Huerta, Schade, Granell (Eds): Connecting a Digital Europe through Location and Place. Proceedings of the AGILE'2014 International Conference on Geographic Information Science, Castellón, June, 3-6, 2014. ISBN: 978-90-816960-4-3

Error propagation in a fuzzy logic spatial multi-criteria evaluation

Lisa Bingham
University of Stavanger
Department of
Petroleum Engineering
Stavanger, Norway
lisa.bingham@uis.no

Derek Karssenberg
Utrecht University
Department of Physical
Geography
Utrecht, the Netherlands
d.karssenberg@uu.nl

Abstract

Quantifying errors in results of spatial multi-criteria evaluation (MCE) techniques is essential to improve the credibility of MCE in planning and decision-making. We present an error propagation procedure using Monte Carlo simulation for fuzzy logic MCE applied in a case study of petroleum exploration which covers northern South America. The fuzzy logic MCE combines data sets to evaluate the favourability of petroleum exploration in a geographic region. Each input data set has associated error models and estimated uncertainty. Two sources of error are investigated: boundary and fuzzy membership. 2000iterations of the model were run. The resulting mean of the 2000 samples and a series of confidence interval maps were analysed. It is concluded that the combination of the MCE analysis and error propagation modelling will support decisions for petroleum exploration.
Keywords: error propagation, fuzzy logic, multi-criteria evaluation, petroleum.

1 Introduction

Multi-Criteria Evaluation (MCE) is a subset of multidimensional decision and evaluation models that essentially are tools to evaluate the trade-offs between alternatives with different impacts [3]. The goal of MCE is to evaluate the outcome of combining different criteria to fulfil one or more objectives that may possibly be conflicting [3]. It is important to understand that all types of MCE are subjective and different methods may give different results (e.g. [5]).

Bingham et al. [1] proposed a spatial MCE method based on a series of user-defined inputs and criteria that can be used to show geographic areas that may be of interest for further investigation for petroleum exploration (Figure 1). By using the results of the analysis, the geologist can support his reasons for requiring more or less investigation in a geographic region for future exploration.

The purpose of the model is to produce maps for petroleum exploration; high-scoring areas are more favorable for petroleum exploration based on the criteria (the input data sets; the top level of boxes shown in Figure 1). The input data sets are assigned fuzzy membership values [0,1] where 0 is unfavorable and 1 is favorable. The MCE uses fuzzy logic operators (FLOs; AND, OR, GAMMA, ALGEBRAIC SUM) to combine the input data sets. The operators work on a cell-by-cell basis (i.e. raster algebra); thus all input data sets must be in the same coordinate reference system (projection), have the same cell size, and cover the same spatial extent [2].

Following the model framework (Figure 1), the favorability of petroleum exploration can be modeled either by specific age intervals which would be appropriate for determining drilling targets or by examining all of the available data (i.e. non-age-specific) which would be appropriate for a general overview of petroleum exploration favorability.

All models, including MCE, have some amount of uncertainty related to their input, parameters, and results, which when unknown decrease the reliability of decisions based on the output [6, 7]. The fuzzy logic MCE proposed has some uncertainty associated with the input data sets. This paper applies error propagation using Monte Carlo simulation to a fuzzy logic multi-criteria evaluation. The original work [1] acknowledged that there is some uncertainty with the results but it was not quantified or evaluated thoroughly. This work aims to rectify this lack of information. It specifically investigates:

- What kind of error models should be applied given the data?
- In the case study, how do the confidence intervals and mean compare to the original result?
- How reliable are the results from the MCE?

The sources of uncertainty are neither gross (i.e. due to negligence or carelessness) nor systematic (i.e. having a functional relationship), but random [6]. The error of the input data sets can be estimated; thus, error propagation is a suitable method for investigating the estimated error effects on the final output [6]. Uncertainty propagation modelling is important in order to assess how model input errors propagate to the model results, in order to quantify the uncertainty in the model results [5, 8]. Any uncertainty propagation modelling is subjective based on the parameters of the uncertainty modelling.

Monte Carlo simulation is a technique, which calculates the model repeatedly using different input values based on an error model [4]. By interpreting the culminated results (e.g. 95% confidence interval, median) of hundreds or thousands of samples, the modeller can determine if the model results are reliable.

2 Approach

Three main sources of uncertainty in a model are parameters, inputs, and model structure [7]. Uncertainty related to inputs is investigated in this study; parameters and model structure are not investigated. Input data uncertainty is common in a GIS environment where the user relies on many different sources for data, especially if data is older or the original source is unknown. Border classification and fuzzy membership assignment are the only sources of uncertainty that are investigated in this paper.

The workflow of the uncertainty error modelling follows a series of steps to create realizations for the investigated uncertainties. After a number of realizations are completed, the composite probabilities, average, and variance are calculated.

3 Case study

The study area focused on northern South America from Colombia in the west to Guyana in the east; the data was compiled and converted to PCRaster format following general conventions. For more details on the data compilation the reader is referred to Bingham et al. [1]. The authors assume that the attribute classifications of the data are correct and do not contain gross or systematic errors. The uncertainty for all input data sets is estimated to be 10%, except in the case of geochemical and seeps data which were estimated to be 30%.

The size of the study area covers more than 2.5 x 10^6 km^2. Data were compiled from sources ranging in scale from 1:100,000 to 1:44,000,000 and converted to raster format using a customized Albers projection and a cell length of 1000 m.

The Monte Carlo simulation ran 2000 samples for a non-age-specific model. Each simulation took approximately 15 hours on PCRaster v. 3.0 with python v. 2.6 on a CentOS 5, 64-bit, 8 Xeon 2.66 GHz processors, and 32 GB RAM.

3.1 Results

The Monte Carlo simulation focused on the non-age-specific favorability model. This model included all of the data meeting the criteria regardless of age. The original model result is shown in Figure 2 for comparison. Figure 3 shows the output map of the average from 2000 samples. The original model result (Figure 2) is very similar to the average; both maps have approximately the same maximum range (0.78 vs 0.79) and reflect the same pattern for high versus low favorability areas. The average of the 2000 samples shows more and larger areas of high favorability. Figure 4 shows the 2000 samples at a 95% confidence interval with a threshold value of 0.7; in the original model results, this value was arbitrarily chosen as an indicator of high favorability where additional exploration may be justified. This combination of confidence interval and threshold value results in areas that are certain to be below a favorability value of 0.70, areas that are certain to be above this value, or areas for which it is uncertain whether the favorability will be above or below 0.70. In this map, areas where it is certain the favorability value is above 0.70 are indiscernible. However, by decreasing the threshold value to 0.50, some areas are now certain to be above the threshold value (Figure 5). If the exploration geologist finds it appropriate, the confidence interval can be increased to 50% and more areas will lie above the threshold value (Figures 6 and 7).

4 Conclusions

The Monte Carlo simulation investigating error propagation of fuzzy logic multi-criteria evaluation shows that the model and its inputs are reasonably reliable; however, the data itself may be improved to reduce the amount of uncertainty. The confidence interval and threshold value must be chosen with care and communicated to the end user. With improved data, the multi-criteria evaluation in conjunction with the Monte Carlo simulation can provide a useful tool for decision-making in petroleum exploration.

References

[1] Bingham, L., R. Zurita-Milla, and A. Escalona, GIS-based Fuzzy Logic for Petroleum Exploration. *AAPG Bulletin*, 96:2121-2142, 2012.

[2] Bonham-Carter, G. F., *Geographic Information Systems for Geoscientists Modelling with GIS: Computer Methods in the Geosciences Volume 13*: Oxford, Pergamon, 1994.

[3] Carter, S., Site search and multicriteria evaluation. *Planning Outlook*, 34:27-36, 1991.

[4] Heuvelink, G. B. M., *Error propagation in environmental modelling with GIS*. Taylor and Francis, London, 1998.

[5] Heywood, I., J. Oliver, and S. Tomlinson, Building an exploratory multi-criteria modelling environment for spatial decision support. *Innovations in GIS*, 2:127-136, 1995.

[6] Thapa, K., and J. Bossler, Accuracy of Spatial Data Used in Geographic Information Systems. *Photogrammetric Engineering and Remote Sensing*, 58:835-41, 1992.

[7] Verstegen, J. A., D. Karssenberg, F. van der Hilst, and A. Faaij, Spatio-temporal uncertainty in Spatial Decision Support Systems: A case study of changing land availability for bioenergy crops in Mozambique. *Computers, Environment and Urban Systems*, 36:30-42, 2012.

Figure 1: Workflow of the model. Data sets are combined starting at the top level and moving subsequently down.

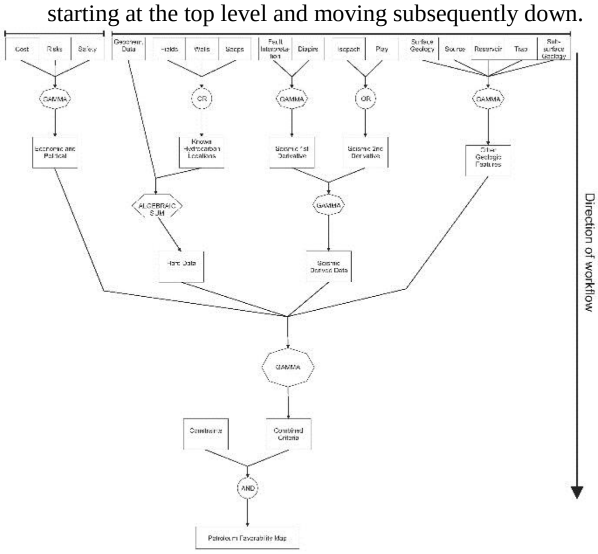

Figure 2: Original result showing petroleum exploration favorability [1]. Darker grey reflects higher favorability..

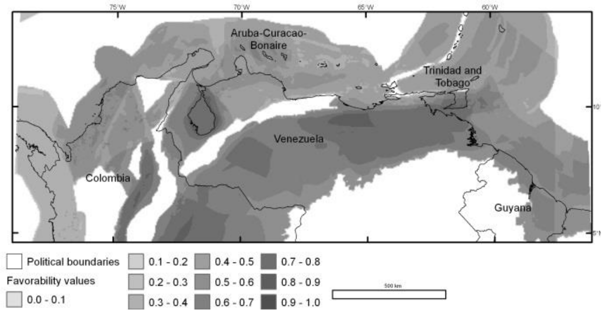

Figure 3: Average result of 2000 samples showing petroleum exploration favorability. Darker grey reflects higher favorability.

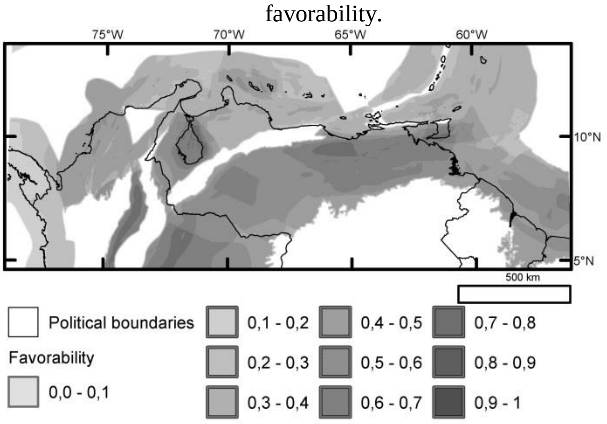

Figure 4: 95% confidence interval with threshold 0.70.

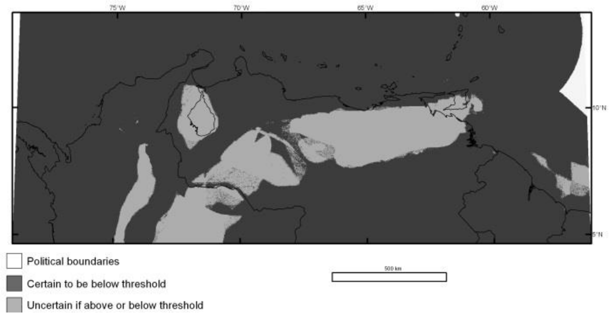

Figure 5: 95% confidence interval with threshold 0.50.

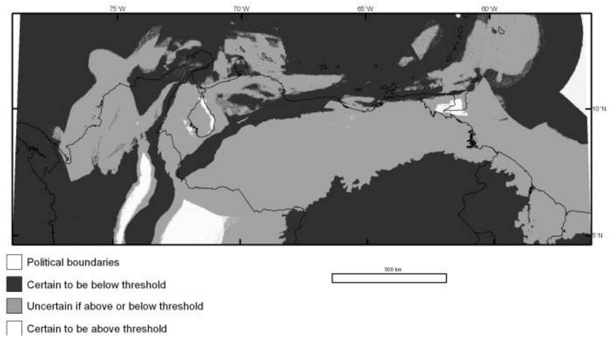

Figure 6: 50% confidence interval with threshold 0.70.

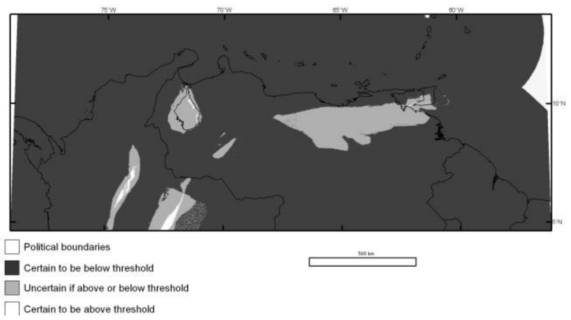

Figure 7: 50% confidence interval with threshold 0.50.