Fuzzy Shannon Entropy: A Hybrid GIS-Based Landslide Susceptibility Mapping Method

Majid Shadman Roodposhti 1, Jagannath Aryal 1,*, Himan Shahabi 2 and Taher Safarrad 3

1 Discipline of Geography and Spatial Sciences, School of Land & Food, University of Tasmania, Hobart 7001, Australia; majid.shadman@utas.edu.au
2 Department of Geomorphology, Faculty of Natural Resources, University of Kurdistan, Sanandaj 66177-15175, Iran; h.shahabi@uok.ac.ir
3 Department of Geography and Urban Planning, Faculty of Humanities and Social Science, University of Mazandaran, Babolsar 47416-13534, Iran; t.safarrad@umz.ac.ir
* Correspondence: jagannath.aryal@utas.edu.au; Tel.: +61-3-6226-2848

Academic Editor: Raúl Alcaraz Martínez
Received: 1 August 2016; Accepted: 19 September 2016; Published: 27 September 2016

Abstract: Assessing Landslide Susceptibility Mapping (LSM) contributes to reducing the risk of living with landslides. Handling the vagueness associated with LSM is a challenging task. Here we show the application of hybrid GIS-based LSM. The hybrid approach embraces fuzzy membership functions (FMFs) in combination with Shannon entropy, a well-known information theory-based method. Nine landslide-related criteria, along with an inventory of landslides containing 108 recent and historic landslide points, are used to prepare a susceptibility map. A random split into training (≈70%) and testing (≈30%) samples are used for training and validation of the LSM model. The study area—Izeh—is located in the Khuzestan province of Iran, a highly susceptible landslide zone. The performance of the hybrid method is evaluated using receiver operating characteristics (ROC) curves in combination with area under the curve (AUC). The performance of the proposed hybrid method with AUC of 0.934 is superior to multi-criteria evaluation approaches using a subjective scheme in this research in comparison with a previous study using the same dataset through extended fuzzy multi-criteria evaluation with AUC value of 0.894, and was built on the basis of decision makers’ evaluation in the same study area.

Keywords: Shannon entropy; fuzzy membership function (FMF); landslide susceptibility mapping (LSM); Izeh

1. Introduction

A landslide is either geophysical or climate-related disaster that is described as a mass movement of earth surface material. This usually involves shear displacement of soil and/or rock masses along one or several slip surfaces [1]. A landslide susceptibility map (LSM) is a promising solution for both understanding and predicting probable future landslides. It assists planners in decision-making phase aimed for further mitigation of landslide consequences. Accordingly, a LSM depicts areas likely to have landslides in the future by correlating some of the principal factors that contribute to landslides with the past distribution of slope failures [2]. In this respect, production of LSM at the early stage of landslide assessments is of crucial importance for safe economic planning, such as urbanization activities and the engineering of structures. However, a standard procedure for the production of landslide susceptibility maps does not exist [3]. Thus, LSM can be accomplished by providing risk managers with easily accessible, continuous, and accurate information about landslide occurrence. The predictive capacity is poorly understood in LSM and is vague. In general, the spatial prediction of landslides is not easy due to the complex nature of landslides [4]. LSM provide important information...
for predicting landslides hazards which include an indication of the time scale within which particular landslides are likely to occur [5]. The associated vagueness can be dealt using fuzzy sets theory.

Introduced by Zadeh, fuzzy set theory handles indefiniteness arising from intrinsic ambiguity than from a statistical variation [6]. A functional defined on the class of generalised characteristic functions (fuzzy sets), called “entropy”, is introduced using no probabilistic concepts in order to obtain a global measure of the indefiniteness connected with the situations described by fuzzy sets [7]. The meaning of this quantity is quite different from the one of classical entropy because no probabilistic concept is needed in order to define it. This function gives a global measure of the “indefiniteness” of the situation of the problem at hand [8]. Although there is a well-defined mathematical theory of probability, there is no universal agreement about the meaning of probability. Thus, for example, there is the view that probability is an objective property of a system and another view that it describes a subjective state of belief of a person. Then there is the frequentist view that the probability of an event is the relative frequency of its occurrence in a long or infinite sequence of trials. Thus, entropy is often used as a characterization of the information content of a data source, this information content is not absolute: it depends crucially on the probabilistic model [9].

Effective LSMs could provide a proper understanding of “susceptible regions” [10]. In order to better assist planners in understanding landslide hazard, a variety of GIS-based susceptibility mapping techniques are employed and developed [11]. These approaches can be classified into three main groups: subjective, objective and hybrid methods. The subjective methods typically include inventory mapping and decision makers’ (DMs) evaluation in both standardisation and weighting of selected criteria [12]. There are various GIS-based studies on LSM through the use of subjective approaches. Some of them used multi-criteria evaluation (MCE) techniques including: simple additive weighting [13], ordered weighted average [14], analytical hierarchy process [15], analytical network process [16], PROMETHEE [17], etc. and some used different heuristic and knowledge driven techniques in order to assess landslide susceptibility mapping [18–20]. Other studies, on the other hand, have shown a variety of objective methods in the assessment of the landslide susceptibility because of some limitations such as insufficient knowledge about the area of interest. The objective methods mostly rely on statistical [21–27], soft computing [4,28,29], deterministic analysis [30], neuro-fuzzy [4,31], artificial neural network [32–34], decision trees [35,36], and index of entropy [37–41], which are more rigorous and mostly relying on objective assessments. On the other hand, there are various hybrid GIS-based methods which are both subjective and objective. In other words, some hybrid GIS-based LSM methods used subjective standardisation and an objective weighing technique [42–44], and vice versa.

The accuracy of LSM mostly depends on the amount and quality of available data, the working scale and the selection of the appropriate methodology for analysis and modelling [17]. In methodology implementation and its assessment, landslide casual criteria play a key role. In this study, we decipher the optimality of predictive solutions for objective criteria weighting. In an attempt to find an optimal solution, we show how modified Shannon entropy algorithm in association with fuzzy set theory can be successfully applied to the numerical solution of the LSM while there is no sufficient knowledge about the area of interest. In other words, the main objective of the present study is to extend a hybrid GIS-based LSM method within which fuzzy membership functions (FMFs) have been applied for criteria standardisation using “global knowledge” about landslides, while no “local knowledge” is utilised for criteria weighting. In literature, although different GIS-based models have been used for landslide susceptibility mapping, however, LSM map extracted from modified Shannon entropy algorithm in association with fuzzy set theory has seldom been carried out. Therefore, this study aims to fill this identified gap in the relevant literature.

Since the LSM deals with a various sets of criteria it can be assumed that integration of fuzzy set theory with information theory, and in particular with Shannon entropy, will assist in performing accurate landslide susceptibility mapping. This accurate LSM is due to the flexibility of fuzzy membership functions and objective evaluation of criteria weights. Based on this assumption, the present research is an attempt to propose a novel hybrid method, which contributes to the objective...
decision making for regional landslide management. In other words, by using only the entropy values of previous landslide events for each criterion and regardless of experts’ opinions, we intend to facilitate criteria weighting process while improving or preserving LSM predictive accuracy compared with accurate subjective methods.

The paper is organized as follows: after a description of the study area in Section 2, a detailed definition of the material and methods of the research is described in Section 3. Section 4 presents results while Section 5 discusses the achieved results and contributions, respectively. At the end, we provide the conclusions of this research in Section 6.

2. Description of the Study Region

Izeh is located in the eastern part of Khuzestan province, in south-western Iran (see Figure 1), where the high susceptibility for a mass movement and in particular landslides is considered as a potential natural hazard for human society and their activities such as the hydropower plants in Izeh. According to the inventory of landslides compiled by the Ministry of Natural Resources [45], there are 108 recorded landslide events in the region.

![Figure 1. Location map of the study area.](image)

The climate is a temperate in north, while in south a warm climate prevails. Similarly, mean annual precipitation within the study area varies from 450 to 700 mm. The region is important in terms of the agricultural activities and in particular hydropower plants. The Karun River, the main and longest river in all of Iran, passes through this area. The suitable topography of Karun canyon provides the possibility of constructing hydropower plants and three main dams have been constructed so far on different branches of the Karun River [44].

Geologically, there are several minor faults and one major thrust in the region along with the 13 types of geologic formations cropping out in the region. The Izeh fault zone is a transverse fault zone with right-lateral strike slip (and some reverse component) in the Zagros Mountains, south-western Iran. That is majority controlled by the subsidence and sedimentation of the embayment. In terms of
13 types of geologic formations, nearly all of them composed of sedimentary rocks including, marl, shale, limestone, gypsum, siltstone and other Quaternary deposits. It also should be mentioned that in the case of any triggering cause, there will be a significant chance of landslide occurrence within the south and south-east where the rough topography coincides with major thrust fault, Karun canyon and susceptible lithology. In other words, where there is susceptible lithology, proximity to faults contributes to slope instability, affecting not just surface structures but also terrain permeability. Eventually, the erosion associated with Karun River in nearby areas further leads to slope instability and generally increases the rate of subsequent slope failure. This is considered another prominent reason for the notable landslide recurrence in the region [44].

3. Materials and Methods

3.1. Landslide Influencing Data Layers

First of all, with respect to the available peer-reviewed GIS-based LSM research, nine criteria of the study area have been employed and prepared (Table 1 and Figure 2).

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Data Source</th>
<th>Former Studies Using the Same Criterion for GIS-Based LSM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slope</td>
<td>30 m, STRM DEM</td>
<td>Lee and Min [21]; Komac [46]; Ayalew et al. [14]; Conoscenti et al. [47]; Thiery et al. [48]; Yalcin [15]; Kayastha et al. [49]; Bennett et al. [50]; Kritikos et al. [51]</td>
</tr>
<tr>
<td>Aspect</td>
<td>30 m, STRM DEM</td>
<td>Ayalew and Yamagishi [52]; Komac [46]; Guzzetti et al. [53]; Thiery et al. [48]; Yalcin [15]; Loffi et al. [54]</td>
</tr>
<tr>
<td>River</td>
<td>1:50,000, Topo-map</td>
<td>Yalcin [15]; Feizizadeh et al. [44]; Faraji Sabokbar et al. [55]</td>
</tr>
<tr>
<td>Drainage</td>
<td>1:50,000, Topo-map</td>
<td>Yalcin [15]; Pareek et al. [56]; Shadman et al. [17]; Feizizadeh et al. [44]</td>
</tr>
<tr>
<td>Fault</td>
<td>1:100,000, Geo-map</td>
<td>Havenith et al. [22]; Kanungo et al. [57]; Lee and Pradhan [58]; Marjanović et al. [59]; Shababi et al. [60]</td>
</tr>
<tr>
<td>Rainfall</td>
<td>30 years, IMO data</td>
<td>Hong et al. [61]; Guzzetti et al. [62]; Feizizadeh et al. [44]</td>
</tr>
<tr>
<td>Road</td>
<td>1:50,000, Topo-map</td>
<td>Ayalew and Yamagishi [52]; Yalcin [15]; Youssef et al. [63]; Bathrellos et al. [64]; Pradhan [31]</td>
</tr>
<tr>
<td>Lithology</td>
<td>1:100,000, Geo-map</td>
<td>Ercanoglu and Gokceoglu [3]; Ayalew and Yamagishi [52]; Thiery et al. [48]; Akgun et al. [65]; Davis and Blesius [43]</td>
</tr>
<tr>
<td>Land use</td>
<td>30 m, Landsat image</td>
<td>Lee and Pradhan [58]; Bathrellos et al. [64]; Feizizadeh et al. [44]</td>
</tr>
</tbody>
</table>

It must be noted that landslide susceptibility map of the study area has been derived from landslide related criteria mentioned in above Table 1. Road, river and drainage input maps were extracted from the topographical map (1:50,000) of the study area, while the fault and lithology maps were obtained from geologic maps (1:100,000). In addition, the slope and aspect criteria were derived from 30 m shuttle radar topography mission (SRTM) digital elevation model (DEM). Land use/cover maps were derived from Landsat ETM+ satellite images with 30 m spatial resolution employing image analysis methods [66].

The average of 30 years mean annual rainfall data from the Iran Meteorological Organization (IMO) was used to create mean annual rainfall map using kriging interpolation methods in the ArcGIS environment. Finally, we also used an inventory of landslides containing 108 recent and historic landslide points which were recorded by GPS in field survey [45] for both geo-data layer weight evaluation and further validation of proposed LSM. The recorded landslides points are centroids of each landslide polygon. Almost all of these landslides belong to slide-type landslides which are down-slope movement of material along a distinctive surface of weakness such as a fault, joint or bedding plane. In terms of landslide inventory, it was randomly split into a train (≈70%) and test (≈30%) samples for training the proposed hybrid model and subsequent validation purpose, prior and posterior map elaboration, respectively.
Figure 2. Nine applied criteria used in LSM of Izeh involving: (a) Slope; (b) Aspect; (c) Distance to river; (d) Drainage Density; (e) Distance to faults; (f) Mean annual rainfall; (g) Distance to roads; (h) Lithology and (i) Land use/cover.
3.2. Proposed Methodology

In order to depict the proposed methodology, it is best to consider a three-step procedure: in step 1, using fuzzy sets theory, data standardisation has been implemented in ArcGIS environment. To this end, a proper FMF is fitted on each selected criterion posterior to preprocessing phase. These FMFs have been selected according to peer literature review of similar LSM studies and local expert opinions. Accordingly, in step 2, the Shannon entropy is used for further evaluation of criteria weights, which determines the subsequent contribution of each landslide related criteria in overall susceptibility. This phase is implemented in MATLAB. Here, Shannon entropy technique is used as an objective-weighting scheme in LSM process. Finally, in the third step, results from above two steps are integrated using ArcGIS software (Figure 3). Further, results were validated using receiver operating characteristics (ROC) curves and simple overlay technique using MATLAB and ArcGIS environments, respectively. Each step is explained as below:

3.2.1. Fuzzy Membership Function (FMF)

A major contribution of fuzzy set theory and related fuzzy membership functions (FMFs) is its capability of representing vague data. The theory also allows mathematical operators and programming to apply to the fuzzy domain. A fuzzy set is a class of objects with a continuum of grades of membership [6]. Such a set is mainly characterised by a membership function, which designates a membership value to every single object ranging from 0 and 1 and vice versa. In terms of LSM, fuzzy sets approve the possibility of partial membership of a considered geographic location to more than one susceptibility class. FMFs accordingly, were used to determine pattern variation forming a gradual class boundaries between each susceptibility class. The shape of each applied FMF determines how the transition between 0 and 1 takes place.

3.2.2. Shannon Entropy

The information theory application, originated from thermodynamics in 1948 [42], is used in diverse disciplines and application areas. In information theory, entropy is the quantitative measure of system disorder, instability, imbalance, and uncertainty and can forecast development trend of specified system [42,54,67,68]. The Shannon entropy usually indicates to quantification of the expected
amount of information enclosed by a message. At present, it has been widely used to determine the weighted index in natural hazards, and in integrated assessment of natural-environmental processes such as debris flows, droughts, sandstorms, etc. [69,70].

In the case of landslides, it measures the dissimilarity or diversity in the environment, indicating the potential of each factor in causing landslides. In other words, the entropy of landslides refers to the extent the various factors influence landslide. Greater is the entropy index, greater is the influence of the factor in causing landslide [68]. Finally, it also should be mentioned that various landslide related criteria are not the same regarding their attributes and dimension. Therefore, it is not possible to conduct a direct comparison between those mentioned criteria which are applied in a LSM process. In order to construct proper comparison, it is necessary to conduct standardisation process in the first step (see Equations (1) and (2)).

\[
\text{Positive effect} = \begin{cases} 
1 & \text{if } x = x_{\text{max}} \\
0.5(1 - \cos(\pi \frac{x - x_{\text{min}}}{x_{\text{mac}} - x_{\text{min}}})) & \text{if } x_{\text{min}} < x < x_{\text{max}} \\
0 & \text{if } x = x_{\text{min}}
\end{cases}
\]

\[
\text{Negative effect} = \begin{cases} 
1 & \text{if } x = x_{\text{min}} \\
0.5(1 + \cos(\pi \frac{x_{\text{mac}} - x}{x_{\text{mac}} - x_{\text{min}}})) & \text{if } x_{\text{min}} < x < x_{\text{max}} \\
0 & \text{if } x = x_{\text{max}}
\end{cases}
\]

Equation (1) is applicable for specific criteria with positive effect on probability of landslide occurrence (such as drainage density and mean annual rainfall). It means the more the value of the considered criteria is, the more the probability of landslide is as a simple rule. However, for some other criteria (viz. distance to river, distance to faults and distance to road) Equation (2) is well-suited, where the reverse condition exists. Then, landslide entropy matrix \( R \) is formed by \( m \) landslide samples and \( n \) geo-data layer:

\[
R = \begin{bmatrix}
\ell_{1,1} & \ell_{1,2} & \cdots & \ell_{1,n} \\
\ell_{2,1} & \ell_{2,2} & \cdots & \ell_{2,n} \\
\vdots & \vdots & \ddots & \vdots \\
\ell_{m,1} & \ell_{m,2} & \cdots & \ell_{m,n}
\end{bmatrix}
\]

Accordingly, Shannon entropy is defined by [42]:

\[
E_j = -k \sum_{i=1}^{m} p_{ij} \ln p_{ij}
\]

where \( E_j \) is entropy value, \( p_{ij} \) is value of \( i^{\text{th}} \) landslide in \( j^{\text{th}} \) criteria and \( k \) is a positive constant, essentially a choice of unit of measure which is given by:

\[
k = (\ln m)^{-1}
\]

where \( m \) is number of occurred landslide events. Accordingly, normalized decision matrix \( p_{ij} \) for each landslide criteria can be defined by:

\[
p_{ij} = \frac{\ell_{ij}}{\sum_{i=1}^{m} \ell_{ij}}
\]
The weights have attributed the role the factors play in the synthesis assessment, and the bigger value indicates that the factor’s function is more important in this index system.

\[ w_j = \frac{v_j}{\sum_{i=1}^{m} v_i} \]  

(7)

where \( W_j \) is weight of \( j^{th} \) geo-data layer and \( V_j \) is defined by:

\[ v_j = 1 - E_j \]  

(8)

3.2.3. Hybrid Landslide Susceptibility Mapping Model

The proposed hybrid model of landslide susceptibility mapping can be defined as:

\[ S = \sum_{i=1}^{n} w_i \otimes x_i \]  

(9)

where \( S \) is a degree of landslide susceptibility, \( W_i \) stands for the weight of each criterion and \( X_i \) is standardised landslide criteria.

3.3. Methodology Implementation

The following 3-steps experimental design is implemented (see Figure 3).

3.3.1. Step 1: Data Standardisation Using FMFs

Considering the fact that GIS-based landslide related criteria measured not only in different units but also in different scales of measurement (data types), such as nominal, ordinal, interval, and ratio scales [71], there is an urgent need for data standardisation. This rises from the inherent need to integrate all landslide criteria into the single output in the evaluation process. In this regard, the fuzzy membership approach is considered one of the frequently applied standardisation methods that have been proposed [72].

The use of fuzzy sets within GIS-based hazard and susceptibility assessment has been demonstrated to have a good effect [71–73]. For this reason, fuzzy sets were used in this study. In this context, all the factors used were standardised to a float-level range of 0–1, where 0 is assigned to the least susceptible areas and 1 to the most susceptible ones. This transforms the different measurement units of all landslide casual criteria into comparable values using FMFs [74]. Figure 4 shows selected and applied FMFs for LSM of the study region.

There is no optimal method for choosing the most appropriate FMF and their respective parameters; these are generally selected according to the preferences of the DMs [17,75]. However, the predictive and causal value of landslide casual criteria seems more or less similar in most of the studies. In this study, three different membership functions have been employed for landslide susceptibility purpose including sigmoidal (s-shaped) FMFs, i.e., monotonically increasing and monotonically decreasing, user-defined fuzzy membership functions along with crisp membership functions are specified for each criterion (see Figure 4). The sigmoidal membership function is likely the most commonly used FMF in fuzzy set theory, and provides a gradual variation from non-membership (zero) to complete membership (one) [6,71,72,76], whereas it is sometimes inevitable to use user-defined FMFs or crisp membership functions. Nevertheless, all applied functions of criteria and the resultant output raster files are shown in Figures 4 and 5, respectively.
Figure 4. All types of used membership functions including: (Type I) User defined FMF (for: (a) slope and (b) aspect), (Type II) Sigmoidal FMF including both monotonically decreasing (for: (c) distance to river, distance to faults and distance to road) and monotonically increasing (for: (d) drainage density and mean annual rainfall) and (Type III) Crisp MF (for: (e) lithology and (f) land use).

Figure 5. Cont.
3.3.2. Step 2: Assessment of Weights with Shannon Entropy

Here, this study selects a total number of 76 landslides to calculate the weights of each landslide criteria in southern Izeh. Following the insertion of all nine selected landslide criteria into the entropy matrix \( R \), in order to standardise the basic data into the mentioned entropy matrix, both Equations (1) and (2) were employed along with user defined FMF and crisp MF (Figure 4). Next, weights of all criteria are determined by the help of Shannon entropy in the successive steps:

\[
R = \begin{bmatrix}
1.000 & 0.583 & 0.747 & 1.000 & 0.000 & 0.143 & 0.970 & 1.000 & 0.300 \\
0.337 & 0.000 & 0.708 & 1.000 & 0.000 & 0.138 & 1.000 & 1.000 & 0.300 \\
0.364 & 0.124 & 0.000 & 0.140 & 0.184 & 0.122 & 1.000 & 0.900 & 0.600 \\
\end{bmatrix}
\]

Entropy values and weights can be calculated using (4)–(7):

\[
E_j = \{ 0.883, 0.874, 0.735, 0.902, 0.715, 0.861, 0.915, 0.927, 0.902 \} \quad (11)
\]
\[
v_j = \{ 0.116, 0.097, 0.264, 0.097, 0.284, 0.138, 0.084, 0.072, 0.097 \} \quad (12)
\]
\[
w_j = \{ 0.090, 0.097, 0.206, 0.075, 0.222, 0.108, 0.066, 0.056, 0.075 \} \quad (13)
\]
Finally, after further calculation of entropy value, we obtain the weights (Table 2) to be used in criteria integration in the next step.

### Table 2. The calculated weight vector from Shannon entropy method.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slope</td>
<td>0.090</td>
</tr>
<tr>
<td>Aspect</td>
<td>0.097</td>
</tr>
<tr>
<td>Distance to river</td>
<td>0.206</td>
</tr>
<tr>
<td>Drainage density</td>
<td>0.075</td>
</tr>
<tr>
<td>Distance to Fault</td>
<td>0.222</td>
</tr>
<tr>
<td>Rainfall</td>
<td>0.108</td>
</tr>
<tr>
<td>Distance to roads</td>
<td>0.066</td>
</tr>
<tr>
<td>Lithology</td>
<td>0.056</td>
</tr>
<tr>
<td>Land use/cover</td>
<td>0.075</td>
</tr>
</tbody>
</table>

#### 3.3.3. Step 3: Integration Phase

The prepared database of LSM is successfully georeferenced using Universal Transverse Mercator (UTM) coordinate system in the ArcGIS environment. The weight derived from Shannon entropy index for each landslide related criteria is calculated using the 76 occurred landslide events and is applied for integration purpose. Afterwards, the resultant susceptibility map is calculated as the summation of the weighted criteria as shown in Equation (9).

Using all the factors (Table 1), susceptibility values range from 0 to 8.00 showing various levels of susceptibility. The higher susceptibility values refer greater probability of expected landslides occurrence in the near future. Finally, the landslide susceptibility map (Figure 6) is divided into five susceptibility classes very low, low, moderate, high and very high using natural breaks classification.

![Figure 6. Final susceptibility map using the proposed hybrid GIS-based method. A, B and C circles are only for better representation of contiguous landslide points density and positions within the study area.](image-url)
The “natural breaks” classifier is based on natural grouping of data values. Normally, the breakpoints are identified by looking for groups and patterns inherent in the data. However, the reasons for using certain methods in previous works are usually not explained by the authors. In this study, the manual classifier method was used to reclassify the LSM values into five different susceptibility zones, according to the classification method that was proposed by [77].

4. Results

After layer standardisation of landslide casual criteria, the susceptibility map was produced based on a hybrid GIS-based LSM technique (see Figure 6). In terms of criteria weighting, typically, in spatial MCDA (multi-criteria decision analysis) problems the greater the value of the entropy corresponding to a spatial attribute, which implies the smaller attribute’s weight, the less the discriminant power of that attribute in decision-making process (see Equation (8)) [54]. Accordingly, fault and river criteria are considered as the two first important landslide criteria. Also, the distance from faults and river criteria both indicate potential trigger factors of the slope failure, are also among the principal indexes of a landslide. As a result, the objectively obtained weights of the landslide related criteria using Shannon entropy index is consistent with the basic rules of identification, characterization and development of landslides.

4.1. Validation of the Results Using ROC Curve

The validation phase could be considered as one of the most fundamental stages in the development of all susceptibility maps and determination of their prediction capability for future usage in any natural hazards study and managements. The prediction efficacy of each LSM and its resultant output is typically evaluated by using available independent information of recorded landslide events, which are not used through LSM process (i.e., test subset of landslide inventory map) [17]. As a result, in the present study, the landslide inventory database has been divided into two parts, including training and test datasets. Therefore, the accuracy of the proposed LSM in the study area was evaluated by calculating relative operating characteristics (ROC) [17,24,78] and percentage of known occurred landslides events in various susceptibility classes using test landslide samples. Here, the Area under the ROC Curve (AUC) value, ranging from 0.5 to 1.0, is a numeric indicator of map accuracy. Meaning that AUC is close to 1, the result of the test is more reliable, while closer the AUC to 0.5 indicates to the less reliable result [17,31].

In pursuance of further implementation of the ROC evaluation technique, a precise and comprehensive test dataset was prepared using 32 landslides and 32 randomly selected non-landslide points of the study area. In this regard, following the early identification of landslide-free area using aerial photo interpretation and field survey, non-landslide points are selected within the boundary of these landslide-free areas. Subsequently, the AUC value of 0.934 has been obtained with standard deviation (area) of 0.034 (see Figure 7).

![Figure 7. ROC curve for the proposed Landslide Susceptibility Map (LSM) of fuzzy Shannon entropy.](image-url)
4.2. Validation of the Results Using Simple Overlay

In the second validation process, the LSM result has been evaluated using the test landslide locations, accordingly, these 32 points were overlaid on the susceptibility map of proposed hybrid GIS-based LSM (see Figure 6). The result shows that approximately about 90 percent of the recorded landslides occurred in the high and very high susceptibility classes, which only cover 30.63% of the study area, while, no recorded landslide appears in the low and very low susceptibility zones. In addition to the above, only three landslide points (≈10% of all recorded landslides) fall into the medium susceptibility zone of the map, which covers about 18.22% of the study area (see Figure 8).

![Figure 8. Histogram of calculated landslide susceptibility map showing the relative areas for each susceptibility class (susceptibility classes are labelled with the numbers of the observed landslide points accordingly).](image)

5. Discussion

The accuracy of predictive models is considered a major concern in the majority of environmental modelling applications including LSM [50]. The predictive accuracy of subjective LSM models can be affected by the inherent bias that emanates from DMs’ point of view during both data standardisation and criteria weighting. Moreover, the absence of expert DMs may be a serious hindrance in the LSM process when using a subjective method. Considering criteria standardization schemes, by applying a more computationally intensive approach we attempted to preserve the original quality of spatial data. In this respect using a variety of FMFs positively affected the validity and accuracy of input spatial criteria. Therefore, missing or generalised values can represent otherwise precise data. Further, the proposed methodology shows promising results to predict landslide susceptibility values regardless of experts’ opinion. According to the obtained results, the accuracy of the proposed hybrid model is improved significantly compared with the accuracy of accurate subjective approaches, which have been previously implemented in the study area using the same dataset [44].

5.1. Obtained Results and Relevance to the Previous Studies

Considering the high frequency of landslides occurring in several areas of southern Izeh, there was a demand to establish an accurate landslide susceptibility map. The expected accuracy of LSM depends not only on the presence of concise and comprehensive data, in terms of data scale and accuracy, but also on the selection of the appropriate methodology of data processing and modelling [15]. Regardless of data scale and accuracy, the present study aimed to explore landslide susceptibility of southern Izeh by developing a hybrid GIS-based LSM that uses neither DM's evaluation nor sophisticated objective methods. This is an integrated strategic LSM framework with an emphasis on structuring the decision-making process problem. Within this approach, Shannon entropy was employed to determine the criteria weightings from an objective evaluation of spatial domain while different fuzzy membership functions were employed for criteria standardization.
Obtained results of ROC curve analysis (AUC = 0.934) (see Figure 7) and simple overlay technique (see Figure 8) signify that the proposed hybrid fuzzy Shannon entropy evaluation technique is a promising tool for integrating multiple raster-based criteria for LSM while there is not sufficient knowledge about the criteria weights with respect to landslide mechanism of the study region. The previous study using the same dataset through extended fuzzy multi-criteria evaluation which was built on the basis of DMs’ evaluation achieved AUC value of 0.894 [44]. This further approves the capability of proposed hybrid model for prediction of landslide susceptibility values. In other words, achieved results of accuracy metrics comparison approves that the proposed LSM model can achieve superior prediction accuracy to what that can be achieved by using DMs’ points of view (Table 3), with significant time saving.

Table 3. Accuracy metrics of implemented data-driven (objective) and expert-driven (subjective) LSMs using fuzzy Shannon entropy and extended fuzzy multi-criteria evaluation methods, respectively.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Objective Weighting Approach</th>
<th>Subjective Weighting Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Cases</td>
<td>64</td>
<td>212</td>
</tr>
<tr>
<td>Number Correct</td>
<td>56 (76% of total)</td>
<td>173 (81% of total)</td>
</tr>
<tr>
<td>AUC</td>
<td>0.93</td>
<td>0.89</td>
</tr>
<tr>
<td>Std. Dev. (Area)</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>Accuracy</td>
<td>76.6%</td>
<td>81.6%</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>100.0%</td>
<td>98.1%</td>
</tr>
<tr>
<td>Specificity</td>
<td>53.1%</td>
<td>65.1%</td>
</tr>
<tr>
<td>Pos Cases Missed</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Neg Cases Missed</td>
<td>15</td>
<td>37</td>
</tr>
</tbody>
</table>

5.2. Spatial Information Extraction and Prediction

This study contributed in the area of the spatially structured dilemma of predicting landslide susceptibility values for specific geographic locations. This may be implemented through standardising and subsequent summing of landslide casual criteria. In this paper, we attempted to present an assessment of LSM, carried out by the implementation of hybrid fuzzy Shannon entropy evaluation within which fuzzy set theory has been used for criteria standardisation, and Shannon entropy algorithm was used for weighting of some factors that may affect the landslide susceptibility. Therefore, the prepared hybrid susceptibility map is the result of a pixel-based combination of nine standardised criteria affecting the degree of landslide susceptibility. The optimal criteria weights are obtained objectively by a precise mathematical solution through the proposed entropy-based model [79]. In this respect, the lower the landslide entropy of a criterion, the higher the weight is. In other words, a lower landslide entropy within certain criteria (i.e., distance to faults and distance to river) indicates the presence of predictive spatial frequency and vice versa.

Further, as expected, the estimated data driven (objective) criteria weights using Shannon entropy algorithm do not conform to the subjective criteria weights estimated using an aggregation of DMs’ votes from our prior research (Figure 9).

According to the obtained fuzzy Shannon entropy criteria weighting scheme results the distance to fault is the most important criterion, followed by distance to river and rainfall criteria, respectively. Therefore, considering the estimated criteria weights, the spatial distribution of landslide susceptibility values is mostly controlled by these mentioned criteria. This may be further proved by the high concentration of recorded landslide events along the Karun River (Figure 6). Nonetheless, considering the DMs’ evaluation slope is referred to as the most significant criterion followed by lithology and distance to road layers. Considering these two weighting approach, fuzzy Shannon entropy seems more realistic for predictive modelling of spatial pattern of landslides compared to the latter method. Even though the slope criterion is of paramount importance in any shape of slope instability, it is not the only constituent of landslides. Accordingly, the spatial pattern of landslides (at least in the study region) is controlled by other important but less geographically available landslide casual criteria (distance
to fault, distance to river and rainfall criteria). In other words, if similar high susceptible values of slope (or any other criteria) are prevailing all around a region while the landslide distribution pattern is represented by a different spatial order (Figure 5a), a secondary criterion (such as distance to river) with less availability may be the determinant factor of landslides’ spatial distribution (Figure 5c,e,f). This indicates the insight of the proposed objective weighting scheme in local evaluation of the landslide casual criteria. In other words, in the current study area, the slope angle is usually sufficient to influence landsliding. Nonetheless, considering the spatial distribution of landslides, there is limited evidence which proves that the slope criterion plays an important role in landsliding. In the present study area, susceptible slope values are distributed almost evenly over the study region; however, actual landslide events are more or less concentrated along the Karun Canyon. This may be due to the fact that the required water for slope failure, as a triggering factor, is controlling landslide events in southern Izeh. Water is not always directly involved as the transporting medium in mass movement processes while it does play an important role. This is not only proved by the obtained criteria weights of the Shannon entropy method but also it can be recognized by visual inspection of landslide spatial patterns and frequency along the Karun River (Figure 6).

![Histogram of estimated data driven (objective) and experts driven (subjective) landslide casual criteria weights for selected landslide criteria.](image)

Further, considering the results of our proposed objective weighting approach lithology criterion is the least important among all selected criteria, while the expert opinion refers to the rainfall layer as the least important. The achieved accuracy value of fuzzy Shannon entropy, however, is still remarkably superior.

5.3. Decision Aiding and Planning

Many researchers, [13,80,81] have pointed out that the traditional subjective weighting schemes usually suffer from sensitivity in decision-making and they are susceptible to intrinsic experts’ knowledge errors. Looking into the contribution to decision aiding, this study presents an integrated strategic weighting procedure using an objective method which determines the criteria weights by solving mathematical models. This is executed without any consideration of the decision maker’s preferences as it is a convention in subjective methods, such as the AHP method, OWA method, Delphi method, etc. In other words, this article introduces an objective approach that integrates fuzzy set theory and information theory algorithm (i.e., Shannon entropy), which could be a useful geospatial tool for integrating multiple features/attributes that affect the LSM process. This can largely compensate for the absence of expert DMs or the lack of local knowledge about study area when it comes to producing quality LSMs.
5.4. Limitation of the Proposed Methodology in LSM

While information theory-based methods such as the one proposed in the present research have shown considerable potential in different predictive spatial modelling scenarios, they do have their own limitations. Even though the application of the proposed methodology as an objective weighting scheme is not dependent on decision maker’s expertise and judgment, it relies on quantification of defined attributes of landslide data points using step by step mathematical computations. This is conditional on the existence of a concise and representative database. In terms of the present research, the availability of a comprehensive and readily accessible landslide inventory database was quite beneficial in achieving the desired outcome.

Another limitation of the implemented methodology is observable in particular in the NE part of the study region where false alarms exist in the form of low slope areas indicated as a high susceptibility class (very few pixels as a very high susceptible class). This is mainly due to the fact that the slope angle is not characterised as a primary criterion shaping the landslide occurrence spatial pattern. Most LSM approaches end up with extremely high false positive rates in terms of high or very high susceptible areas compared to the total landslide areas. This problem is not only limited to our study, therefore, we would like to call the attention of the physical geography community, in particular methodological development researchers, to exploring ways to reduce the problem of over-estimated susceptibility in future studies.

Further, after fitting the desired membership function, the proposed fuzzy Shannon entropy technique considers the dataset as a collection of distributions, which may not be suitable to extract specific spatial structures embedded in the underlying features/attributes [82]. Even though datasets with the same histogram certainly have the same entropy (i.e., distance to river and distance to fault in the present study), the distributions of their data values in space could be totally different. In addition, the result can be sensitive to the level of discretization caused by different membership functions (i.e., crisp or fuzzy) when using the histogram. We believe that further interest from researchers with access to larger data sample sizes is vital for developing more robust entropy-based LSM methods that can incorporate generalizable results.

6. Conclusions

This study showed promising results for GIS-MCDA tackling two major limitations. Firstly, the inherent subjectivity which emanates from decision makers’ (DMs’) preferences is diminished during the criteria standardization phase. Secondly, intrinsic biases and probable errors of DMs’ preferences corresponding to the subjective weighting approaches are also eliminated using the proposed LSM model. This LSM approach involves a thoughtful selection and elaborative standardization of landslide casual criteria while weighting procedures are accomplished using an objective method. This is performed by constructing a mathematical approach without any consideration of the DMs’ preferences from the beginning to the end of model implementation. Our results show that the integration of fuzzy sets with Shannon entropy can contribute to the production of landslide susceptibility maps with a reasonably high level of reliability. Finally, considering the fact that the proposed hybrid method has the advantage of objective weight evaluation, it can be used not only in similar areas of geo-hazard risk assessment and mapping, such as land subsidence, earthquake and flood risk mapping, but also in multi-hazard risk assessment for a further combination of risk elements. However, in order to apply the proposed objective weighting approach more generally by conducting different case studies, new hybrid models of GIS-based landslide susceptibility mapping need to be developed.

Acknowledgments: This research was supported by CSIRO Australian Sustainable Agriculture Scholarship (ASAS) as a top-up scholarship to Majid Shadman Roodposhti, a PhD scholar, at the University of Tasmania (RT109121), School of Land and Food. We thank three anonymous reviewers for their suggestions in improving the manuscript. Special thanks to Brett A. Bryan and Jeff Connor from CSIRO, Australia for their support.

Author Contributions: This paper was prepared using the contributions of all the authors. All authors have read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.
References

3. Ercanoglu, M.; Gokceoglu, C. Use of fuzzy relations to produce landslide susceptibility map of a landslide prone area (West Black Sea Region, Turkey). *Eng. Geol. 2004, 75*, 229–250. [CrossRef]
5. Atkinson, P.M.; Massari, R. Autologistic modelling of susceptibility to landsliding in the Central Apennines, Italy. *Geomorphol. 2011, 130*, 55–64. [CrossRef]


36. Hong, H.; Pradhan, B.; Xu, C.; Bui, D.T. Spatial prediction of landslide hazard at the Yihuang area (China) using two-class kernel logistic regression, alternating decision tree and support vector machines. *Catena* 2015, 133, 266–281. [CrossRef]


43. Davis, J.; Blissius, L. A hybrid physical and maximum-entropy landslide susceptibility model. *Entropy* 2015, 17, 4271–4292. [CrossRef]


47. Conoscenti, C.; Di Maggio, C.; Rotigliano, E. GIS analysis to assess landslide susceptibility in a fluvial basin of NW Sicily (Italy). Geomorphology 2008, 94, 325–339. [CrossRef]
54. Lotfi, F.H.; Fallahnejad, R. Imprecise Shannon’s entropy and multi attribute decision making. Entropy 2010, 12, 53–62. [CrossRef]
58. Lee, S.; Pradhan, B. Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models. Landslides 2007, 4, 33–41. [CrossRef]
60. Shahabi, H.; Ahmad, B.B.; Khezri, S. Evaluation and comparison of bivariate and multivariate statistical methods for landslide susceptibility mapping (case study: Zab basin). Arab. J. Geosci. 2013, 6, 3885–3907. [CrossRef]


71. Akgun, A.; Türk, N. Landslide susceptibility mapping for Ayvalik (Western Turkey) and its vicinity by multicriteria decision analysis. *Environ. Earth Sci.* **2010**, *61*, 595–611. [CrossRef]


