can have a strong impact on the maps obtained. This study is also useful as a reference document for standardization in management zone delineation, and can be used as a guide to choose the right tools, right data and methods, allowing the optimization of the management zone delineation and improving the quality of the maps obtained. The use of Business Process Model and Notation (BPMN) showed to be adequate to represent the steps, their flows and the decisions to be made along the process.

Keywords: Management zone, Reference model, Process, Spatial variability, Clustering, Business process model and notation, Brazil.

1. INTRODUCTION

The main objective of precision agriculture is to enable field management considering spatio-temporal variability. For many years, soil management in the field was performed by the average of soil nutrients required, but this method results in insufficient nutrients application in some regions and excessive amounts in others, wasting them and, consequently, increasing production costs without maximizing productivity.

One way of dealing with this problem in the context of precision agriculture is to use management zones, which means dividing the area into sub-areas that present homogeneity of the need for input application.

The proper delineation of management zones, through the study of the relationships between yield, soil properties, relief and others, prevents chemical exhaustion and degradation of the soil physical attributes, seeking the maximum sustainable productivity and profitability. Several techniques have been used for delineating management zones, each of which is suitable for different types of data gathered. However, before using a technique for delineating management zones, many steps are required to prepare and to analyze data and to evaluate the management zones.

R. T. Santos, J. P. Molin, A. M. Saraiva. "A Reference Process for Management Zones Delineation". EFITA-WCCA-CIGR Conference "Sustainable Agriculture through ICT Innovation", Turin, Italy, 24-27 June 2013.

Although there is a diversity of tools, the correct selection of methods, data and tools is a complex task due to the large number of parameters and data interrelated and their characteristics.

A process reference model represents many dynamic aspects of a business organization, such as sequence of activities, flow control between activities and restrictions to particular premises (Becker *et al.*, 2003) *apud* (Fettke *et al.*, 2006). Our aim is to propose a reference model for management zone delineation.

2. MATERIAL AND METHODS

To propose this model, many papers about management zone delineation were analyzed for checking different steps in the process, data, methods, algorithms, tools, equipment and software used in studies. Many steps were identified, from data collection to generation and evaluation of management zone maps. However, there are different numbers of steps shown by researchers. Our reference model process was divided into some steps observed in most of the works analyzed: data collection, data selection, data filtering, clustering and map evaluation.

We used Business Process Model and Notation (BPMN) version 2.0 to present a general description of the steps, flows, methods, data, tasks, events and conditions used in the management zone delineation process. The BPMN is a standard proposed by the Object Management Group (OMG) for business process modeling. The main aim of BPMN is to work as a standard bridge filling the gap between the business process and its implementation. It provides a unique notation to business users engaged in different phases of the business process life cycle: project, implementation, management, monitoring and analysis (OMG, 2011).

We used Bonita Studio software for designing the reference model process for management zone delineation. The software used is free and provides many resources, including java call functions, interface design and other functionalities. It works with BPMN 2.0.

3. RESULTS

The reference model process for management zones delineation composed of five sub processes is presented in Figure 1. The steps are described in the next subsections.

Figure 1. Reference model process for management zone (MZ) delineation modelled in BPMN.

3.1 Data Collection

The data collection step is the beginning of the process for management zones delineation. In this step some tasks are needed for right data collection. The step begins with data types selection, followed by sampling methods, definition of the sampling strategy and then it ends with de data collected (Figure 2).

The sampling plan must be set in order to obtain the lowest density sampling that best characterizes the spatial variability in a field. Particular attention is necessary for soil data samples to determine an appropriate sampling scheme which normally contains limited density points due to the high cost of soil analysis, but should contain enough points to adequately characterize field spatial variability (Kerry *et al.*, 2010). Soil variability may result in different effects on yield (Zhang *et al.*, 2011).

According to Kerry *et al.* (2010), different nutrients require different sampling schemes; thus, the number of samples must be based on the nutrients that mostly affect crop growth. The geostatistic variogram can be used to provide knowledge about the spatial variation between samples (Kerry *et al.*, 2010) helping optimize the sampling strategy.

Data collection may be performed in the field as defined in the sampling plan. Soil electric conductivity may be considered as it enables a large sampling density. Another alternative is the use of images, due to their capacity of covering large areas with reduced cost. Therefore, researchers have used images for analyzing the need of nutrients in crops based on canopy reflectance analysis (Zhang *et al.*, 2011). In this context, the types of data to be used depend on the target of the collection, technology and financial resources available.

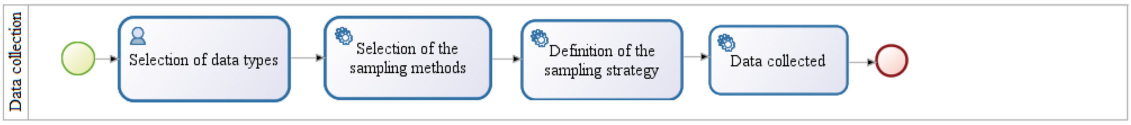

Figure 2. Reference process for data collection modelled in BPMN.

3.2 Data Filtering

The data filtering step aims to correct errors occurred during the data collection process. In this step, precision agriculture specialists choose the filters they will apply on the data collected and then the set of points resulting from each filter are intersected to obtain a filtered map (Figure 3).

Some errors found are: outside field points, GPS positioning error, outliers, incorrect or missing data due to calibration and operation errors of yield monitors, for example. Those errors can drastically reduce the quality of the management zone maps and the magnitude of this reduction depends on the methods used in the next steps. Some methods are more sensitive to errors than others. Due to the negative impact of the errors, many methods in the literature are found to remove them, from simple removal of points outside the range based on mean and standard deviation until methodologies with many and complex steps.

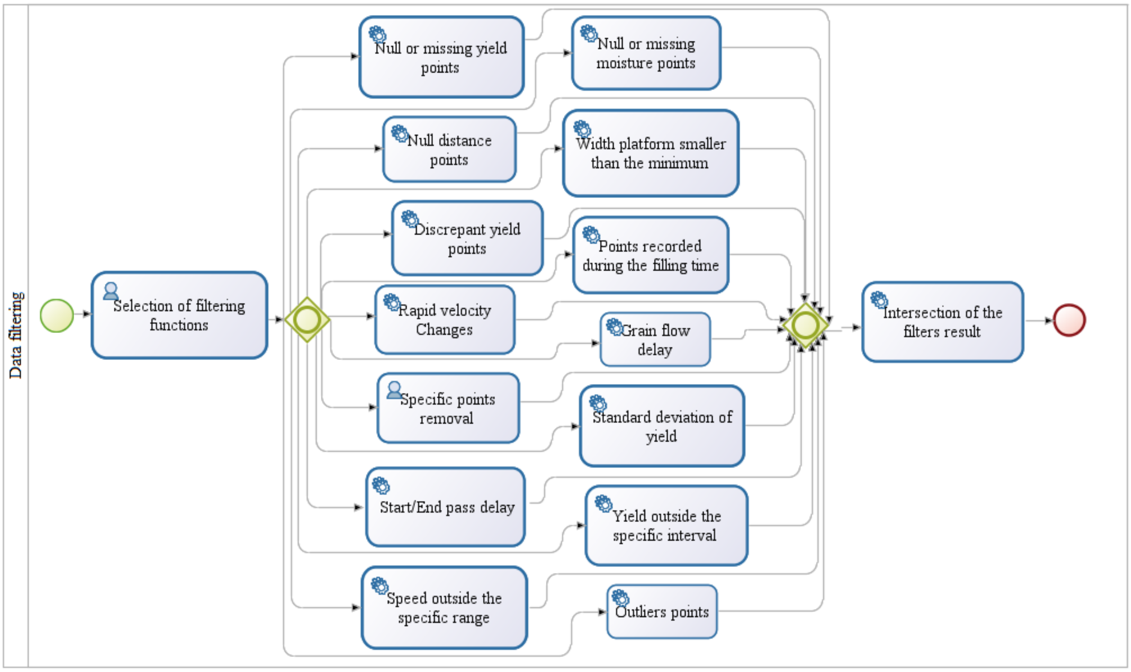

Figure 3. Reference process for data filtering modelled in BPMN.

In our model, we proposed the use of the following filters: outliers, null or missing moisture points, platform width smaller than the minimum, null distance points, points recorded during the filling time, discrepant yield points, grain flow delay, start/end pass delay, speed outside the specific range, rapid velocity changes, yield outside the specific interval, standard deviation of yield and specific points removal (Menegatti and Molin, 2004; Sudduth and Drummond, 2007). In some cases, a unique filter can provide maps with quality for management zone delineation, depending on the user´s experience.

3.3 Data Selection

The data collected can include inconsistent, irrelevant or redundant data, impacting the quality of management zones (Witten *et al.*, 2011). Due to the negative effect of redundant, irrelevant or inconsistent attributes, it is important to select an adequate subset of attributes for delineating management zones. In addition, it is important to eliminate or to reduce the correlation between attributes (Jiang *et al.*, 2011) for avoiding distorted or biased management zones.

This step (Figure 4) is used to determine the most relevant data for management zone delineation. It is divided into some tasks starting with checking for grid adjustment before data analyzing, when needed. Our suggestion to solve this problem is using interpolation after spatial analysis. After this, the specialist should apply algorithms to remove inconsistent, irrelevant and redundant attributes besides an algorithm to identify the most relevant attributes.

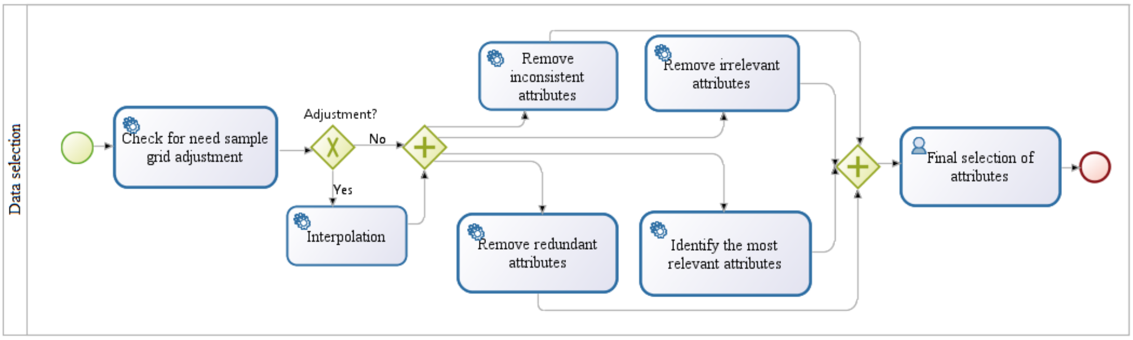

Figure 4. Reference process for data selection modelled in BPMN.

Due to the importance of data selection, there are many strategies and methods for executing these tasks. Witten *et al.* (2011) described strategies and methods for attribute selection, many of them available in the Weka Software. Besides, other statistical software can help select attributes, such as R and SPSS. One method often used is the principal component analysis.

Another option is the choice of attributes based on criteria such as the stable nature of the attributes, as in Derby *et al.*, 2007 using elevation and soil electrical conductivity due to their stable nature. At the end of this step, attributes that best represent field characteristics are selected.

3.4 Clustering

The soil in agricultural areas has different physical, chemical and relief characteristics, resulting in different yield in the same field. To homogenize the field, to maximize yield and/or profitability and to enable variable rate management, agriculture specialists often use soil data and plant yield to identify homogeneous areas in the field.

Therefore, the grouping or management zone delineation maps aim to group the sample points considering the spatial variability of the attributes used. This task results in groups/management zones with higher homogeneity and higher heterogeneity between groups, resulting in the best possible division of the field. Due to the unsupervised classification characteristic, the task results in grouping by similarity points, allowing the use of a wide range of clustering algorithms to generate management zone maps.

Among the different clustering methods presented in Witten *et al.* (2011), the most widely used in studies of management zone delineation is fuzzy c-means (Santos *et al.*, 2012). Another approach is to consider all the points as an image, allowing the application of segmentation algorithms and image processing, such as the watershed algorithm.

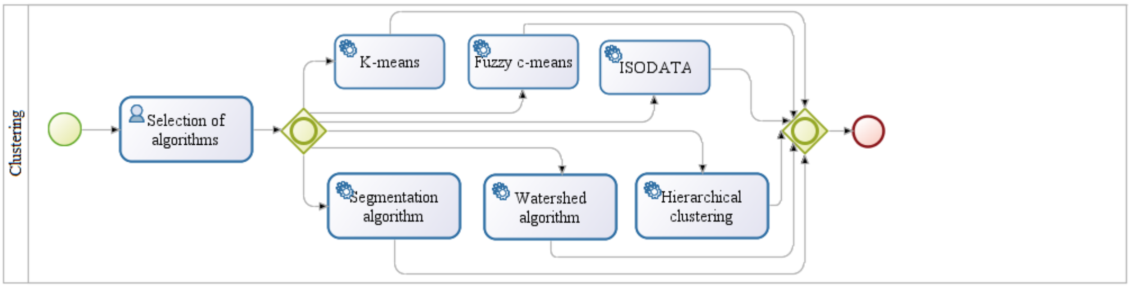

Figure 5. Reference process for clustering modelled in BPMN.

In our model for this step (Figure 5), the specialist should choose one or more clustering algorithms to identify management zones. We are suggesting k-means algorithms, fuzzy c-means algorithms, Iterative Self-Organizing Data Analysis Technique (ISODATA), segmentation algorithms, watershed algorithm and Hierarchical clustering algorithms.

3.5 Maps Evaluating

At the clustering step, the specialist chooses one or more clustering algorithms to identify management zones. In this step, the aim is to determine the quality of these maps allowing the comparison between maps obtained with different methods and/or different parameters. Different statistical methods are used for evaluating and comparing, as shown in Figure 6.

It is important to know what kinds of characteristics are more important in the map obtained to, then, choose the statistical methods for map evaluation. The results provide information to identify the best map for the management of the field or the need to choose other methods or parameters to get better management zone maps.

The evaluation can range from simple aspects, such as variance analyses between groups, to more complex, such as statistical comparison between maps to quantify how significantly different they are. An alternative is to evaluate the correlation between points of management zone maps generated using soil attributes in comparison with yield maps divided into yield range.

Among the statistical methods for comparing groups, we can mention Hotelling-T2, McNemar's test, chi-square test, kappa coefficient, T-test, One-way ANOVA, F Pseudo-test and Tukey test, but besides the statistical methods, the methodology for comparison may be different. The specialist should choose one or more of those methods to identify the best map.

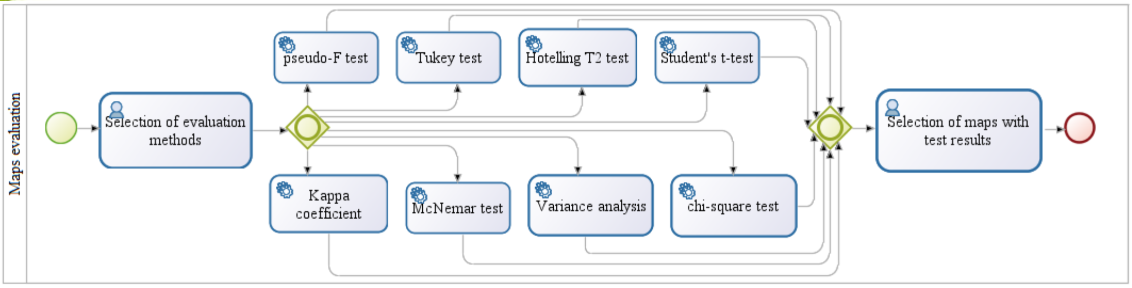

Figure 6. Reference process for maps evaluation modelled in BPMN.

4. CONCLUSION

The reference process for delineating management zones identified the steps necessary from data collection to generation and evaluation of management zone maps. At each step, we identified knowledge areas, methods and tools used for the appropriate treatment of data and also emphasized the importance of each step.

The results showed that despite the differences in the steps used by each author, some are recurrent, but on the other hand, steps not considered can have a strong impact on the maps obtained. Due to the importance of the site-specific treatment in field, this reference process proposed aimed at helping to generate management zone maps and, consequently, the rational treatment of the soil and increase in yield.

The process can be used to guide agronomists and farmers to obtain management zone maps, helping with the choice of data, better tools and methods for obtaining management zone maps that best represent variations in the field. In addition, this study is useful as a reference document for standardization in management zone delineation, and can be used as a guide to choose the right tools, right data and methods, allowing the optimization of management zone delineation and improving the quality of the maps obtained.

The use of BPMN showed to be adequate to represent steps, their flows, decisions to be made along the process and indicating software service tasks and human tasks. This notation is known to decrease the time spent for software development, and consequently the time spend for development of software for management zone delineation using our model.

5. ACKNOWLEDGMENTS

Fundação de Amparo à Pesquisa do Estado de Mato Grosso (FAPEMAT), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Universidade Federal de Mato Grosso (UFMT), Financiadora de Estudos e Projetos (FINEP).

6. REFERENCES

BECKER, J.; KUGELER, M.; ROSEMANN, M. 2003. Process management. Springer.

DERBY, N. E.; CASEY, F. X. M.; FRANZEN, D. W. 2007. Comparison of nitrogen management zone delineation methods for corn grain yield. Agronomy Journal, v. 99, n. 2, p. 405-414.

FETTKE, P.; LOOS, P.; ZWICKER, J. 2006. Business process reference models: Survey and classification. Business Process Management Workshops, Edited by Ekkart Kindler and Markus Nüttgens, Germany. pp.1-88.

JIANG, Q.; FU, Q.; WANG, Z. 2011. Delineating site-specific irrigation management zones. Irrigation and Drainage, v. 60, n. 4, p. 464-472.

KERRY, R.; OLIVER, M. A.; FROGBROOK, Z. L. 2010. Sampling in Precision Agriculture - Geostatistical Applications for Precision Agriculture. In: (Ed.): Springer Netherlands, p.35-63.

MENEGATTI, L. A. A.; MOLIN, J. P. 2004. Remoção de erros em mapas de produtividade via filtragem de dados brutos. (Removing errors in yield maps by filtering on raw data) Revista Brasileira de Engenharia Agrícola e Ambiental, v.8, n.1, p. 126-134.

OMG. Business Process Model and Notation (BPMN), Version 2.0. 2011-01-03 2011. Disponível em: < http://www.omg.org/spec/BPMN/2.0/PDF >.

SANTOS, R. T.; SARAIVA, A. M.; MOLIN, J. P. 2012. Comparison of algorithms for delineating management zones. In: 11ª International Conference on Precision Agriculture, Indianapolis, Indiana, USA. p. 1-15.

SUDDUTH, K. A.; DRUMMOND, S. T. 2007. Yield Editor: Software for removing errors from crop yield maps. Agronomy Journal, v. 99, n. 6, pp. 1471-1482.

WITTEN, I. H.; FRANK, E.; HALL, M. A. 2011. Data mining: practical machine learning tools and techniques. Burlington, MA: Morgan Kaufmann. pp 1-629.

ZHANG, H.; LAN, Y.; LACEY, R.; HOFFMANN, W. C.; WESTBROOK, J. K. 2011. Spatial analysis of ndvi readings with different sampling densities. Transactions of the Asabe, v. 54, n. 1, p. 349-354.