A hybrid GIS-assisted framework to integrate Dempster–Shafer theory of evidence and fuzzy sets in risk analysis: an application in hydrocarbon exploration

Sahand Seraj, Mahmoud Reza Delavar & Reza Rezaee

To cite this article: Sahand Seraj, Mahmoud Reza Delavar & Reza Rezaee (2019): A hybrid GIS-assisted framework to integrate Dempster–Shafer theory of evidence and fuzzy sets in risk analysis: an application in hydrocarbon exploration, Geocarto International, DOI: 10.1080/10106049.2019.1622602

To link to this article: https://doi.org/10.1080/10106049.2019.1622602

Accepted author version posted online: 09 Jun 2019.
Published online: 24 Jun 2019.

Submit your article to this journal

Article views: 10

View Crossmark data
A hybrid GIS-assisted framework to integrate Dempster–Shafer theory of evidence and fuzzy sets in risk analysis: an application in hydrocarbon exploration

Sahand Seraj\textsuperscript{a}, Mahmoud Reza Delavara and Reza Rezaee\textsuperscript{b}

\textsuperscript{a}GIS Department, School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran; \textsuperscript{b}Department of Petroleum Engineering, Curtin University, Perth, Australia

\textbf{ABSTRACT}
One of the most important reasons for the existence of geologic risk during the hydrocarbon exploration process is related to uncertainties in geospatial data and models employed for data fusion. This study proposes a geospatial information system-assisted approach integrated with soft computing methods to manage spatial uncertainties during the hydrocarbon exploration process. A framework was designed to illustrate the process of calculating the geologic risk interval of each hydrocarbon structure and its estimation of uncertainties. The model enhances the geologic risk analysis of a Dempster–Shafer data-driven method by a fuzzy logic approach. The resultant hybrid method showed high predictive power with the area under the success and predictive curves being 82.2 and 75.9\%, respectively. According to the results, the proposed hybrid method has improved the quality of risk analysis.

\textbf{ARTICLE HISTORY}
Received 29 March 2019
Accepted 3 May 2019

\textbf{KEYWORDS}
Geospatial information system; uncertainty; hydrocarbon exploration; Dempster–Shafer theory; fuzzy set

1. Introduction

Hydrocarbon exploration is a costly and time consuming process. Notwithstanding, this large investment, it has been referred to as a conventional situation for making decisions under risk and uncertainty conditions. The process of hydrocarbon exploration is categorized in three levels which are basin focus, play focus and prospect focus (Royal Dutch Shell 2014). It is admitted that specifying a play to drill is the most vital decision in the hydrocarbon exploration process (Seraj and Delavar 2018a, 2018b). During the process of play based hydrocarbon exploration, petroleum system is a vital concept. It embraces the entire necessary elements and processes which are required for the existence of the oil and gas accumulations (Magoon and Dow 1994). The indispensable components include the source, reservoir, seal and overburden rocks whereas the processes consist of trap formation and the generation–migration–accumulation of the petroleum (Figure 1).

Geologic risk occurs due to the uncertain areas in spatial data layers related to petroleum system components and uncertainty associated with the employed model for data fusion. In order to handle these uncertainties, a number of techniques such as weights of
evidence (e.g. Bonham-Carter 1994; Tangestani 2009; Agterberg 2011; Carranza 2015; Zhou et al. 2016; Hong et al. 2017a, 2017b), evidential belief functions (e.g. Park 2011; Mousavi 2012; Amiri et al. 2015; Ford et al. 2016; Chen et al. 2017; Seraj and Delavar 2018a, 2018b) and fuzzy logic methods (e.g. Nikravesh 2005; Nykanen et al. 2008; Lusty et al. 2012; Shokouh Saljoughi and Hezarkhani 2018; Xiong and Zuo 2018; Ziyong et al. 2018), have been applied.

Most real world multiple attribute decision analysis (MADA) problems, such as hydrocarbon exploration, involve different forms of uncertainties, which further complicate the decision analysis process. The solution for handling such uncertainties requires powerful soft computing methods that need to deal with attributes from various types of uncertainties. The integration of Dempster–Shafer theory (DST) and fuzzy set theory, which is proposed in this article, provides a pliable and principled framework to support hydrocarbon exploration as a MADA procedure under both aleatory and epistemic uncertainties. Aleatory uncertainty reflects our inability to predict random observable events, whereas epistemic uncertainty represents the analyst’s lacking knowledge about the values (Kiureghian and Ditlevsen 2009). It is seen that accessible data are deciphered in a probabilistic sense because the traditional probability theory is an extremely solid and entrenched mathematical tool that deals with aleatory uncertainty. To remove the limitations of the classical probabilistic method, such as definitions and one-value probability allocation instead of interval reasoning, several studies including Baudrit et al. (2007), recognized that it may be more appropriate to utilize a group of probability distributions to represent the incomplete and imprecise information related to a parameter instead of using a unique presumed probabilistic distribution. This concept can be represented by Dempster–Shafer functions within the paradigm of its theory (Dempster 1967; Shafer 1976).

Another paramount point in hydrocarbon exploration is that the geologic risk of each hydrocarbon structure would not be a crisp value and is a fuzzy range having lower and upper approximation borders due to inherent vagueness in hydrocarbon system parameters. In a crisp set, elements have a Boolean state of obedience that implies either membership or not. Integrated formation of source rock, reservoir rock, trap, seal and hydrocarbon migration, in a few kilometres underground, is not a case which could be handled by a 0 or 1 Boolean membership function. Therefore fuzzy-set theory, introduced by Zadeh (1965) was selected as an alternative to Boolean sets for geologic risk modeling of hydrocarbon structures. The advantage of fuzzy sets over Boolean sets for this research is related to the membership type of an object in fuzzy sets, which is a real number in the range of [0, 1] that is compatible with our purpose in petroleum system assessment.

In this study, the geologic risk assessment problem has been investigated using the DST and fuzzy logic method, which was employed to cover the uncertain factors in hydrocarbon exploration risk analysis. This included petroleum system parameters, such as the source rock, the reservoir rock as well as the trap and seal in a spatial framework. Dempster’s rule of combination (Dempster 1967; Shafer 1976) provides the capability of integrating maps of evidences in a more meaningful way for spatial data associated with hydrocarbon exploration. In addition, it assists with the filtering of unrelated or spurious evidences by assigning evidential weights to segments of a certain study area in which spatial data are lacking. It was able to model the uncertainty interval of data while facilitating the way in which we quantify and interpret the spatial associations that exist among the hydrocarbon deposits for every class of input spatial data. Furthermore, the method is able to integrate dependent assumptions of intuition sources, which are popular in spatial data. The main question that this research aims to answer was ‘what is the geologic risk value for each hydrocarbon structure and how much is the uncertainties related to these
values’. To answer this question, applying spatial data layers had an essential role. According to its definition, geospatial information system (GIS) ‘as a system designed to capture, store, manipulate, analyze, manage and present spatial data’ was used for spatial data fusion and producing final geologic risk maps in this research.

The golden point here is that geologic risk of hydrocarbon structures is fuzzy. It is determined by the proposed GIS-assisted hybrid model. The scenario is that the model calculates two maps that are consistent with the gained evidential belief functions and two fuzzy maps obtained from the fuzzification process, which are corresponding fuzzy membership values for the adjacency concepts by applying ArcGIS software. In the next step, in order to take advantage of both methods and their capabilities in handling uncertainty, the obtained values from the two methods were multiplied. The main idea is that the belief values obtained from the Dempster–Shafer method and Bad Geologic Chance (BGC) gained from the fuzzy logic approach have a similar pessimistic nature. Furthermore, plausibility and Good Geologic Chance (GGC) follow the same concepts. Finally, the uncertainty interval was estimated for each structure by subtracting the pessimistic from optimistic values.

This research proposes a GIS-assisted framework on the Dempster–Shafer and fuzzy set theory, to produce different risk maps related to each petroleum system parameters for anticipating the geologic risk associated with each structure. So the main contributions of the work can be categorized in three main parts. First of all, the entire petroleum system parameters are involved together in calculating geologic risk maps for each structure instead of surveying parameters separately. Secondly, the proposed model applied different soft computing methods to take advantages of all the exclusive capabilities of each of these methods by producing different geologic risk maps. The last contribution is related to applying GIS-assisted framework for geologic risk assessment in petroleum exploration process.

As the proposed model is not only valid for a unique area, it can be used for all other locations. The results of this research will help experts to explore hydrocarbon structures of sedimentary basins with significantly increased certainty. This would provide pessimistic and optimistic values related to the geologic chance of each hydrocarbon structure, which will be accompanied by an uncertainty value.

The remaining parts of the article are organized as follows. Section 2 describes the study area, Section 3 discusses the methods used and the proposed model and Section 4 provides data analyses and results. Section 5 provides a discussion and concluding remarks.

2. Study area

The Fars area situated in the Zagros fold belt in Iran was selected for hydrocarbon exploration, which is hypothesized to contain 102 hydrocarbon structures. This area covers roughly 15% of the world’s proven gas reserves (Motamedi et al. 2012). Based on the geologic classification, the study area encompasses Interior Fars, Sub Coastal Fars, Coastal Fars and Bandar–Abbas area (Figure 2).

Several studies have been investigating the petroleum system in the Fars embayment during the last few years. The tectonic and stratigraphic characteristics of Zagros and Makran during the Mesozoic–Cenozoic era were scrutinized by Leturmy and Robin (2010) in order to create a program demonstrating the evolution of the basins in the Middle East. Moreover, the sedimentary evolution of Fars region was investigated by Piryaei et al. (2011), who examined the marked changes in fades and thicknesses of the
upper cretaceous succession in the Bandar Abbas area located in the south of Iran. Furthermore, the petroleum systems in addition to the distribution of the oil and gas fields in the Iranian part of the Tethyan Region have been described by Bordenave (2014). Other studies that have revealed invaluable information concerning the petroleum system in Fars Region include the ones conducted by (Motamedi et al. 2012) and
Tavakoli-Shirazi et al. (2013). A play-based investigation of hydrocarbon system and its associated uncertainties was undertaken by Seraj and Delavar (2018a, 2018b).

3. Methods

The foundation of this research is established on the basis of the geologic chance of success in hydrocarbon exploration. The risk that occurs during the process of finding hydrocarbons is known as the geologic risk, which is the risk associated with the accumulation of a producible hydrocarbon. A producible accumulation needs to be a stabilized flow of hydrocarbons. The geologic risk is evaluated by taking into account the probability that the subsequent four independent petroleum system factors as source rock, reservoir rock, trap rock and dynamic conditions of hydrocarbon structures can be modeled by Rose formula (Rose 2000):

\[
P_g = \frac{P_{\text{source}}}{C^3} \frac{P_{\text{reservoir}}}{C^3} \frac{P_{\text{trap}}}{C^3} \frac{P_{\text{dynamics}}}{C^3}
\]

where \(P_{\text{source}}\) represents the probability of the hydrocarbon generation from organic materials that are cooked over time under proper temperature and pressure conditions, which can be ultimately converted to oil or gas. \(P_{\text{reservoir}}\) expresses the possibility of a subsurface pool of hydrocarbons being contained in porous or fractured rock formations. The probability of the configuration of rocks being suitable for containing hydrocarbons and sealed by a relatively impermeable formation through which hydrocarbons cannot migrate is described as \(P_{\text{trap}}\). Finally, \(P_{\text{dynamics}}\) illustrates the likelihood of the movement of hydrocarbons from their source into reservoir rocks. In order to display the geographical distribution associated with existing chance for each of the above-mentioned factors, several evidence maps produced in a GIS environment. Moreover, every element encompasses diverse parameters that are inevitably associated with substantial uncertainties. To handle these uncertainties, some soft computing methods were applied in this research. The main methods used in this study are briefly described in Sections 3.1 and 3.2.

3.1. GIS-assisted evidential theory

The theory of belief functions is known as an evidence theory or DST (Dempster 1967; Shafer 1976), which is a mathematical theory of evidence that is employed to integrate discrete pieces of information (evidences) to estimate the probability of an incident. Furthermore, it is asserted that it is an approach that lends its foundation to the Bayes’ rule, which combines data with the aim of predicting the occurrence of events (Armas 2012). This method takes advantage of the concept of prior or unconditional probability and posterior or conditional probability. Its functions are comprised of degree of belief (Bel), degree of disbelief (Dis), degree of uncertainty (Unc) and degree of plausibility (Pls) in the range of [0, 1], (Althuwaynee et al. 2014). When compared to other data-driven methods, the functions yield outputs that contain complementary information as disbeliefs, cautionary information as plausibility and relevant uncertainties. These outputs can each be separately mapped in the GIS environment although only two quantities are independent. This contrasts with the fuzzy logic output that consists only of a single map, which is also known as the combined fuzzy membership (Tangestani 2009).

Dempster’s combination rule has a substantial role in the process of information fusion in the DST. In order to employ the evidence theory for mineral prospectivity analysis in a geospatial environment, a novel method was recommended by Carranza and Hale (2003).
A high correlation is observed between mineral exploration and hydrocarbon exploration, which is chosen as the basis of the proposed model in this article. The method reveals that in the case of dividing an area into unit cells (pixels) with a fixed size (s) and a total area (t), \( N\{T\} = (t)/(s) \) is the total number of pixels in the study area. As such, if there are several pixels, \( N\{D\} \), with an occurrence of D, the prior probability of an occurrence is estimated as Equation 2 (Carranza and Hale 2003; Carranza et al. 2008):

\[
P\{D\} = \frac{N\{D\}}{N\{T\}}
\]

It is presumed that a binary predictor pattern B occupying \( N\{B\} \) pixels arises in the area and that numerous known boreholes arise within this pattern. Essentially, this means that \( N\{D\cap B\} \). As such, the conditional probability can be used to express the favorability of the location of an occurrence given the presence of a predictor pattern.

It is assumed that a set of petroleum system factors are employed in the hydrocarbon exploration process as \( \theta = \{A_1, A_2, \ldots, A_n\} \) where \( \theta \) describes the frame of discernment. Suppose a function is \( m: 2\theta \rightarrow [0, 1] \) where \( 2\theta \) is the set of all subsets of \( \theta \), including the empty set and \( \theta \) itself. This function is called a basic probability assignment when it satisfies \( m(\Phi) = 0 \) and \( \sum_{A \subseteq \theta} m(A) = 1 \), where \( \Phi \) is an empty set and \( A \) is a subset of \( \theta \). Furthermore, \( m \) is also referred to as a mass function and \( m(A) \) measures the extent to which the evidence supports \( A \). Finally, this is denoted Bel \( (A) \), which is a belief function (Tien Bui et al. 2012).

The difference between PIs and Bel is uncertainty (Unc) and the plausibility is equal to \( 1 - \text{Dis} \). Consequently, the sum Bel + Dis + Unc = 1 (Carranza and Hale 2003; Carranza et al. 2008).

Because an ordinary kriging was required to convert the discrete vector data related to each petroleum system parameter to a continuous raster form, the extended formulas of Dempster–Shafer in our model are raster-based (Seraj and Delavar 2018a). To explain these extended formulas, suppose that the maps of \( A_i \) \( (i = 1, 2, \ldots, n) \) are provided, which form the spatial evidence for petroleum system, each with a variable number of \( C_{ij} \) \( (j = 1, 2, \ldots, m) \) classes. Moreover, it is assumed that in the study area, \( N\{T\} \) denotes the total number of pixels, \( N\{P\} \) denotes the total number of hydrocarbon pixels, \( N\{C_{ij}\} \) stands for the number of pixels in the variable class \( C_{ij} \) and \( N\{C_{ij} \cap P\} \) is the number of hydrocarbon pixels in the class of petroleum system parameter \( C_{ij} \). The degree of belief for the class \( C_{ij} \) (Bel\( C_{ij} \)) is calculated as Equations 3 and 4 (Carranza et al. 2008; Carranza 2009):

\[
\text{Bel}_{C_{ij}} = \frac{W_{C_{ij}P}}{\sum_{j=1}^{m} W_{C_{ij}P}}
\]

where

\[
W_{C_{ij}P} = \frac{N\{C_{ij} \cap P\}}{N\{C_{ij}\}}
\]

The degree of disbelief for the class \( C_{ij} \) (Dis \( C_{ij} \)) is calculated as Equations 5 and 6 (Carranza et al. 2008; Carranza 2009):

\[
\text{Dis}_{C_{ij}} = \frac{W_{C_{ij}P}}{\sum_{j=1}^{m} W_{C_{ij}P}}
\]

where
As we now have two maps of A and B with the evidential belief functions for each, Dempster’s rule of combination for calculating the combined belief, disbelief, plausibility and uncertainty would be needed. Dempster’s rule of combination is applied by employing either the AND or the OR operations with the intention of calculating the combined belief functions for the two evidences (Carranza and Hale 2003). The list of the formulas needed to combine evidential belief functions of the two spatial layers by the AND operation is provided in Equations 7 and 8 (Carranza et al. 2008):

\[
\text{Bel}_{A_1A_2} = \frac{\text{Bel}_{A_1}\text{Bel}_{A_2}}{1 - \text{Bel}_{A_1}\text{Dis}_{A_2} - \text{Dis}_{A_1}\text{Bel}_{A_2}} \tag{7}
\]

\[
\text{Dis}_{A_1A_2} = \frac{\text{Dis}_{A_1}\text{Dis}_{A_2}}{1 - \text{Bel}_{A_1}\text{Dis}_{A_2} - \text{Dis}_{A_1}\text{Bel}_{A_2}} \tag{8}
\]

Finally, the integrated evidential belief functions of the four petroleum system parameters were combined by means of the AND operation in order to represent the integrated spatial evidence of ‘hydrocarbon potential’ in the final map. The AND operation was employed due to the fact that all the pieces of evidences were required to support the proposition.

### 3.2. GIS-assisted fuzzy logic

Membership of an object in the general set theory is a crisp concept. Specific functions would be able to assign the objects of a defined universe to a certain subset of this universe. In such a crisp set, an element can plainly belong to a set or not. If this membership comes with uncertainty, it could be represented by its membership function. These values can range between 0 and 1, where a value of 0 represents full non-membership and a value of 1 represents full membership.

As described by Jamshidi et al. (2016), a fuzzy logic-based model is comprised of the following feedforward modules: a fuzzifier (encoder); an inference engine (processor) and a defuzzifier (decoder). A fuzzifier has the function of encoding input crisp values into output fuzzy values. Fuzzification is the most decisive process in fuzzy modelling since these fuzzy values propagate through the model and finally characterize the output. A membership function is required for handling the fuzzification procedure. There are two sources for defining the membership function: prior knowledge of a system and using input data.

Fuzzification is carried out by a membership function, which can be derived either from a priori knowledge of a system or by using input data.

A fuzzy set A of a universe X is defined by a membership function \( \mu_A \) such that \( \mu_A : X \rightarrow [0, 1] \) where \( \mu_A(x) \) is the membership value of \( x \) in A. The universe X is always a crisp set.

If the universe is a finite set \( X = \{x_1, x_2, \ldots, x_n\} \), a fuzzy set A on X is expressed as Equation 9 (Tsoukalas and Uhrig 1997):

\[
A = \sum_{i=1}^{n} \mu_A(x_i)/x_i \tag{9}
\]

If the universe is an infinite set \( X = \{x_1, x_2, \ldots, x_n\} \), a fuzzy set A on X is expressed as
Equation 10 (Tsoukalas and Uhrig 1997):

\[ A = \int_{x} \mu_{A}(x)/x \]  

Fuzzy logic model can utilize any information from any estimation scale and weighting of evidence which is completely controlled by the expert. The fuzzy logic model takes into account the more adaptable fusion of factor maps and could be promptly executed within a GIS environment (Biswajeet Saro and Manfred 2009). A fuzzy membership value can be assigned either by experts or through a fuzzification process (Tsoukalas and Uhrig 1997; Burrough and McDonnell 2015). In this study, fuzzy membership values were calculated through a fuzzification process. In terms of human thinking, uncertain or vague concepts are frequently used. In real life applications, such hydrocarbon exploration, we might search for a suitable point in a hydrocarbon structure for drilling. An appropriate location must include the following criteria:

- be a moderate source of hydrocarbon
- have favourable reservoir rocks
- should not be near a fault
- have suitable traps
- be located in a closure
- have a favourable seal rock

All the above-mentioned conditions are vague related to the existence of a petroleum system in the area. Using the first law of geography (Tobler 1970), which declares that ‘everything is related to everything else, but near things are more related than distant things’, the adjacency concept related to topological spatial relations (Egenhofer and Franzosa 1991) for adjacent structures with a common petroleum system was applied in a GIS environment to compute two fuzzy membership values for each hydrocarbon structure, which were BGC and GGC. The BGC and GGC represent the pessimistic and optimistic values for the geologic chance of hydrocarbon exploration, respectively.

3.3. The proposed method

A framework was designed to illustrate the process of handling geologic risk analysis and estimation of uncertainties related to hydrocarbon exploration. In this section the different phases of the proposed model (Figure 3) are described.

During the first step, Dempster–Shafer functions were applied to handle the uncertainty. Two maps were produced for each of the petroleum system parameters, which were consistent with the gained evidential belief functions in Fars sedimentary region. The proposition evaluated by belief and plausibility maps with the help of the DST is that ‘a proper petroleum system exists for drilling the structure’. This proposition is appraised in the following way. The belief map represents a conservative support of proposition and the degree to which the evidence provides concrete support for the proposition. The plausibility map shows an optimistic estimate of the belief and the degree to which the evidence does not refute the proposition. Thus, the Rose formula (Equation 1) (Rose 2000) is modified and is represented as follows:

\[ \text{Belief (g)} = \text{Belief (source)} \ast \text{Belief (reservoir)} \ast \text{Belief (trap)} \ast \text{Belief (dynamics)} \] (11)
Plausibility \( g \) = \text{Plausibility (source)} \times \text{Plausibility (reservoir)} \times \text{Plausibility (trap)}
\times \text{Plausibility (dynamics)}

\hspace{1cm} (12)

However, the ‘\(*\)’ sign in Equation (11) and Equation (12) is not a simple multiplication sign and it refers to the Dempster–Shafer combination rule and applies an AND operator, which is customized by Equations 2–8.

During the next step, the fuzzy set capabilities were utilized and a fuzzifier was defined by Thiessen Polygon analysis in GIS. It was proved that the petroleum system in Fars area was spatially clustered, which was substantiated by assessing Moran’s I index as calculated by the spatial autocorrelation tool applying GIS for the sample drilled wells (Figure 4). Its \( P \) value was roughly 0.09 and the \( Z \) score was around 1.7, implying that the probability of the incidence of the petroleum system being spatially and randomly distributed was less than 10%. Consequently, consistent with the popular petroleum systems, the adjacent structures would probably have petroleum.

Consequently, in the Fars domain, it can be determined if a structure contains hydrocarbons with a high or low geologic chance through the findings of the hydrocarbon exploration in the adjacent structures.

The adjacency concept explained earlier could be modelled through the fuzzy sets employed as the decision attribute at two levels, namely BGC and GGC. The fuzzy membership could be specified as a scenario that is detailed as follows.

Primarily, the geologic structures in the study area could be categorized into two namely the ones with and without hydrocarbons. A new attribute field ‘Adjacent Structures’ would be created for every structure by making use of the Thiessen Polygon analysis in the GIS environment (Figure 5), which counts the number of the adjacent structures with hydrocarbons. For every likely value of the ‘Adjacent Structures’, the quantity of the structures that contain hydrocarbons could be estimated. The quantity was utilized to signify the corresponding fuzzy membership.
Figure 4. Spatial distribution of petroleum system in the study area.

Figure 5. Thiessen polygons for calculating the fuzzy membership function.
During Steps 3 and 4 (Figure 3), in order to take advantage of both of the DST and fuzzy methods in terms of their capabilities in handling uncertainty, the obtained values from the two methods were multiplied. The main idea is that the belief values obtained from Dempster–Shafer method and BGC gained from fuzzy logic approach have a similar pessimistic nature. Plausibility and GGC also follow the same concepts.

During the last step, the uncertainty interval was estimated for each structure by subtracting the values obtained in step 4 from those from step 3.

Result validation is the next step, which includes a comparison between the values gained from the proposed model and the reality of the exploration well that was drilled on each structure. This comparison can be qualitative by an overlay operation, or quantitative, which is performed using functions such as the Area Under Curve (AUC) (Remondo et al. 2004). AUC represents the area under the curve related to the classification accuracy of the proposed model. A higher AUC value indicates better model functionality. We applied success rate curve (SRC) and prediction rate curve (PRC) to measure the performance of our conceptual model. The SRC is based on the comparison between the predicted values and real values used in the modelling (Remondo et al. 2004). Furthermore, the success rate method can help to determine how well the results have classified the exploration area. The prediction rate method performs the comparison by partitioning the data related to the hydrocarbon structures; training data that are used for obtaining a prediction image; and test data that are compared with the prediction results for validation (Seraj and Delavar 2018a). This clarifies the accuracy to which the proposed model predicts the location of hydrocarbon accumulation.

4. Data analysis and results

A new fusion of the DST and fuzzy sets for risk analysis in hydrocarbon exploration was proposed and implemented in this research on a real dataset of 102 exploration wells in the Fars sedimentary region. Table 1 tabulates the multiple datasets that were employed. Having conducted the pre-processing, the entire thematic data sets were compiled into a raster-based database with a resolution of 250 m. All the required spatial data sets were projected to Universal Transverse Mercator (UTM) coordinate system (Zone 39N, Datum WGS_1984).

The Dempster–Shafer functions in the GIS environment generate two maps as belief and plausibility intervals for each petroleum system. The values of functions in these maps are continuous. So we produced 8 maps (4 factor /C3 maps) totally as illustrated in Figure 6.

The final belief and plausibility maps (Figure 7) were produced by applying modified Dempster–Shafer combination using Equations 2 to 9 on the four belief factor maps and four plausibility factor maps separately. Figure 7 reports the calculated quantity of the

<table>
<thead>
<tr>
<th>Petroleum System</th>
<th>Factor</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source</td>
<td>Total organic carbon</td>
<td>(1:100,000)</td>
</tr>
<tr>
<td></td>
<td>Potential of hydrocarbon generation</td>
<td>(1:100,000)</td>
</tr>
<tr>
<td></td>
<td>Maturity</td>
<td>(1:100,000)</td>
</tr>
<tr>
<td></td>
<td>Porosity</td>
<td>(1:250,000)</td>
</tr>
<tr>
<td></td>
<td>Permeability</td>
<td>(1:250,000)</td>
</tr>
<tr>
<td></td>
<td>Thickness of trap formation</td>
<td>(1:250,000)</td>
</tr>
<tr>
<td></td>
<td>Geologic formations map</td>
<td>(1:250,000)</td>
</tr>
<tr>
<td></td>
<td>Anticline axes map</td>
<td>(1:250,000)</td>
</tr>
<tr>
<td></td>
<td>Hydrodynamic flow map</td>
<td>(1:250,000)</td>
</tr>
</tbody>
</table>

| Reservoir        |                                          |              |
| Trap             |                                          |              |
| Dynamics         |                                          |              |
Figure 6. Evidence maps for (a) source (b) reservoir (c) trap (d) dynamics.
belief and plausibility of the petroleum system in Fars Domain for each pixel in the final map. The experts would be able to employ these maps to recognize the range of uncertainties for the presence of hydrocarbons at each point of the area. The final maps (Figure 7) were produced with the fusion of the entire factor maps, which is consistent with the gained Dempster–Shafer combination rule in Fars sedimentary region.

The BGC and GGC as fuzzy membership function values were calculated with the help of Thiessen polygons (Figure 8). According to the results from a Thiessen polygon analysis, we supposed that there were 12 adjacent structures for a specified structure. Among these 12 structures, there were 8 structures with oil wells and 4 structures with dry wells. Therefore, the membership function corresponding to BGC is 0.33 and that corresponding to GGC is 0.67.

An information table (Table 2) was compiled in order to show the results from the proposed model. In Table 2, the first two columns were obtained from the produced maps of Dempster–Shafer’s combination rule. The next two columns represent the fuzzy membership values calculated by the GIS analysis. The pessimistic and optimistic fields were added to the table because there are common concepts between belief and BGC in terms of pessimistic notions in addition to common concepts between plausibility and
GGC from an optimistic point of view. The values of pessimistic field were calculated by multiplication of belief in BGC. Optimistic field values were appraised by multiplication of plausibility in GGC. The last column shows the uncertainty related to hydrocarbon exploration risk for each structure and their values are obtained from the subtraction of the two upper and lower approximations, which are optimistic and pessimistic fields.

The results described that approximately 34.9% of the study area had a high risk for hydrocarbon exploration, which was able to predict beyond the 82.2% for the training and 75.9% for the testing data, respectively (Figure 9).

The results of the proposed model were compared from three aspects of the implemented methods. They are depicted in Figure 10 as the Dempster–Shafer, fuzzy logic and hybrid methods. It illustrated that each method puts a different number of hydrocarbon structure in the geologic risk classes. It was found that the hybrid method had more accurate results than two others according to post drill analysis of exploration wells in the study area.

### 5. Discussion

Finding oil and gas reservoir is a complicated and costly procedure. Apart from requiring a large investment, hydrocarbon exploration has been considered as a risky process because it includes several parameters with uncertainties. Most of these parameters have spatial properties and consequently their uncertainties could be modeled in a spatial framework like GIS.

It is verified that there are some gaps in previous researches which have been considered in this study. Primarily, it was evident that previous studies have not examined the uncertainty handling in the hydrocarbon exploration while most of them have considered limited number of major petroleum system parameters in the process. This implies that the mentioned literature dealt with only one element of petroleum system, such as reservoir characterization. The other major difference between this study and previous research is attributed to concurrent integration of the Dempster–Shafer and fuzzy sets methods with reference to risk analysis in hydrocarbon exploration process. This means that combining the two soft computing methods has enhanced the quality of our results related to exploring hydrocarbons in a sedimentary basin which leads to lower uncertainty. The next fundamental contribution of this research is the application of a novel GIS-assisted model in the hydrocarbon exploration process. The advantages of this method can be observed by experts when they look at four values namely belief, plausibility, BGC and GGC, in the final maps instead of just one value for the probability of success for each point. This capability allows for more reliable decisions to be made in a GIS environment, which is more flexible because it employs uncertainty intervals. In the GIS-assisted model

### Table 2. Geologic chance values of hydrocarbon structures, gained by applying the proposed method.

<table>
<thead>
<tr>
<th>Structure</th>
<th>Belief</th>
<th>Plausibility</th>
<th>BGC</th>
<th>GGC</th>
<th>Pessimistic</th>
<th>Optimistic</th>
<th>Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bamu</td>
<td>0.31</td>
<td>0.87</td>
<td>0.72</td>
<td>0.28</td>
<td>0.22</td>
<td>0.24</td>
<td>0.02</td>
</tr>
<tr>
<td>Gisakan</td>
<td>0.26</td>
<td>0.85</td>
<td>0.69</td>
<td>0.31</td>
<td>0.18</td>
<td>0.26</td>
<td>0.08</td>
</tr>
<tr>
<td>Lavan</td>
<td>0.67</td>
<td>0.88</td>
<td>0.55</td>
<td>0.45</td>
<td>0.36</td>
<td>0.39</td>
<td>0.03</td>
</tr>
<tr>
<td>Holur</td>
<td>0.64</td>
<td>0.76</td>
<td>0.70</td>
<td>0.75</td>
<td>0.45</td>
<td>0.57</td>
<td>0.12</td>
</tr>
<tr>
<td>Pichakan</td>
<td>0.41</td>
<td>0.89</td>
<td>0.68</td>
<td>0.32</td>
<td>0.28</td>
<td>0.29</td>
<td>0.01</td>
</tr>
<tr>
<td>Burkh</td>
<td>0.33</td>
<td>0.65</td>
<td>0.40</td>
<td>0.6</td>
<td>0.13</td>
<td>0.39</td>
<td>0.26</td>
</tr>
<tr>
<td>Gashu</td>
<td>0.62</td>
<td>0.77</td>
<td>0.48</td>
<td>0.53</td>
<td>0.30</td>
<td>0.40</td>
<td>0.10</td>
</tr>
<tr>
<td>Bonaruyeh</td>
<td>0.74</td>
<td>0.82</td>
<td>0.36</td>
<td>0.64</td>
<td>0.26</td>
<td>0.52</td>
<td>0.26</td>
</tr>
<tr>
<td>Hendurabi</td>
<td>0.44</td>
<td>0.83</td>
<td>0.18</td>
<td>0.71</td>
<td>0.08</td>
<td>0.59</td>
<td>0.51</td>
</tr>
<tr>
<td>Paskuhak</td>
<td>0.55</td>
<td>0.67</td>
<td>0.54</td>
<td>0.66</td>
<td>0.30</td>
<td>0.43</td>
<td>0.13</td>
</tr>
<tr>
<td>Dashtak</td>
<td>0.63</td>
<td>0.68</td>
<td>0.58</td>
<td>0.62</td>
<td>0.36</td>
<td>0.42</td>
<td>0.06</td>
</tr>
</tbody>
</table>
of this study, a hydrocarbon structure may be a member of both the lower and upper approximations of the concepts to some degree. In other words, they are fuzzy concepts.

This research has another valuable role in the development of a GIS-based framework for hydrocarbon well drilling priority planning. The proposed method outstanding distinction is related to uncertainty quantification. Table 2 presents the uncertainty of the proposed. Suppose that the geoscience experts decide to produce a priority road map for drilling in an area. If they do not access to uncertainty quantification of each hydrocarbon structure, they will select the wrong structure to drill. For example, if their knowledge was limited to belief and plausibility of the structures, they would not be able to distinguish the more proper structure for drill among ‘Pichakan’ and ‘Hendurabi’ structures because the values are close to each other. If experts just concentrate on BGC and GGC, they will select

![Figure 9](image_url)
the second structure certainly which will be a wrong choice according to the real drilling results. However, if they consider the uncertainty values obtained from our proposed model, the ‘Pichakan’ structure will be elected because it has a 50% less uncertainty than ‘Hendurabi’ structure. The prominence of this research is increased when considering that the cost of drilling an off-shore well would be roughly 10 million US Dollars.

Some issues restrict implementing the proposed model in real-world applications although this had no impact on our study. The first and foremost limitation is that the dependency among diverse evidences has to be regarded prior to the application of the combination rule (Su et al. 2016). This problem was resolved as there was no overlap and dependency between the petroleum system parameters whose factors for all the four layers of the petroleum systems were independent in nature. The other important point is that an extensively practicable method of producing basic probability assignment has to be devised in order to model the uncertain information for each application (Zhou et al. 2016). A modified method was used in this study as a conceptual model to tackle such an issue.

The risk of hydrocarbon exploration in fields was compared using DST, fuzzy and hybrid models. It was observed that among 102 prospects, all of the prospects that were categorized in the ‘Very High’ and ‘Very Low’ classes remained in the same risk classes. However, 14 prospects were categorized in different risk classes as eleven prospects that were categorized in the ‘Low risk’ class were ranked as ‘Moderate risk’ class and three prospects that were in the ‘High risk’ class were classified into the ‘Moderate risk’ class.

By comparing the results with the drilled hydrocarbon wells in the study area, it was observed that a larger number of hydrocarbon structures were categorized by the hybrid GIS-assisted method to have accurate results compared to the other two methods. As shown in Figure 10, in the first and last category where the uncertainty of model is less than other groups, the fusion of two methods could be neglected. However, in the three other classes, the proposed model enhanced the accuracy of classification by up to 19, 27 and 24%, respectively.

6. Conclusion

Selecting an appropriate spot for drilling the exploration well on a structure has been claimed to be essential in that any fault in computing the drilling place can result in
wasting the capital and time. The results of this research could improve the process of making decision under risk and uncertainty conditions by facilitating the prioritizing of hydrocarbon structures for the drilling phase, which provides necessary information about uncertainty and risk that exist in the area.

The current research demonstrated that using GIS as a foundation to model geologic risk and accompanied uncertainties would be an invaluable scientific approach. GIS offers an environment that allows spatial analysis, such as Thiessen polygon, kriging analysis and compiling maps with better accuracy. Modified Dempster–Shafer functions were used as a means to model the reasoning behaviour under uncertainty with the purpose of integrating all the relevant spatial data for hydrocarbon exploration in a GIS environment. The model enhances the geologic risk analysis of a Dempster–Shafer data-driven method by a fuzzy logic approach. The most momentous privilege of the proposed method is that it models both the belief and plausibility of the evidences related to the hydrocarbon existence and common concepts between them and BGC as well as GGC obtained as fuzzy membership values. This research delineated a hybrid GIS-assisted model, which considers the uncertainty of hydrocarbon exploration and minimizes the geologic risk.

Developing a GIS-based software for modelling spatial uncertainties in an exploration process is a useful plan for future works. Producing an interactive and location based tool in order to create geologic risk maps quickly would be of immense help to hydrocarbon exploration decision makers. Future studies should investigate and assess other types of uncertainties in risk analysis of hydrocarbon exploration by applying other soft computing methods in order to reduce them. The fuzzification process in this study was performed in combination with a Thiessen polygon analysis in the GIS environment but future researches could focus on other methods for applying different fuzzifiers.

Acknowledgments

All the respectable reviewers are acknowledged for their fruitful comments and suggestions about the paper.

Disclosure Statement

The authors declare no conflict of interest.

Notes on contributors

Mahmoud Reza Delavar designed the initial idea of the research and helped Sahand Seraj in the methodology development. Sahand Seraj gathered the data. Both Sahand Seraj and Mahmoud Reza Delavar analyzed the data; Sahand Seraj undertook the computational parts and the experiments. Sahand Seraj wrote the draft paper. Mahmoud Reza Delavar and Reza Rezaee critically reviewed and extended all the sections of paper. Sahand Seraj finalized the paper.

Funding

The authors would like to thank the University of Tehran Vice President for Research to fund the research with the Grant N. 8103002/6/46 awarded to the second author.
References


Amiri MA, Karimi M, Alimohammadi A. 2015. Hydrocarbon resource potential mapping using the evidential belief functions and GIS, Ahvaz/Khuzestan province, southwest Iran. Arab J Geosci. 8: 3929–3941.


Hong H, Naghibi SA, Dashtpadderdi MM, Pourghasemi HR, Chen W. 2017a. A comparative assessment between linear and quadratic discriminant analyses (LDA-QDA) with frequency ratio and weights-of-evidence models for forest fire susceptibility mapping in China. Arab J Geosci. 10:167.


