Taylor & Francis
Taylor & Francis Group

Error propagation in a fuzzy logic multi-criteria evaluation for petroleum exploration

L. Bingham, A. Escalona & D. Karssenberg

⊞ View supplementary material ↗

▦ Published online: 18 Feb 2016.

✎ Submit your article to this journal ↗

▥ Article views: 152

Q View related articles ↗

⊞ View Crossmark data ↗

Taylor & Francis
Taylor & Francis Group

Error propagation in a fuzzy logic multi-criteria evaluation for petroleum exploration

L. Bingham[a], A. Escalona[a] and D. Karssenberg[b]

[a]Department of Petroleum Engineering, Faculty of Science and Technology, University of Stavanger, Stavanger, Norway; [b]Department of Physical Geography, Faculty of Geosciences, Utrecht University, Utrecht, The Netherlands

ABSTRACT

This article applies error propagation in a Monte Carlo simulation for a spatial-based fuzzy logic multi-criteria evaluation (MCE) in order to investigate the output uncertainty created by the input data sets and model structure. Six scenarios for quantifying uncertainty are reviewed. Three scenarios are progressively more complex in defining observational data (attribute uncertainty); while three other scenarios include uncertainty in observational data (position of boundaries between map units), weighting of evidence (fuzzy membership assignment), and evaluating changes in the MCE model (fuzzy logic operators). A case study of petroleum exploration in northern South America is used. Despite the resources and time required, the best estimate of input uncertainty is that based on expert-defined values. Uncertainties for fuzzy membership assignment and boundary transition zones do not affect the results as much as the attribute assignment uncertainty. The MCE fuzzy logic operator uncertainty affects the results the most. Confidence levels of 95% and 60% are evaluated with threshold values of 0.7 and 0.5 and show that accepting more uncertainty in the results increases the total area available for decision-making. Threshold values and confidence levels should be predetermined, although a series of combinations may yield the best decision-making support.

1. Introduction

Multi-criteria evaluation (MCE), also referred to as multi-criteria analysis and multi-criteria decision analysis, is a subset of multidimensional decision and evaluation models that essentially are tools to evaluate the trade-offs between alternatives with different impacts (Carter 1991). The goal of MCE is to evaluate the outcome of combining different criteria to fulfill one or more objectives that may possibly be conflicting (Carter 1991, Heywood *et al.* 2006); MCE can be with or without a spatial component and is applicable in any discipline or field.

Supplemental data for this article can be accessed here. Uncertainty probability distribution graphs and maps for all input data sets for Scenario 3.

Spatial MCE is used in many different disciplines in order to evaluate suitabilities (or favorabilities) of locations. This article follows Bonham-Carter *et al*. (1988) in using the term *favorability*. Defining the most favorable location may be positive as in the best location or negative as in the worst location. Voogd (1983), Carver (1991), and Bonham-Carter (1994) have conducted comprehensive and in-depth studies of spatial MCE; Malczewski (2006) provides a recent survey of literature from 1990 to 2004. Previous spatial MCE studies focusing on economic geology have used either Bayesian methods or fuzzy logic (e.g. Bonham-Carter *et al*. 1988, An *et al*. 1994, Wright and Bonham-Carter 1996, Tangestani and Moore 2002, Tounsi 2005). Both methods are knowledge-driven relying on information from an expert. Because Bayesian methods require prior probabilities, they are most appropriate for areas with well-distributed and sufficient data; thus, they are less appropriate for frontier exploration areas where data distribution is sparse (Bonham-Carter 1994). Therefore, fuzzy logic MCE is most appropriate for a variety of exploration settings.

Spatial MCE is useful for modeling favorable locations on a set of criteria. However, it is extremely important that the modeler and the decision-maker understand the implications of subjectivity within the method (Heywood *et al*. 1995) and the uncertainty related to the MCE model input, parameters, structure, and results (Thapa and Bossler 1992, Heywood *et al*. 1995, Karssenberg and De Jong 2005, Li *et al*. 2012). Because of the inherent subjectivity of MCE and the uncertainty both from the data and the parameters used in the MCE, there exists a need to quantify the uncertainty.

Data uncertainty is common in any spatial analysis, including MCE modeling, where the user relies on many different sources for data, especially if data is older or the original source is unknown. Model input uncertainty is generally associated with data collection or processing; data completeness for time and space; data resolution; and errors from digitizing data (Heuvelink 1998, Heywood *et al*. 2006). Model parameters and structure may have some uncertainty associated with them due to the decisions made by the modeler in how to combine the data sets (Heuvelink 1998, Heywood *et al*. 2006).

There are several methods for investigating uncertainty in models. Error propagation is an analysis used to investigate uncertainty and is a suitable method for calculating the uncertainty of the final output (Thapa and Bossler 1992). The uncertainty, therefore, influences all subsequent calculations in the model. Monte Carlo simulation, used here, is an error propagation technique that calculates the model repeatedly using different input values based on an error model (Heuvelink 1998). By interpreting the culminated results (e.g. 5th percentile, 95th percentile, and median) of hundreds or thousands of realizations, the modeler can evaluate spatial variations for further analysis (i.e. confidence level tests).

There are a limited number of published examples of spatial MCE and error propagation combined. Moon (1998) combined spatial MCE, fuzzy logic, and error propagation in a theoretical presentation to show their importance and possible theoretical application in geosciences. The paper did not produce examples or models but an umbrella idea of applications. Davis and Keller (1997) applied fuzzy logic with Monte Carlo simulation in a spatial setting in order to investigate uncertainty related to attribute boundaries and attribute assignment for slope stability. Fernández and Lutz (2010) and Feizizadeh and Blaschke (2014) are two recent papers which focus on both spatial MCE and error propagation. Fernández and Lutz (2010) apply Taylor's series for error propagation and use a combination of defined and assumed uncertainties. Feizizadeh and Blaschke (2014) focus on comparing different MCE methods.

This article focuses on applying error propagation using Monte Carlo simulation to a spatial fuzzy logic MCE for petroleum exploration using geologic data. A case study from northern South America is used because the region is oil-rich and has underexplored areas. This article addresses three questions:

- What are possible methods for defining uncertainty in the various components of the MCE model and can similar results be found with different methods requiring less effort?
- How do the different scenarios for defining uncertainty compare to each other?
- What is the MCE error magnitude (on average) from the defined uncertainty on the results?

The innovation of this article is the combination of spatial MCE, fuzzy logic, and error propagation applied to petroleum exploration.

The remainder of the article is divided into five sections. First, the spatial MCE is reviewed in Section 2; Section 3 reviews the error propagation and six scenarios for evaluating the MCE. Sections 4–6 of the article present and discuss the results and provide the conclusions.

2. Spatial MCE

Bingham *et al*. (2012) proposed a spatial-based fuzzy logic MCE with applications to petroleum exploration such that the investigating geologist may choose a geographic area for further and more intensive investigation for proposed hydrocarbon prospects, but does not pinpoint specific locations for drilling. The method is intended to reflect the data evaluation process of a geologist in a recordable and repeatable manner. The method can be applied to geologic age-specific data (e.g. Cretaceous interval consisting of substrata that fall within the Cretaceous age between ca. 145 Ma and 65 Ma) or to non-age-specific data by combining all of the available age-specific data (e.g. Cretaceous to recent). This article focuses only on the non-age-specific evaluation of a case study of northern South America (Bingham *et al*. 2012).

2.1. *Data*

The MCE combines data covering five criteria: economic and political, hard (or directly measured), seismic derived, other geologic, and constraints. The criteria are composed of 17 subcriteria; Table 1 lists the criteria and subcriteria with their original sources. All subcriteria were either compiled or created by the researchers as part of the Caribbean Basins, Tectonics, and Hydrocarbons consortium (www.cbth.uh.edu) and were collected from scales varying from 1:100,000 to 1:44,000,000. Bingham *et al*. (2012) discuss the data selection process. The case study extends from approximately 80°W to 55°W and 3° N to 15°N (Figure 1). A customized Albers projection with central meridian 67.5°W, two standard parallels of 7°N and 11°N, and WGS1984 as the datum is used to preserve the relative distance and area (USGS 1989). A cell size of 100 km^2 is used due to the extensive size of the study area (2.5×10^6 km^2), the lack of large-scale data, dense data sets, and the goal of the MCE to have a general idea of favorable prospect areas rather than pinpointing specific drilling sites.

Table 1. Criteria sources.

5 Criteria	17 Subcriteria	Sources and information	No. of data sets
Economic and political (E)	Costs (C)	A highly general map intended to reflect relative general costs only	1
	Safety (Y)	A highly general map intended to reflect relative safety levels only	1
	Risks (R)	Data set created by AON Group, Inc. (2010)	1
Hard (H)	Geochemical (G)	A general map showing total organic content	1
	Fields (F)	Data sets collected from several sources (DM2 Project, 2005; Castellanos *et al.*, 2006; Wood McKenzie, 2006; Staatsolie, 2007; EPIS and ANH, 2008)	1
	Wells (W)	Data sets collected from several sources (French and Schenk, 2004; DM2 Project, 2005; Wood McKenzie, 2006; Staatsolie, 2007; EPIS and ANH, 2008)	1
	Seeps (S)	Data sets collected from several sources (DM2 Project, 2005; Staatsolie, 2007; EPIS and ANH, 2008)	1
Seismic derived (D)	Fault density (U)	Fault densities are calculated for 10,000 square kilometers from faults identified on 2D seismic data; there are four stratigraphic surfaces: Cretaceous, Paleogene, Miocene, surface/seafloor	4
	Diapirs (A)	Data sets collected from several sources (Kugler, 1959; Valery et al., 1985; Brown and Westbrook, 1988; Beltran, 1993; Deville *et al.*, 2003; DM2 Project, 2005; Sullivan, 2005; ANH, 2007; Duerto, 2007)	1
	Isochrons (I)	Three isochrons (basement/Cretaceous-Paleogene; Paleogene-Miocene; Miocene-Recent/surface) calculated from interpolated 2D and 3D seismic data	3
Other geologic (O)	Plays (P)	A general map showing play risks	1
	Surface geology (Q)	Data sets collected from several sources (Geologisch Mijnbouwkundige Dienst, 1977; Walrond, 1987; Saunders and Snoke, 1998; Schenk *et al.*, 1999; French and Schenk, 2004; Garrity *et al.*, 2006; Gomez Tapias *et al.*, 2007)	1
	Source rock (X)	Two data sets, Cretaceous and Lower Cenozoic, show interpretive extents based on several publications (Curet, 1992; Ysaccis, 1997; Di Croce *et al.*, 1999; Sanchez, 2007; Kroehler, 2007; Yang and Escalona, 2011; and references cited therein)	2
	Reservoir rock (V)	Three data sets, Cretaceous, Paleogene, and Miocene, show interpretive extents	3
	Traps (T)	A general map showing trap presence/absence	1
	Subsurface lithology (L)	Three data sets, Cretaceous, Paleogene, and Miocene, based on paleogeographic maps	3
Constraint	Sedimentary basin (N)	Data set edited from Fugro Data Services, AG. (2005) and a buffer of five kilometers was applied	1
TOTAL			27

Letters in parentheses by criteria and subcriteria are the equation symbols.

The subcriteria maps were created using expert opinion by assigning fuzzy membership values [0,1], where 0 is unfavorable, and 1 is favorable. Table 2 lists the subcriteria, attributes, and corresponding fuzzy membership values. The fuzzy membership allows for partial truth where Boolean membership only allows true or false (Bonham-Carter 1994, Jiang and Eastman 2000, Tounsi 2005). A fuzzy membership value of 0 is avoided due to the mathematics (i.e. multiplication) of the MCE. The fuzzy membership values and classifications are not rigidly or universally set as the values may be altered for other geographic regions, with additional data, and/or more expertise.

2.2. MCE framework

The subcriteria are combined in aggregation stages using fuzzy logic operators (FLO; MIN (minimum), MAX (maximum), GAMMA, ALGEBRAIC SUM). The operators work on a
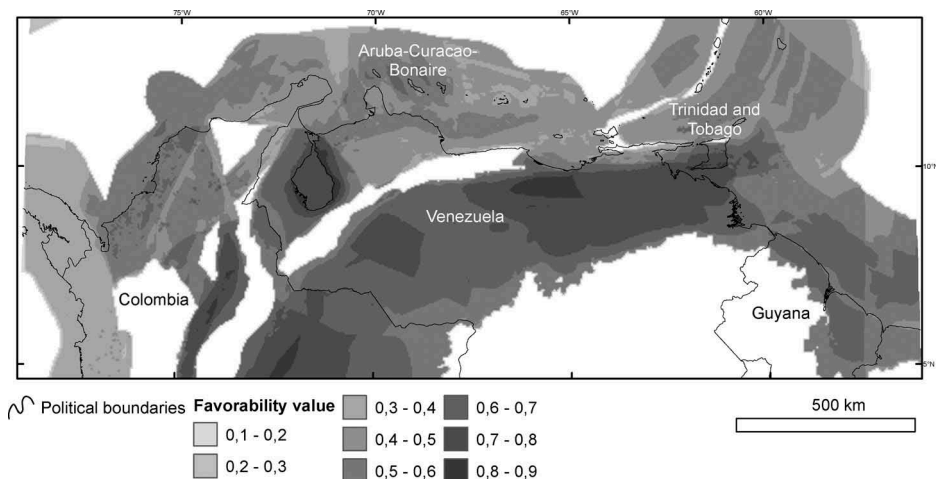
Figure 1. Case study results (Bingham *et al.* 2012) of non-age-specific map from fuzzy logic MCE. Color scale represents more to less favorable, where light gray is least favorable and black is most favorable. White areas were removed by the constraints.

cell-by-cell basis (i.e. raster algebra). Bonham-Carter (1994) and Zimmerman and Zysno (1980) provide thorough explanations of the operators, while a simplified explanation of GAMMA and ALGEBRAIC SUM follows here. The GAMMA FLO uses an exponent called gamma with a value between 0 and 1; it can result in fuzzy membership values that lie between the extremes of the other FLOs (Bonham-Carter 1994). If the GAMMA FLO is used in more than one instance in a model, the gamma value does not need to be the same for each instance. The algebraic sum 'is always larger (or equal to) the largest contributing fuzzy membership value' (Bonham-Carter 1994). The MCE workflow is shown in Figure 2. The subcriteria are combined in three stages. Stage 1 aggregates the subcriteria to the criteria level. Stage 2 combines the criteria together. Stage 3 removes the constraint areas producing the final favorability map.

A brief description of the proposed MCE follows; for a full explanation of the method, equations, and input data sets, see Bingham *et al.* (2012). The first step (Figure 2) combines the economic and political subcriteria using the GAMMA FLO to allow for flexibility (by changing the gamma value) based on a company's risk policy:

$$E = (1 - (1 - C)(1 - R)(1 - Y))^{\gamma_1}(CRY)^{(1-\gamma_1)} \tag{1}$$

where E is the economic and political criteria, C is cost, R is risks, Y is safety, and γ_1 is the gamma value applied to the economic and political subcriteria.

The second step combines the hard subcriteria, which are considered 'hard data' because they are associated with direct measurements (e.g. oil is or is not present in a well vs. the interpreted location of a fault on seismic data). As the subcriterion with the highest favorability should be assigned, the MAX FLO is used to combine fields, wells, and seeps. The result is combined with the geochemical data using the ALGEBRAIC SUM FLO, which results in the highest possible output.

Table 2. Fuzzy membership values applied in case study.

Subcriteria	Attribute	FM value	Subcriteria	Attribute	FM value
Costs (C)	Very low	1.0	Diapirs (A)	Present	0.7
	Low	0.7		Absent	0.3
	Intermediate	0.5	Isochrons in milliseconds	0	0.1
	High	0.3	two-way-time (I)	0.0–1000	0.3
	Very high	0.1		1000–3000	0.5
Risks (R)	Low	0.9		3000–7000	0.8
	Low-medium	0.7		7000–9000	0.5
	Medium	0.5		>9000	0.3
Safety (Y)	Very high	1.0	Plays (P)	None	0.01
	High	0.9		Low	0.9
	Med-high	0.7		Low-Medium	0.7
	Low	0.3		Medium	0.5
	Very low	0.1		Med-high	0.3
Geochemical data (G)	Low-med	0.4		High	0.1
	Medium	0.5	Surface geology (Q)	Igneous	0.1
	Med-high	0.7		Metamorphic	0.1
	High	0.9		Sedimentary	0.9
	Very high	1.0		Mix	0.5
	None	0.01	Source rock (X)	Present	1.0
Oil and gas fields (F)	Present	1.0		Absent	0.1
	Absent	0.1	Reservoir rock (V)	Present	1.0
Oil and gas wells (W)	Dry well	0.01		Absent	0.1
	Water well	0.01	Traps (T)	Present	1.0
	Show	0.7		Mix	0.5
	Producing	1.0	Subsurface lithology (L)	Limestone	0.8
	Other	0.01		Shale	0.8
Oil and gas seeps (S)	Other seep	0.5		Sandstone	0.8
	Oil or gas seep	0.9		Combination	0.8
	No seep	0.1		No deposition	0.1
Fault density, surface (U)	0	0.9		Other	0.1
	0.0–0.42*	(0.9, 0.1)	Sedimentary basins (N)	Present	1.0
	0.42+	0.1		Absent	0.0
Fault density, subsurface (U)	0	0.1			
	0.0–0.42*	(0.1, 0.9)			
	0.42+	0.9			

*A linear relationship between the attribute and the fuzzy membership value.

$$H = 1 - (1 - (\text{MAX}(F, W, S)))(1 - G) \qquad (2)$$

where H is the hard criteria, F is hydrocarbon fields, W is drilled wells, S is hydrocarbon seeps, and G is geochemical subcriterion.

Three equations (Equations (3)–(5)) combine the seismic-derived subcriteria. Equation (3) combines seismic derived interpretation that can be directly observed in the seismic data and uses the GAMMA FLO to be able to reflect the seismic interpreter's knowledge or even to reflect if the interpretation is considered thorough or superficial. The subcriteria in Equation (4) represent a further state of interpretation and cannot be seen directly from the seismic data; the subcriterion with the higher favorability is chosen in case of null values. The outputs from Equations (3) and (4) are combined for the seismic-derived favorability value (Equation 5) and use the GAMMA FLO to reflect the overall confidence or exploration policy.
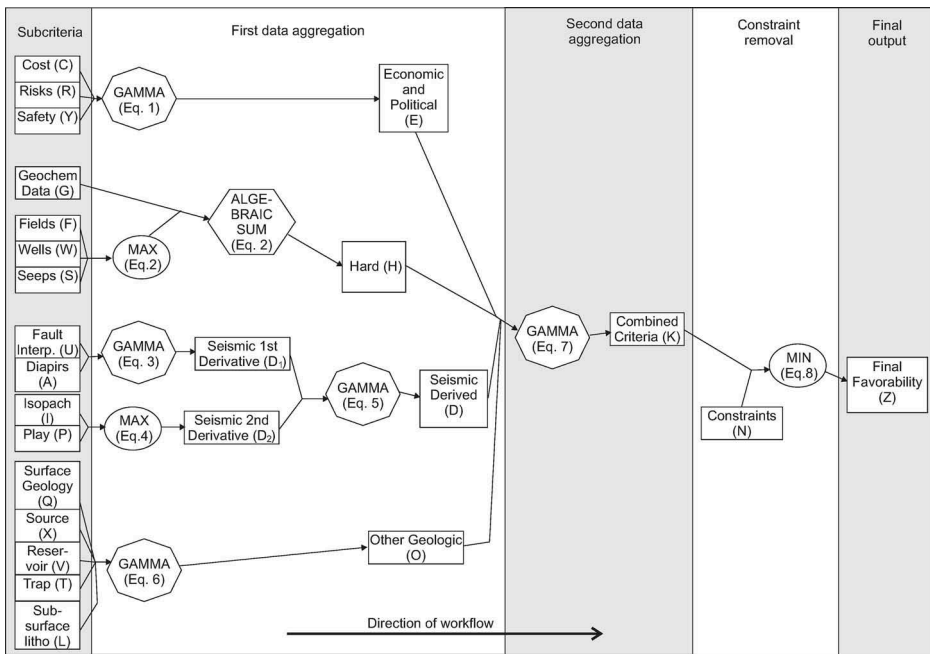
Figure 2. MCE workflow. The left column of boxes shows the subcriteria. All successive boxes are the combined data at different aggregation stages. The other shapes (hexagon, octagon, and circle) match a different fuzzy logic operator and contain the equation number in the text as well. All shapes contain the letter used in the equations in the text. The subcriteria are combined from left to right.

$$D_1 = (1 - (1 - U)(1 - A))^{\gamma_2} (UA)^{(1-\gamma_2)} \tag{3}$$

$$D_2 = MAX(I, P) \tag{4}$$

$$D = (1 - (1 - D_1)(1 - D_2))^{\gamma_3} (D_1 D_2)^{(1-\gamma_3)} \tag{5}$$

where D_1 is the combined first seismic derived subcriteria; U is fault interpretation; A is diapirs; D_2 is the combined second seismic derived subcriteria; I is isopach/isochron; P is plays; and D is the seismic derived criteria.

The final criteria group, other geologic features, is composed of more interpretive geologic subcriteria that are used for exploration, and thus uses the GAMMA FLO to reflect the degree of interpretation more flexibly.

$$O = (1 - (1 - Q)(1 - X)(1 - V)(1 - T)(1 - L))^{\gamma_4} (QXVTL)^{(1-\gamma_4)} \tag{6}$$

where O is the combined other geologic features criteria, Q is surface geology, X is source rock, V is reservoir rock, T is trap, and L is subsurface lithology.

All of the criteria are then combined using the GAMMA FLO.

$$K = (1 - (1 - E)(1 - H)(1 - D)(1 - O))^{\gamma_5} (EHDO)^{(1-\gamma_5)} \tag{7}$$

where K is the combined criteria.

To obtain the final favorability, constraints must be applied. The exploration should be restricted within the limits of sedimentary basins. While it is acknowledged that areas excluded in the case study may contain some sedimentary deposits, the sedimentary layers are not currently thought to be thick enough to contain mature hydrocarbon deposits. Therefore, for simplicity, these areas are considered nonsedimentary basins in this article. Using the MIN FLO ensures the selection of the favorability within the sedimentary basin if the true value is 1 (i.e. sedimentary basin is 1, nonsedimentary basin is 0, and a favorability value within the sedimentary basin [0,1] will be chosen).

$$Z = MIN(N, K) \tag{8}$$

where Z is the final favorability and N is the constraints criteria.

Figure 1 is the final favorability map of the non-age-specific petroleum exploration potential of northern South America. The highest scoring favorability areas are found in Venezuela and Colombia. Non-age-specific favorability will indicate whether the area in general, regardless of possible exploration targets (age intervals of subsurface sediments), is of low to high favorability for exploration. Exploration may occur in well-known and established areas of production or in areas where there is little or no production (i.e. frontier exploration).

3. Error propagation

The workflow of the uncertainty modeling follows a series of steps to create realizations for the investigated uncertainties. The following steps outline the workflow of the uncertainty modeling with Monte Carlo simulation (Heuvelink 1998).

(1) Error models are defined for each input criterion.
(2) Create 2000 sets of realizations of all inputs by drawing from the probability distributions (defined by the error models in step 1).
(3) Run the MCE model (Equations (1)–(8)) for each set of realizations and store results.
(4) For final favorability (Z), calculate sampling statistics over realizations.

In order to investigate systematically the sources of uncertainty, six scenarios are created and evaluated. The scenarios evaluate three basic groups of uncertainty. Group A focuses on uncertainty in observational data, specifically addressing the uncertainty related to feature attribute classification (Column 2 'Attribute' in Table 2). This uncertainty addresses the possible misclassification of a feature to the wrong attribute class by either negligence or mapping error. Group B has two parts and focuses on uncertainty of weighting of evidence with a component of observational data. The two parts are fuzzy membership assignment (Column 3 'FM Value' in Table 2; weighting of evidence) and boundary transition zone (observational data). The uncertainty of the fuzzy membership assignment is related to the expert opinion that a specific attribute should be assigned a specific fuzzy membership value. The boundary transition zone is the uncertainty that the boundary between two classified areas is geographically mislocated either due to mapping errors, classification errors, or rasterization processes. Group C focuses on changing model parameters (i.e. FLOs). The different scenarios

combine the groups in a successive manner. Scenarios 1–3 focus on Group A; Scenario 4 combines groups A and B; Scenario 5 combines groups A and C; and Scenario 6 combines all groups. The results of the different scenarios are evaluated by comparing two different threshold values (0.5, 0.7) and two different confidence levels (60%, 95%).

3.1. *Scenario 1: 'simple' uncertainty*

Scenario 1 acts as the base case of the six scenarios. It assumes that all input data set feature classifications (i.e. Group A) have an uncertainty probability of 0.1 for being misclassified. This value was chosen as most appropriate based on informed expert knowledge. In Scenario 1, this uncertainty is formally represented as follows: class numbers are defined $i = 1, ..., N$ with N, the total number of classes. X is a discrete random variable representing the class of a feature. The probability of the originally mapped class $i = k$ is:

$$\Pr(X = k) = u \tag{9}$$

The probability of the remaining classes $i \neq k$ is:

$$\Pr(X = i) = \frac{1 - u}{N - 1} \quad \text{forall } i \neq k \tag{10}$$

The purpose of this simple scenario is to evaluate the necessity of a more detail-oriented expert-defined uncertainty, i.e. are similar results possible with less effort. Its significance is to compare the analysis results of increased resources and more in-depth modeling. The originally defined classification is assigned a probability of 0.9 (i.e. $u = 0.9$). A probability of 0.1 is distributed equally among the remaining classes as described by $\Pr(X = i)$. For example, for subcriterion geochemical (TOC), polygon 964 was originally classified as high TOC by an expert (Figure 3D). Scenario 1 models with 0.1 probability that the area included in polygon 964 is not classified correctly and may be one of the other classifications (none, low, low-medium, medium, medium-high; see Table 2). A 0.1 probability is distributed equally among these five classifications (Figure 3A). Polygons 963, 968, 969, 970, and 971 were originally noted as 'unknown TOC' (Figure 3D); there-fore, these areas are equally probable as being one of the six classifications: none, low, low-medium, medium, medium-high, and high (Figure 3A).

3.2. *Scenario 2: expert-defined uncertainty spread across neighboring classes*

Scenario 2 aims to be slightly more complex than Scenario 1 in order to evaluate the difference of a detailed expert-defined uncertainty and a limitation to possible misclas-sification. Just like Scenario 1, it focuses on Group A uncertainty only; groups B and C uncertainties are ignored. In this scenario, a geological expert with experience in the case study area defined the uncertainty of each input data set by reviewing the attribute classification and the geographic location of the features. Every feature in every data set was assigned a probability that the feature classification is incorrect. A probability distribution is created using this assigned value. The probability that the feature is misclassified is split equally between the class neighbors. The number of class neighbors
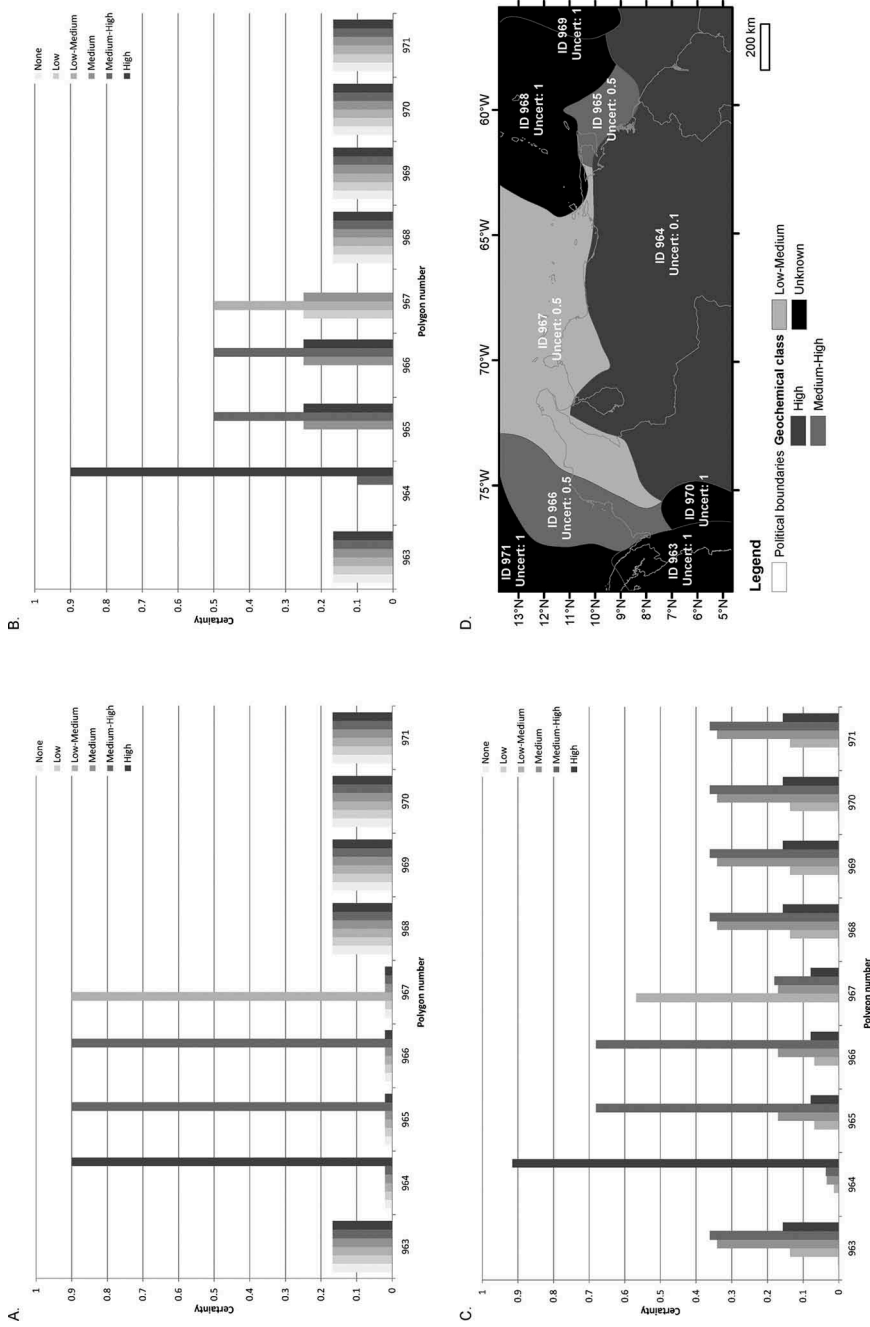
Figure 3. (A) Graph showing the discrete probability distributions representing uncertainty in the geochemical subcriteria following Scenario 1. The *x*-axis is the polygon ID shown in D. The *y*-axis is the probability of a class. Each polygon number shows the distribution of six classes. (B) Graph showing the uncertainty distribution of the geochemical subcriteria following Scenario 2. Each polygon number shows the distribution of two, three, or six classes as the case may be. (C) Graph showing the uncertainty distribution of the geochemical subcriteria following Scenarios 3 through 6. Each polygon number shows the distribution of six classes. (D) Map showing the geochemical subcriteria. Each feature polygon is colored by class and labeled with the polygon ID and expert-assigned uncertainty. Supplemental data provides the probability distributions of other input variables.

is defined $i = 1, ..., N$ with N the total number of classes. X is a discrete random variable representing the class of a feature. The probability of the originally mapped class $i = k$ is:

$$Pr(X = k) = u \tag{10}$$

The probability of the class neighbors $i = k + 1$ and $i = k - 1$ is:

$$Pr(X = k - 1) = Pr(X = k + 1) = \frac{1 - u}{2} \tag{11}$$

The probability of the remaining classes is zero.

For example, polygon 965 was originally classified as 'medium-high' TOC (Figure 3D). The expert-defined uncertainty is a probability of 0.5 that the area is either 'medium' or 'high' TOC (Figure 3B) and is split equally between the two classes (i.e. 0.25 medium, 0.5 medium-high, 0.25 high, because $i = 2$). In the case of polygon 964, the uncertainty is a probability of 0.1 that the area is not 'high' TOC (Figure 3D); however, the classification 'high' is an end-member and in this case 'medium-high' classification is assigned a probability 0.1 because $i = 1$ (Figure 3B). Polygons 963 and 968–971 are treated the same as in Scenario 1 because the original classification was 'unknown' TOC; thus each attribute is equally likely (i.e. $i = N$) (Figure 3B and D).

3.3. *Scenario 3: expert-defined uncertainty in pseudo-normal distribution*

Scenario 3 continues to focus only on Group A uncertainty like Scenarios 1 and 2. It uses expert-defined uncertainty like Scenario 2. However, unlike Scenario 2 where reclassification is limited, Scenario 3 distributes uncertainty among all of the classes in a pseudo-normal distribution.

Class numbers are defined $i = 1, ..., N$ with N the total number of classes. X is a discrete random variable representing the class of a feature. The probability of the originally mapped class $i = k$ is:

$$Pr(X = k) = u \tag{13}$$

The remaining probability $1-u$ is distributed over all the remaining classes $i \neq k$. The probability distribution was created with expert opinion to follow a smooth, bell-curve shape taken from Gaussian distribution sigma values ($1\sigma = 0.341$, $2\sigma = 0.136$, $3\sigma = 0.21$, $4\sigma = 0.001$); hence called 'pseudo-normal distribution' as the originally classified attribute will disrupt the normal bell curve (Figure 3C). The classes nearer to the originally assigned class will be more probable than those classes further away and all classes will have some probability of being the correctly assigned classification. In situations where the class is unknown, the probability distribution follows a normal distribution fitted to N.

Figure 3C shows the probability distribution of the geochemical subcriterion. For example, polygon 964 has a probability of 0.1 of being misclassified, which is divided among the classes in a pseudo-normal distribution. The polygons as well as the uncertainty are labeled in the map (Figure 3D). All uncertainty probability distributions can be found in Supplemental data.

3.4. *Scenario 4: expert-defined uncertainty over pseudo-normal distribution with fuzzy membership assignment and boundary transition zone*

Scenario 4 builds on the uncertainty modeling of Scenario 3 by using the same input data uncertainty and probability distributions. Scenario 4 additionally investigates the uncertainty effects of fuzzy membership assignment and a boundary transition zone of 10 km. This scenario combines Group A and Group B uncertainties. Fuzzy membership uncertainty is represented using a stochastic fuzzy membership value. The fuzzy membership values originally assigned to each feature are $j \in g$, $g = [0,1]$; j may be different for each feature. In order to represent uncertainty, each of these values is replaced by a stochastic variable j_Y, derived from the original fuzzy membership value as:

$$j_Y = j + Y \tag{14}$$

In Equation (14), Y is a Gaussian distributed random variable with a mean 0 and a standard deviation 0.1. In each Monte Carlo realization, a random variable is drawn from j_Y.

Input data sets were assigned a boundary transition zone of 10 km because of the scale of the map and cell size. This means that for every neighboring area, the actual boundary may be shifted up to 5 km into one area or another as the boundary transition zone is bisected by the original boundary. The boundary transition zone is randomly reassigned a class membership between the two neighbors (Figure 4), having equal probability of occurrence of 0.5. Boundary transition zones are not applicable to all input data sets (Table 3). Examples of exceptions are where the boundary is already certain or the original format is point data.
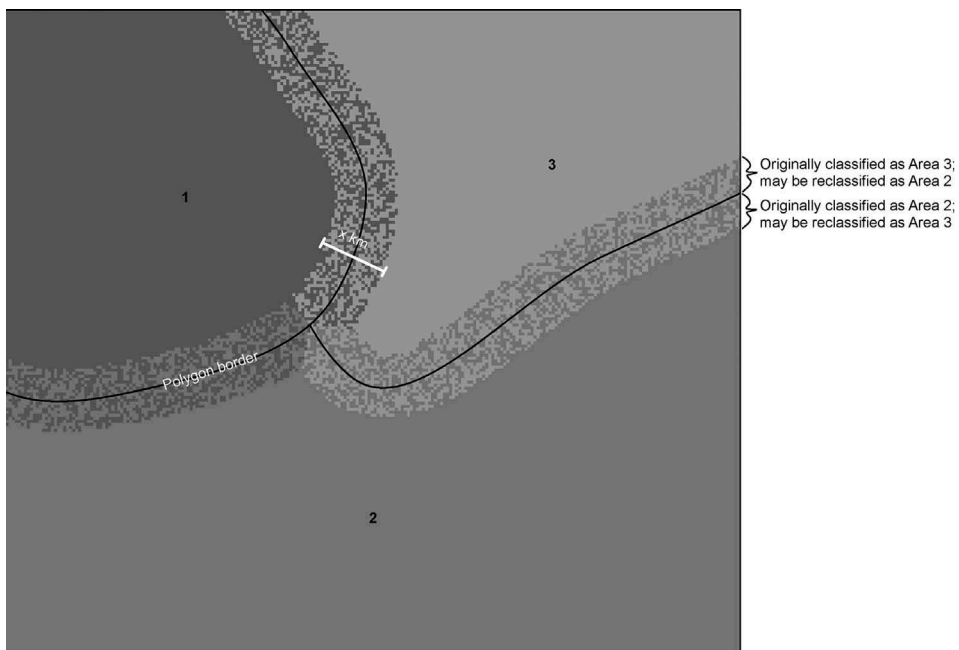

Figure 4. Boundary transition zone example. The polygon border may shift between neighboring areas. The cells within the transition zone are randomly reassigned.

Table 3. Error propagation parameters.

Subcriteria	Border classification	Fuzzy membership assignment
Costs (C)	Yes	Yes
Safety (S)	Yes	Yes
Risks (R)	No[+]	Yes
Geochemical (G)	No[‡]	Yes
Fields (F)	No[‡]	Yes
Wells (W)	No[‡]	Yes
Seeps (W)	No[‡]	Yes
Fault density (U)	No[‡]	Yes
Diapirs (A)	Yes	Yes
Isochrons (I)	No[†]	Yes
Plays (P)	Yes	Yes
Surface geology (Q)	Yes	Yes
Source rock (X)	Yes	Yes
Reservoir rock (V)	Yes	Yes
Traps (T)	Yes	Yes
Subsurface lithology (L)	Yes	Yes
Sedimentary basin (N)	Yes	No*

[+]Risks were not included in the border classification error because the boundaries follow the accepted political and economic zone boundaries.
[‡]These data sets were not included because the random error model was applied.
[†]The isochrons were not included because it is a continuous data set and has no borders.
*Sedimentary basins are not included in fuzzy membership assignment or error modeling because value is not uncertain; either the location is a sedimentary basin or it is not and this is well-established information.

3.5. Scenario 5: parameter investigation

Scenario 5 builds on Scenario 3 ignoring Group B uncertainty, but Scenario 5 introduces Group C uncertainty (model structure uncertainty). Scenario 5 changes all of the FLOs to the gamma operator. Gamma values are $q \in g$, $g = [0,1]$. q_z is a random variable representing the gamma value, q is the originally assigned gamma value, and Z is a Gaussian random variable with mean zero and standard deviation 0.2.

$$q_z = q + Z \qquad (15)$$

3.6. Scenario 6 – parameters with fuzzy membership uncertainty and boundary transition zone

Scenario 6 builds on Scenario 5 adding fuzzy membership uncertainty and a boundary transition zone. This scenario takes all sources of uncertainty of an MCE into account. It addresses groups A, B, and C uncertainties; thus, it applies Equations (13)–(15).

3.7. Confidence levels and thresholds

This article does not intend to give a complete review of statistical theory behind confidence levels and threshold values; however, it is deemed necessary for a brief overview and application in order that the reader is sure to understand the presentation of results and the discussion. Confidence levels are used when the most effective way to describe the error associated with an event cannot be precisely determined. A confidence level provides 'an interval and a probability that the unknown value falls within the interval' (Isaaks and Srivastava 1989, p. 494). Geoscience applies confidence levels on events that do not necessarily repeat over time (i.e. a well is only drilled once), but over
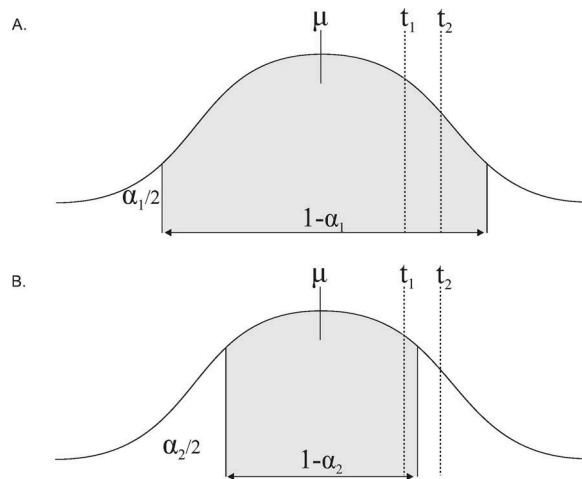
Figure 5. Confidence interval graphs with mean μ. Two thresholds are shown by dashed lines (t_1 and t_2). (A) Boundaries of confidence interval a_1 shown with solid black lines. (B) Boundaries of confidence interval a_2 shown with solid black lines.

space (i.e. multiple wells may be drilled in a prospect). However, it is further understood that the data collected (e.g. from wells) are skewed towards a bias of good outcomes (Isaaks and Srivastava 1989, Goovaerts 1997). The confidence intervals are calculated using the percentiles from the Monte Carlo simulation.

Figure 5 shows a graph illustrating two different confidence intervals. The area within the confidence interval is shaded gray. By changing the confidence level from $1-a_1$ to $1-a_2$, the confidence level is lower. In this study, 95% and 60% confidence levels are calculated and compared. A confidence level of 95% is used because of the conventional use of the interval as well as the lower risk associated with the outcomes; 5% risk is considered exceptional for petroleum exploration. Because the data used in the study is not industry-grade (i.e. data available for university research and public use) and to allow for comparison in a high-risk fashion, a confidence level of 60% is presented for comparison.

The concept of a threshold value is such that it is an indicator of a critical behavior (Isaaks and Srivastava 1989). Figure 5 shows two threshold values, t_1 and t_2. In the case of this study, the threshold value indicates the required favorability for petroleum exploration. The favorability value chosen to indicate possible exploration is 0.7 because the fuzzy membership values in Table 2 have this value as a medium-high qualitative description. A lower favorability value of 0.5 is presented for comparison based on the favorability values in Table 2, because the contextual relationship of 0.5 to 'medium' quality outcomes. It is assumed that less than 'medium' quality outcomes are not preferred for exploration decisions.

The confidence level is combined with the threshold value to indicate three different classifications of locations. Those areas that are (1) confident to be above the threshold value; (2) confident to be below the threshold value; or (3) ambiguous (i.e. it is unknown if the location is above or below the threshold and no decision can be made). Figure 5 illustrates how the confidence level and threshold value may restrict the results. The

interval for a level of $1-\alpha_1$ straddles t_1 and t_2 and is therefore ambiguous (Figure 5A); the interval of $1-\alpha_2$ straddles t_1 and is ambiguous (Figure 5B). However, $1-\alpha_2$ is completely below t_2, so it is confident to be below the threshold (Figure 5B).

4. Results

The median values of the realizations for all scenarios are shown in Figure 6. Areas in white are excluded from the results because the areas are not sedimentary basins, thus not appropriate for petroleum exploration activities. Scenario 3 (Figure 6C) has higher favorability in areas of known production in the Magdalena Valley basin, Llanos basin, Maracaibo basin, Guarico subbasin, Eastern Venezuelan basin, and Trinidad. Relatively high values are also shown in Aruba, Curaçao, and Bonaire, which are not producing areas, and offshore Guyana where there has been a recent oil discovery (Liza-1; Rosati and Carroll 2015) after several unsuccessful attempts of exploration. A comparatively low area is just northeast of Guajira Peninsula. All scenarios reflect similar spatial patterns of favorability trends, i.e. higher favorability onshore than offshore and higher favorability in the nine mentioned areas. Scenarios 1, 2, 3, and 5 have similar ranges of favorability from ~0.4 to ~0.8. Scenario 4 has favorability range from ~0.1 to ~0.8 and Scenario 6 has a range from ~0.0 to ~0.8. However, there are two identifiable ranges of values that appear more often – one ranging from ~0.0 to ~0.1 and a second ranging from ~0.5 to ~0.8 for both scenarios (4 and 6). Scenarios 4 and 6 include low favorability, likely dominating the lower value range histogram, in the boundary transition zone between the sedimentary and nonsedimentary basins.

The standard deviation of the final favorability Z was calculated for each scenario and is shown in Figure 7. Scenario 3 (Figure 7C) has an overall low standard deviation ranging from 0.02 to 0.09. There is a small area in Guyana near the border to Suriname that has a standard deviation ~0.3 because of the uncertainty related to various data sets overlapping there. Disregarding the nonsedimentary basins (shown with cross hatch marks), the lowest standard deviation is onshore and increases towards the offshore because of increased uncertainty; however, there is not much variation between the two regions. Scenarios 2 and 4 exhibit similar spatial trends and ranges of standard deviation to Scenario 3 with the exception of the border region in Scenario 4, which has high standard deviation due to the boundary transition zone. Scenario 1 has high standard deviation between 0.15 and 0.24 with higher values concentrated onshore and in areas of known hydrocarbon production. Scenarios 5 and 6 exhibit similar spatial trends to each other and have similar ranges from ~0.20 to ~0.40. These scenarios also show the greatest amount of spatial variation within the map results.

Following Scenario 3 through the changes in thresholds and confidence intervals (Figures 8–12), some general results can be related to the other five scenarios. A threshold value of 0.7 and a confidence interval with a confidence level of 95% results in 41% of the cells being either above or below the threshold; 59% of the cells cannot be used to support decisions (Figures 8C and 9). By narrowing the confidence level to 60%, the number of cells that can be used for decision-making increases to 67% (Figures 9 and 10C). By increasing the acceptable risk of making a wrong decision (i.e. by decreasing the confidence level of the confidence interval), the number of cells meeting the criteria increases. By lowering the threshold value to 0.5 and using a confidence level of 95%,
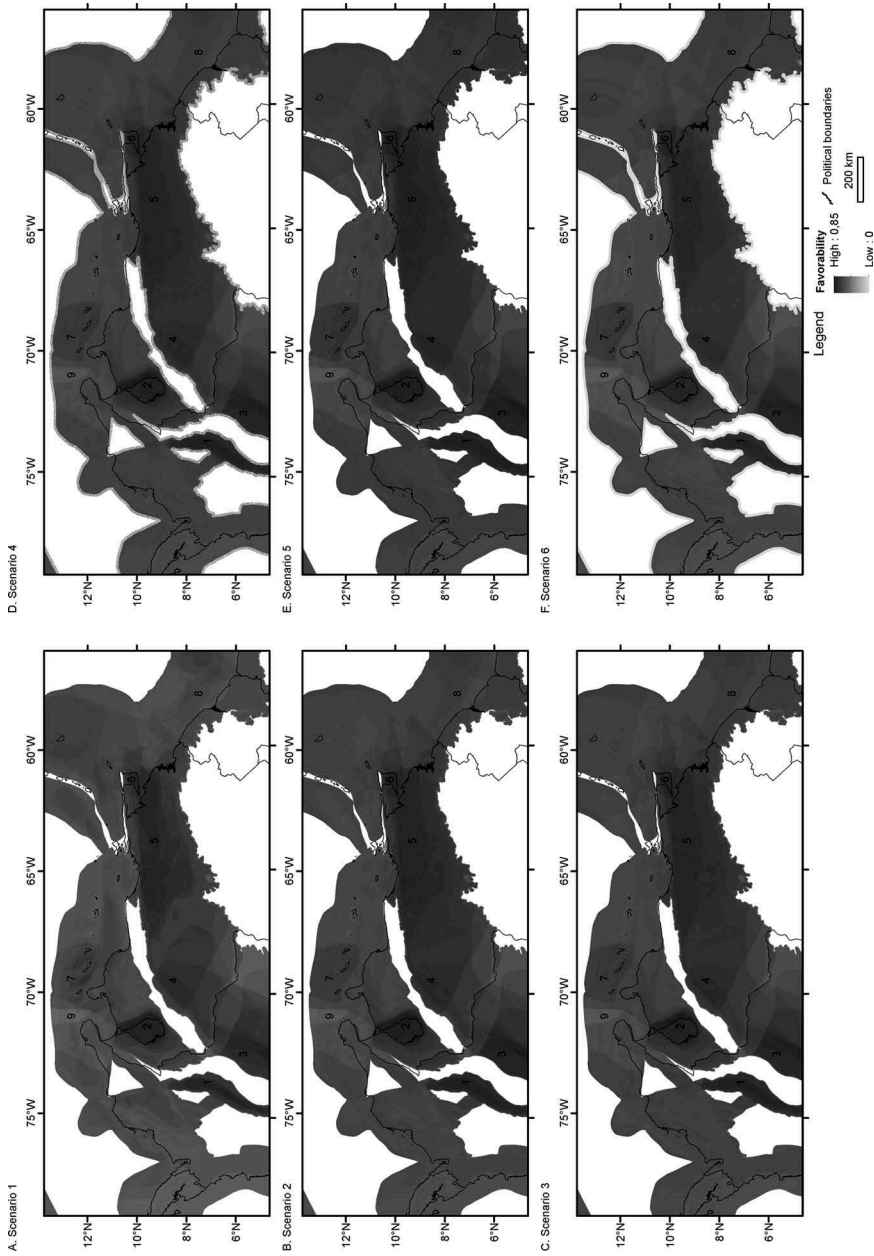
Figure 6. Median favorability of final favorability of all realizations. (A) Scenario 1. (B) Scenario 2. (C) Scenario 3. (D) Scenario 4. (E) Scenario 5. (F) Scenario 6. 1: Lower Magdalena Valley basin; 2: Maracaibo basin; 3: Llanos basin; 4: Guarico subbasin; 5: Eastern Venezuela basin; 6: Trinidad; 7: Aruba, Bonaire, and Curaçao; 8: Guyana; 9: northeast of Guajira Peninsula.
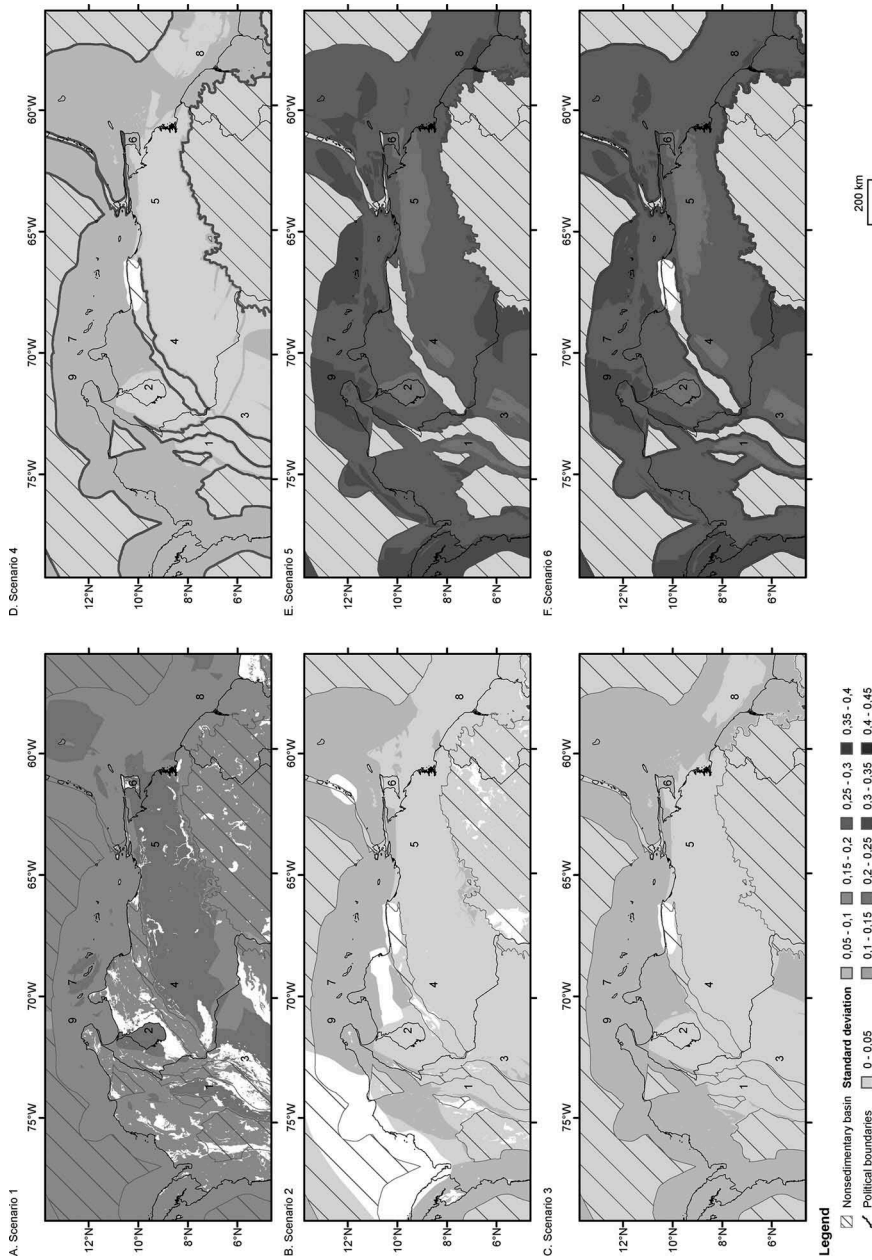
Figure 7. Standard deviation of final favorability of all realizations. (A) Scenario 1. (B) Scenario 2. (C) Scenario 3. (D) Scenario 4. (E) Scenario 5. (F) Scenario 6. 1: Lower Magdalena Valley basin; 2: Maracaibo basin; 3: Llanos basin; 4: Guarico subbasin; 5: Eastern Venezuela basin; 6: Trinidad; 7: Aruba, Bonaire, and Curaçao; 8: Guyana; 9: northeast of Guajira Peninsula.
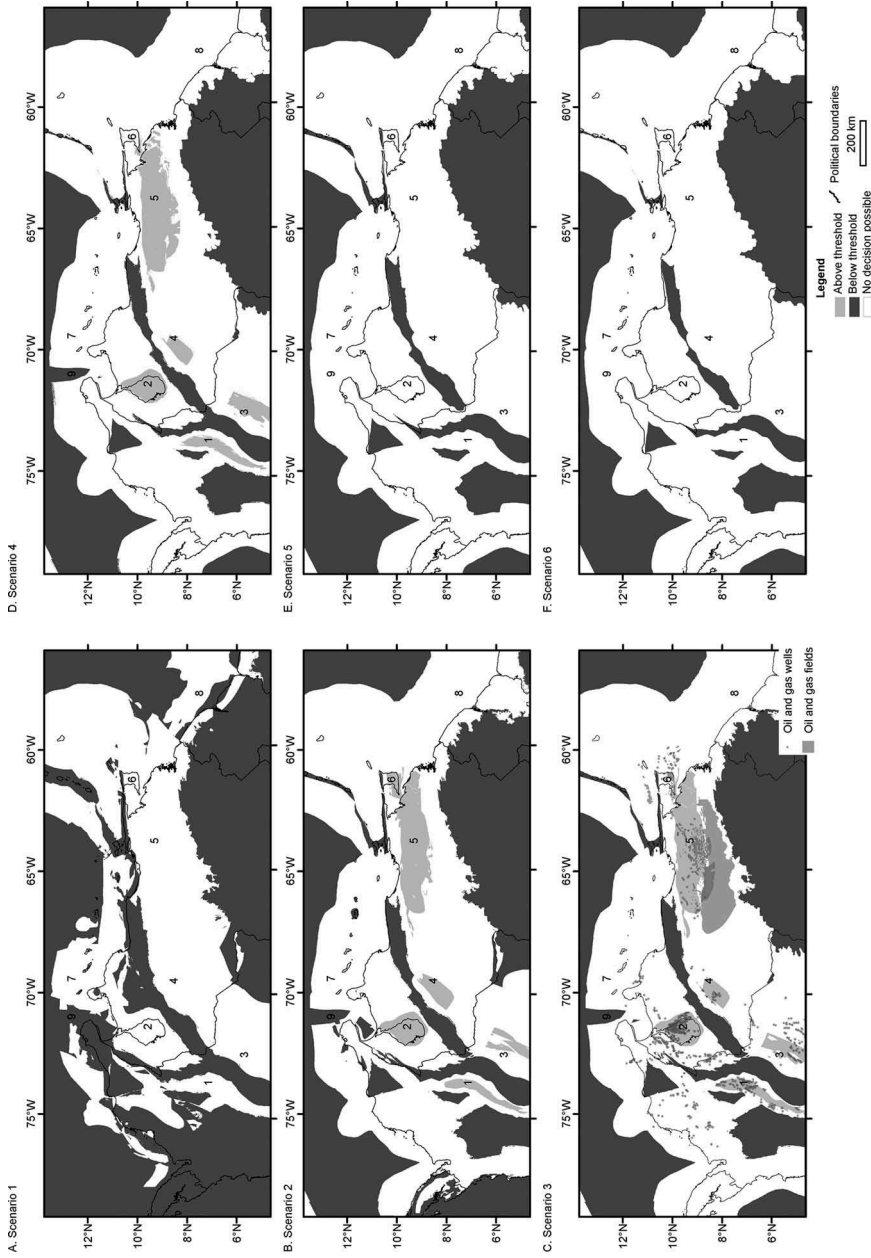
Figure 8. Threshold value 0.7 and confidence interval 95%. (A) Scenario 1. (B) Scenario 2. (C) Scenario 3. Medium gray areas are oil and gas wells and fields. (D) Scenario 4. (E) Scenario 5. (F) Scenario 6. White represents ambiguous (i.e. no decision possible); dark gray represents above the threshold; light gray represents below the threshold. 1: Lower Magdalena Valley basin; 2: Maracaibo basin; 3: Llanos basin; 4: Guarico subbasin; 5: Eastern Venezuela basin; 6: Trinidad; 7: Aruba, Bonaire, and Curaçao; 8: Guyana; 9: northeast of Guajira Peninsula.
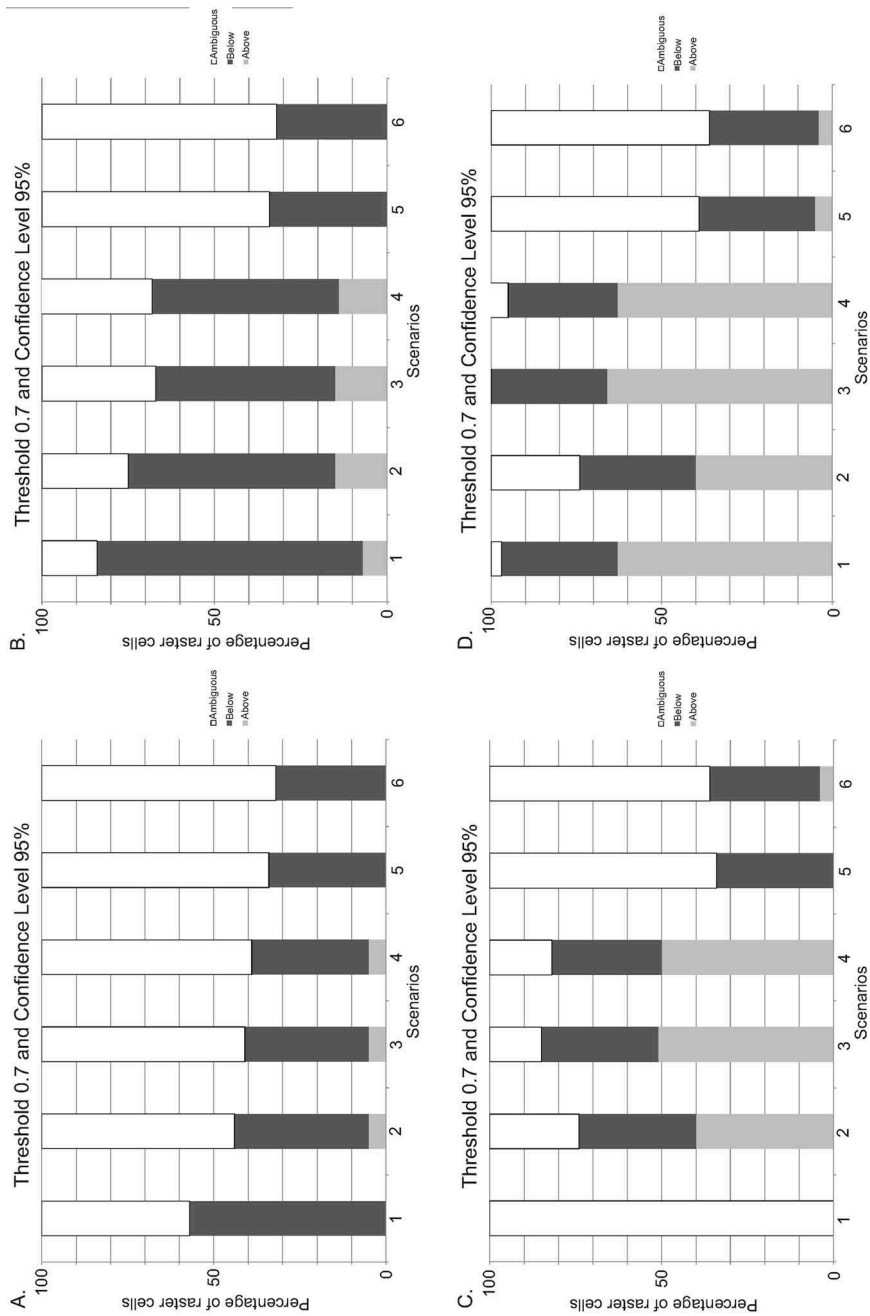
Figure 9. Graph summarizing all scenario results for threshold value and confidence level combinations. (A) Threshold value 0.7 and confidence level 95%. (B) Threshold value 0.7 and confidence level 60%. (C) Threshold value 0.5 and confidence level 95%. (D) Threshold value 0.5 and confidence level 60%. Light gray bars show the percentage of the map with the lower boundary of the confidence interval above the threshold value; dark gray bars show the percentage of the map with the upper boundary confidence interval below the threshold value; white bars show the percentage of the map where it is ambiguous (confidence interval straddles threshold).
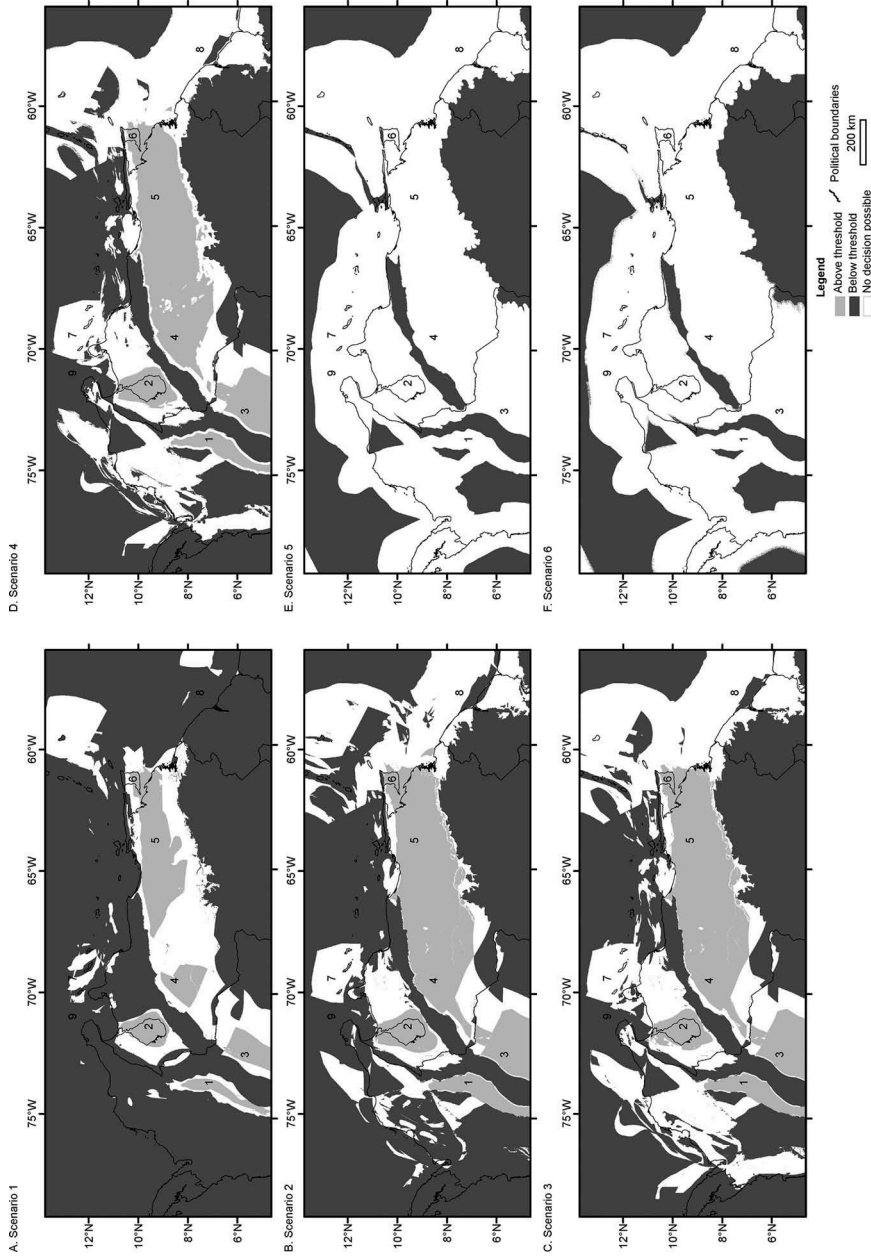
Figure 10. Threshold value 0.7 and confidence interval 60%. (A) Scenario 1. (B) Scenario 2. (C) Scenario 3. (D) Scenario 4. (E) Scenario 5. (F) Scenario 6. White represents ambiguous (i.e. no decision possible); dark gray represents above the threshold; light gray represents below the threshold. 1: Lower Magdalena Valley basin; 2: Maracaibo basin; 3: Llanos basin; 4: Guarico subbasin; 5: Eastern Venezuela basin; 6: Trinidad; 7: Aruba, Bonaire, and Curaçao; 8: Guyana; 9: northeast of Guajira Peninsula.

Figure 11. Threshold value 0.5 and confidence interval 95%. (A) Scenario 1. (B) Scenario 2. (C) Scenario 3. (D) Scenario 4. (E) Scenario 5. (F) Scenario 6. White represents ambiguous (i.e. no decision possible); dark gray represents above the threshold; light gray represents below the threshold. 1: Lower Magdalena Valley basin; 2: Maracaibo basin; 3: Llanos basin; 4: Guarico subbasin; 5: Eastern Venezuela basin; 6: Trinidad; 7: Aruba, Bonaire, and Curaçao; 8: Guyana; 9: northeast of Guajira Peninsula.
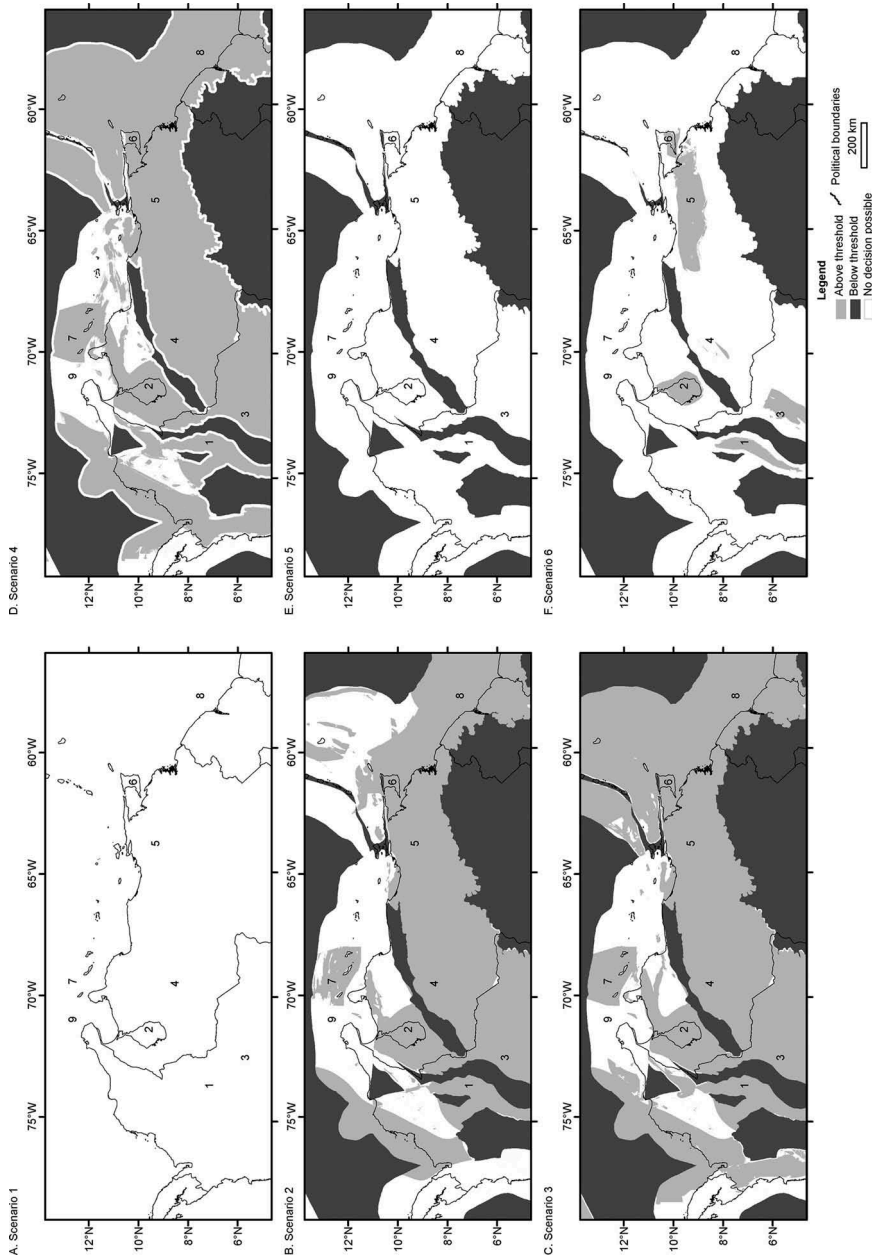
Figure 12. Threshold value 0.5 and confidence interval 60%. (A) Scenario 1. (B) Scenario 2. (C) Scenario 3. (D) Scenario 4. (E) Scenario 5. (F) Scenario 6. Medium gray areas are oil and gas wells and fields. White represents ambiguous (i.e. no decision possible); dark gray represents above the threshold; light gray represents below the threshold. 1: Lower Magdalena Valley basin; 2: Maracaibo basin; 3: Llanos basin; 4: Guarico subbasin; 5: Eastern Venezuela basin; 6: Trinidad; 7: Aruba, Bonaire, and Curaçao; 8: Guyana; 9: northeast of Guajira Peninsula.
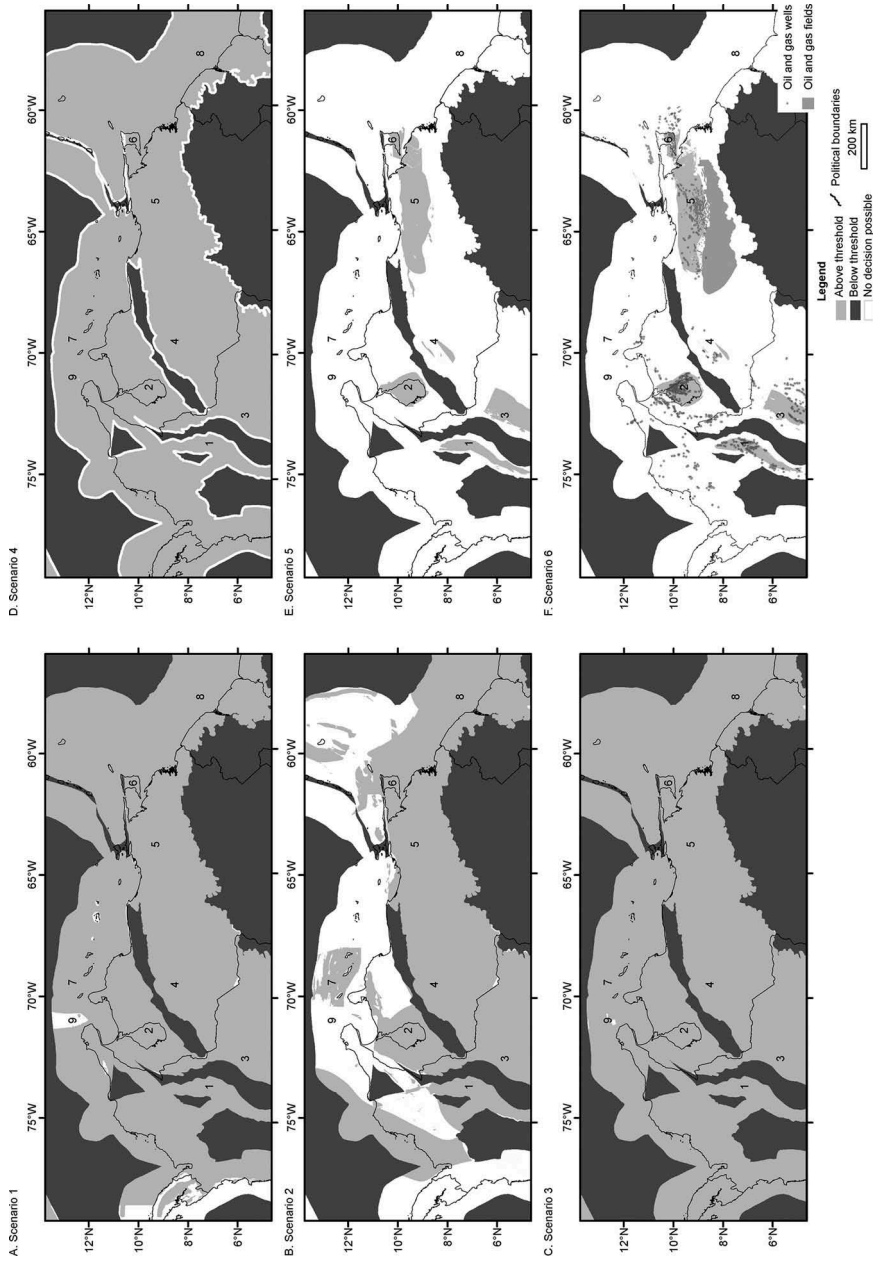
the total area available for decision-making increases; 85% of the cells are either above or below the threshold (Figures 9 and 11C). Accepting more risk by reducing the confidence level to 60%, the total area available for decision-making increases and virtually all of the cells are either above or below the threshold (Figures 9 and 12C). In general, all of the scenarios follow the same pattern: narrowed confidence level and lowered threshold value increases the total area available for decision-making (Figures 8–12).

5. Discussion

From the uncertainty applied to the data sets in this study, the results, in general, show that lowering the confidence level from 95% to 60% results in more cells for decision-making. Changing the threshold value from 0.7 to 0.5 results in more cells that are considered more favorable (i.e. number of cells above the threshold increase). With a threshold value of 0.7 and confidence level of 95% (Figures 8 and 9), regardless of the scenario, the majority of the results either do not support exploration or are inconclusive. In the case that accepts the most uncertainty in the results with threshold value of 0.5 and confidence level of 60%, more areas are above the threshold and support exploration than in previous cases (Figures 9 and 12). However, it is not necessarily good that all of an area could be explored because it is not probable that petroleum exists in the entire study area. In the same case, if all of the uncertainty sources are taken into account (Scenario 6), there are few areas where exploration is supported (Figures 9 and 12F). By overlaying known oil and gas wells and fields, it is shown that many existing production areas are in the ambiguous areas (Figure 12F), although new areas are also predicted. In many of the cases, depending on the sources of uncertainty taken into account, decisions are not possible. Because of this conundrum, it is difficult to choose only one case for making decisions. It would most likely suit the realistic situation that areas are further classified for decision-making and that all four cases are used for a specified scenario. In short, areas would be given a grade based on the threshold value and confidence interval.

Group A (uncertainty of observational data with respect to attribute classification) with expert-defined uncertainty in either Scenarios 2 or 3 yield overall similar spatial results with local variation and have lower standard deviation than Scenario 1 (Figure 7). Scenario 2 has more cells available for decision-making when the threshold value is 0.7 regardless of the confidence interval, and Scenario 3 has more cells available for decision-making when the threshold value is 0.5 regardless of the confidence interval (Table 9). Because Scenarios 2 and 3 yield similar results and rely on expert opinion, they are favored over Scenario 1, which produces very different results.

Scenario 4 combines groups A (uncertainty of observational data with respect to attribute classification) and B (uncertainty of weighting of evidence with respect to fuzzy membership assignment and observational data with respect to boundary transition zones) uncertainties and presents more uncertainty in its results as expected since more uncertainties are quantified. However, fuzzy membership and border transition uncertainty (Group B) contribute less to the output uncertainty than the uncertainty of the attribute classification in the input data sets (Group A). It is concluded that there is more uncertainty related to the input data set attribute classification rather than the fuzzy membership values for the defined uncertainties, because the addition of Group B

uncertainty has little impact on the results of Scenario 4 compared to Scenario 3. It is possible that differently defined uncertainties for either Group A or Group B could yield different results.

Group C uncertainty (model uncertainty of fuzzy logic operators) overshadows the other uncertainties defined in groups A or B. Scenarios 5 and 6 have similar spatial patterns and values both in the standard deviation and in threshold results. The results of Scenarios 5 and 6 verify that changes in the gamma value have large impacts on the results. This confirms the known subjectivity and uncertainty related to MCE models and reinforces the necessity to construct models based in logical grounding of the real-world phenomena and data evaluation of an expert. This is a sensitive aspect of the model.

The weaknesses in this approach are the subjectivity related to every aspect of the analysis: the input data, the MCE methodology, the fuzzy membership values, the defined uncertainties, the accepted threshold value, and the acceptable confidence level. All of these aspects of the analysis may change depending on the expert opinion or depending on the company policy that an exploration geologist would refer to for decisions. However, every aspect of the analysis would not necessarily change or change drastically.

When comparing known locations of oil and gas wells and fields (Figures 8C and 12F), it may be concluded that the uncertainties are underestimated for some subcriteria. The coverage of the oil and gas wells and fields data set is not complete; however, it can be used as a guide for validation. Some areas where petroleum exists coincide with the ambiguous zone. In Figure 8C, the confidence level of 95% means that 5% of the cells are incorrectly classified as either above or below the threshold value. In Figure 12F, 40% of the cells are incorrectly classified as being above or below the threshold. While not all of the existing wells and fields are shown to be in the confidence levels and above the threshold value, some validation is shown. Further work is needed therefore to better constrain the defined uncertainties in order to reflect better-known production information. Despite the weaknesses, statistical information is given that can be used to support decisions which would not be present otherwise. Additionally, the MCE and Monte Carlo simulation could be repeated for age-specific analysis and validation of producing fields may reflect better the reality of production information.

6. Conclusions

This article applies a Monte Carlo simulation to a fuzzy logic MCE to determine the best scenario for defining uncertainty in the input data and model structure. It addresses three groups of uncertainty sources (attribute classification – Scenarios 1–3; fuzzy membership assignment and boundary transition – Scenario 4; fuzzy logic operator – Scenario 5), combines the uncertainty sources in different ways (Scenarios 4–6), and proposes three methods for defining input data attribute classification uncertainty (Scenarios 1–3). It finds that similar results can be obtained in either of the two expert-defined uncertainty scenarios (Scenarios 2 and 3) and thus are the favored uncertainty definition approaches. Uncertainty from fuzzy membership assignment did not affect the results as much as attribute classification uncertainty (Scenario 4 vs. Scenarios 1–3). Changing fuzzy logic operators to gamma and changing the gamma values significantly changed the output, indicating that the choice of fuzzy logic

operator and gamma value must be chosen in a clear and logical manner (Scenario 5). Threshold values and confidence levels need to be predetermined prior to analyzing the results to prevent bias in the interpretation; however, the most useful interpretation may include several cases (i.e. multiple thresholds or multiple confidence levels) which grade the locations. Future work may better constrain uncertainties in order to match known production locations better and to focus on age-specific evaluations. By incorporating the uncertainties of the subcriteria and defining them within the provided MCE framework, the uncertainty analysis can be incorporated in forecasts used for decision-making.

Acknowledgements

Disclosure statement

References

An, P., Moon, W.M., and Bonham-Carter, G.F., 1994. Uncertainty management in integration of exploration data using the belief function. *Natural Resources Research*, 3, 60–71.

Beltran, C., 1993. *Mapa Neotectonico de Venezuela (1:2,000,000)*. FUNVISIS and Departamento de Ciencias de la Tierra.

Bingham, L., Zurita-Milla, R., and Escalona, A., 2012. Geographic information system–based fuzzy-logic analysis for petroleum exploration with a case study of northern South America. *AAPG Bulletin*, 96, 2121–2142. doi:10.1306/04251212009

Bonham-Carter, G.F., 1994. Geographic information systems for geoscientists modelling with GIS. *In*: D.F. Merriam, ed. *Computer methods in the geosciences*. Oxford: Pergamon, Vol. 13, 398 p.

Bonham-Carter, G.F., Agterberg, F.P., and Wright, D.F., 1988. Integration of geological datasets for gold exploration in Nova Scotia. *Photogrammetric Engineering and Remote Sensing*, 54, 1585–1592.

Brown K., and G.K. Westbrook, 1988, Mud diapirism and subcretion in the Barbados Ridge accretionary complex: the role of fluids in accretionary processes. *Tectonics*, 7, 613–640.

Carter, S., 1991. Site search and multicriteria evaluation. *Planning Outlook*, 34, 27–36.

Carver, S.J., 1991. Integrating multi-criteria evaluation with geographical information systems. *International Journal of Geographical Information Systems*, 5, 321–339. doi:10.1080/02693799108927858

Castellanos, H., *et al.*, 2006. Tectonic and stratigraphic controls on the Orinoco heavy oil belt, eastern Venezuela foreland basin. *AAPG annual meeting*, 9–12 April, Houston, TX.

Curet, E., 1992. Stratigraphy and evolution of the Tertiary Aruba basin. *Journal of Petroleum Geology*, 15, 283–304.

Davis, T.J. and Keller, C.P., 1997. Modelling uncertainty in natural resource analysis using fuzzy sets and Monte Carlo simulation: slope stability prediction. *International Journal of Geographical Information Science*, 11, 409–434. doi:10.1080/136588197242239

Deville, E., *et al.*, 2003. Lateral changes of frontal accretion and mud volcanism processes in the Barbados accretionary prism and some implications. *AAPG Bulletin*, 79, 121–124.

Di Croce, J., Bally, A.W., and Vail, P., 1999. Sequence stratigraphy of the Eastern Venezuelan Basin. *In*: P. Mann, ed. *Caribbean basins: sedimentary basins of the world*. Amsterdam: Elsevier Science B. V., Vol. 4, p. 419–476.

DM2 Project, 2005, *Deep marine depositional margins consortium GIS database*. Austin: The University of Texas at Austin.

Duerto, L., 2007. *Shale tectonics, eastern venezuelan basin*. London: Royal Holloway university of London, 424 p.

Exploration and Production Information Service (EPIS), and Agencia Nacional de Hidrocarburos (ANH), 2008, *Exploration and Production Information Service*. Bogata: EPIS and ANH.

Feizizadeh, B. and Blaschke, T., 2014. An uncertainty and sensitivity analysis approach for GIS-based multicriteria landslide susceptibility mapping. *International Journal of Geographical Information Science*, 28, 610–638. doi:10.1080/13658816.2013.869821

Fernández, D.S. and Lutz, M.A., 2010. Urban flood hazard zoning in Tucumán Province, Argentina, using GIS and multicriteria decision analysis. *Engineering Geology*, 111, 90–98. doi:10.1016/j.enggeo.2009.12.006

French, C., D. and C.J. Schenk, 2004, *Map showing geology, oil and gas fields, and geologic provinces of the Caribbean region* [online]. Open file report 97-470-K. Denver, CO: United States Geological Survey. Available from: http://pubs.usgs.gov/of/1997/ofr-97-470/OF97-470K/

Garrity, C. P., Hackley, P. C., and Urbani, F., 2006. *Digital Geologic Map and GIS Database of Venezuela - USGS data series 199* [online]. Denver, CO: United States Geological Survey. Available from: http://pubs.usgs.gov/ds/2006/199/

Geologisch Mijnbouwkundige Dienst, 1977. *Geological Map of Suriname (1:500,000)*: *Ministerie Van Opbouw*. Paramaribo: Ministerie Van Opbouw.

Goovaerts, P., 1997. *Geostatistics for natural resources evaluation*. Oxford: Oxford University Press, 483.

Heuvelink, G.B.M., 1998. *Error propagation in environmental modelling with GIS*. London: Taylor & Francis, 127.

Heywood, I., Cornelius, S., and Carver, S., 2006. *An introduction to geographical information systems*. Harlow: Pearson Prentice Hall, 426.

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

Isaaks, E.H. and Srivastava, R.M., 1989. *An introduction to applied geostatistics*. Oxford: Oxford University Press, 561.

Jiang, H. and Eastman, J.R., 2000. Application of fuzzy measures in multi-criteria evaluation in GIS. *International Journal of Geographical Information Science*, 14, 173–184. doi:10.1080/136588100240903

Karssenberg, D. and De Jong, K., 2005. Dynamic environmental modelling in GIS: 2. Modelling error propagation. *International Journal of Geographical Information Science*, 19, 623–637. doi:10.1080/13658810500104799

Kroehler, M., 2007. *Tectonics and sequence stratigraphy of the Venezuelan basin, Caribbean sea: master of science thesis*. Austin, TX: The University of Texas at Austin, 108.

Kugler, H. G., 1959. *Geological map of Trinidad and geological section through Trinidad (1:100,000)*. London: Orell Fussli S.A., Zurich and E. Stanford Ltd.

Li, D., Zhang, J., and Wu, H., 2012. Spatial data quality and beyond. *International Journal of Geographical Information Science*, 26, 2277–2290. doi:10.1080/13658816.2012.719625

Malczewski, J., 2006. GIS-based multicriteria decision analysis: a survey of the literature. *International Journal of Geographical Information Science*, 20, 703–726. doi:10.1080/13658810600661508

Moon, W.M., 1998. Integration and fusion of geological exploration data: a theoretical review of fuzzy logic approach. *Geosciences Journal*, 2, 175–183. doi:10.1007/BF02910163

Rosati, A. and Carroll, J., 2015, *Exxon's Guyana Oil Discovery may be 12 times Larger than Economy*. Rigzone. Available from: http://www.rigzone.com/news/oil_gas/a/138742/Liza_Discovery_Could_Put_Guyana_on_Oil_Gas_Map/?pgNum=0 [Accessed 1 October 2015].

Sanchez, R. J. P. y., 2007. *Evolucion geologica del sureste mexicano desde el Mexozoico al presente en el contexto regional del Golfo de Mexico*. Boletin de la Sociedad Geologica Mexicana, LIX, 19–42.

Saunders, J. and Snoke, A., 1998. *Geologic map of Trinidad and Tobago (1:100,000)*. Port-of-Spain: Ministry of Energy and Energy Industries, Government of the Republic of Trinidad.

Schenk, C. J., Viger, R. J., and Anderson, C. P., 1999. *Maps showing geology, oil and gas fields, and geologic provinces of South America* [online]. Open File Report 97-470D. Denver, CO: United States Geological Survey. Availbale from: http://pubs.usgs.gov/of/1997/ofr-97-470/OF97-470D/

Staatsolie, 2007. *Staatsolie offshore Suriname exploration presentation*. Paramaribo: Staatsolie, p. 17.

Sullivan, M., 2005. *Geochemistry, sedimentology, and morphology of mud volcanoes, eastern offshore Trinidad*. Austin, TX: University of Texas at Austin. 107 p.

Tangestani, M.H. and Moore, F., 2002. The use of Dempster–Shafer model and GIS in integration of geoscientific data for porphyry copper potential mapping, north of Shahr-e-Babak, Iran. *International Journal of Applied Earth Observation and Geoinformation*, 4, 65–74. doi:10.1016/S0303-2434(02)00008-9

Thapa, K. and Bossler, J., 1992. Accuracy of spatial data used in geographic information systems. *Photogrammetric Engineering and Remote Sensing*, 58, 835–841.

Tounsi, M., 2005. An approximate reasoning based technique for oil assessment. *Expert Systems with Applications*, 29, 485–491. doi:10.1016/j.eswa.2005.05.001

USGS, 1989. *Map projections*. Reston, VA: US Department of the Interior and US Geological Survey.

Valery, P., Nely, G. Mascle, A. Biju-Duval, B. Le Quellec, P. and Berthon, J. L., 1985. Structure et Croissance D'un prisme d'accretion tectonique proche d'un continent: la ride de la barbade au sud de l'arc antillais. *Carribean Geodynamics*, 27, 173–186.

Voogd, H., 1983. *Multicriteria evaluation for urban and regional planning*. London: Pion Limited, 367 p.

Walrond, G. W., 1987. *Geological map of Guyana (1:1,000,000)*. Georgetown, TX: Guyana Geology and Mines Commission.

Wood, M., 2006. *South America Oil and Gas Map (scale not provided)*. Edingburgh: Wood McKenzie Research.

Wright, D.F. and Bonham-Carter, G.F., 1996. VHMS favourability mapping with GIS-based integration models, Chisel Lake-Anderson Lake area. *In*: G.F. Bonham-Carter, A.G. Galley, and G.E.M. Hall, eds. *EXTECH I: a multidisciplinary approach to massive sulphide research in the Rusty Lake-Snow Lake Greenstone Belts, Manitoba: Geological Survey of Canada Bulletin 426*. Ottawa, Ontario: Natural Resources Canada, 339–376.

Yang, W. and Escalona, A., 2011. Tectono-stratigraphic evolution of Guyana Basin. *AAPG Bulletin*, 95, 1139–1369.

Ysaccis, R., 1997. Tertiary evolution of the northeastern Venezuela offshore. Ph.D. dissertation thesis. Houston, TX: Rice University, 285.

Zimmerman, H. and Zysno, P., 1980. Latent connectives in human decision making. *Fuzzy Sets and Systems*, 4, 37–51. doi:10.1016/0165-0114(80)90062-7