Dear Editor,

We would like to thank you and two anonymous reviewers for their helpful comments which helped us to improve the quality of the manuscript.

We have revised the manuscript GMD-2014-168 - entitled “**An integrated user-friendly ArcMAP tool for bivariate statistical modeling in geoscience applications**" to incorporate minor revision, strictly based on the reviewer's report.

We have revised the paper based on the feedback and comments given by the anonymous reviewers. As a result, you can see that there are changes in the new submission. All the comments and feedback have been taken very seriously and hence addressed them carefully in the revised manuscript. Please find the next page onwards of this document wherein we have answered to the issues raised by the referees in a point format. Additionally, I am uploading the manuscript with "track changes" in order to view those changes made during the new submission.

We sincerely expect that this revised manuscript can be published in “GMD”.

With best regards,

Prof. Dr. Biswajeet Pradhan

(Corresponding author)

Comments from First Reviewer

1. Same area has been studied previously by three authors of this study (MN Jebur, B Pradhan, MS Tehrany, 2014. Optimization of landslide conditioning factors using very high-resolution airborne laser scanning (LiDAR) data at catchment scale. Remote Sensing of Environment, 152, 150 – 165pp).

   Author’s response:

   Thank you for the comment. Actually, the main aim of this research is not related to the specific study area. The study area was selected in order to examine the proposed tool. Therefore, there was no need to test it in the new area because the main focus and contribution of this paper is the tool itself.

2. We can observe the behavior of this study area against the employed methods more or less from this previous one. The novelty of this study should be the developing of BSA as ArcMAp tool, but there are very limited details about this point.

Author’s response:

Thanks for the comment. We have provided some additional text which describes the novelty of this study which is the development of the proposed Arcmap tool. Additional section was added to the manuscript calls “2.2. Code description” to describe the code thoroughly. Moreover, the literature review has been modified and some information regarding GIS tools was added.

3. Pseudo codes of EBF, FR and WOE have not been presented.

Author’s response:

The code has been added as an appendix in the end of the manuscript. Moreover, the tool has been submitted with the modified edition.

4. Most of paper consists of the methodological details of BSA, experimental results and study area. To make complex methods more practically applicable on the natural hazards is an important point and deserves to work. This point of view makes this study interesting. However, authors preferred to present the tool as a detail and the method (BSA) as the major point.

Author’s response:

Thank you very much for your constructive comment. As it has been mentioned in comment 2, additional section was added to the manuscript calls “2.2. Code description” to describe the code thoroughly. Moreover, the literature review has been modified and some information regarding GIS tools was added.

5. The last point is that, there are too much self-citations in the references (almost 21 of total 43).

Author’s response:

Thank you for this observation. The references have been reformatted and the self-citation has been significantly reduced.

Comments from Second Reviewer

General Comments: Authors add bivariate statistical modules to ArcGIS using the python language. This GIS tool includes three models: evidential belief function (EBF), frequency ratio (FR) and weights-of-evidence (WoE). They are all data-driven techniques aiming to predict the hazard susceptibility and other geoscience applications. Such tool can assist geo-scientists in performing statistical analyses in the GIS environment since its procedure is simple and efficient. However this module could be improved by following:

(1) The Area under the curve (AUC) should be calculated in the ArcMAP tool so that three models can be compared directly in ArcGIS.

Author’s response:

Thanks for the comment. This paper aims to deal with the bivariate analysis only. The AUC is one of the various available techniques for validating the results. We have used this validation technique to check the efficiency of FR, WoE, and EBF algorithm in the particular application. Therefore, creating AUC would be out of the scope of this paper.

(2) It would be better to provide the functions that could select the ratio of training set to the testing set. It will be convenient for users to test the models.

Author’s response:

Actually, this would be valuable if the tool create the training and testing or doing the sampling itself. However, the aim of this paper is to run the bivariate techniques in GIS after the training and testing layer are created. Moreover, selecting the ratio depends of the quantity of the samples and the application. Therefore, the authors did not include them in the tool.

(3) It would be better to provide functions for testing the modeling results. For example, the Bootstrapping (statistics) approache can be adopted. This method allows assigning measures of accuracy (defined in terms of bias, variance, confidence intervals, prediction error or some other such measure) to sample estimates. It will facilitate selecting the model and judging the modeling results judge.

Author’s response:

Judging the modeling result may be varying from one application to another. The main point here is that, authors didn’t propose new method for performing susceptibility or other mapping. We programmed the BSA algorithms which can be used in numerous modeling. Hence, the accuracy of the results that the user can attain is only related to the data used, nature of the study area and precision of their original data. For the users who attempt to use this tool, a full understanding of the provided bivariate technique should be attained before using the tool. To do such, a literature review is required; moreover, running the success rate using AUC algorithm would help for judging the modeling results judge.

Specific comments: There are also some typos or grammar problems in the MS.

For example: (1) On page 7241, line 18, ‘is’ should be deleted.

Author’s response:

Thanks, correction is made.

(2) On page 7246, line 25, the equation (2) was wrong.

Author’s response:

Thanks for the comment. Amended as suggested. Moreover, the manuscript was reviewed by the authors and many errors were corrected with track change and it can be seen in the revised version.

An integrated user-friendly ArcMAP tool for bivariate statistical modeling in geoscience applications

Mustafa Neamah Jebur, Biswajeet Pradhan*, Helmi Zulhaidi Mohd Shafri, Zainuddin Md. Yusoff, Mahyat Shafapour Tehrany

Department of Civil Engineering, Geospatial Information Science Research Center (GISRC),
Faculty of Engineering,

University Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia,

Tel. +603-89466383; Fax. +603-89468470

Email. biswajeet24@gmail.com or biswajeet@lycos.com (corresponding author)

Abstract

Modeling and classification difficulties are fundamental issues in natural hazard assessment. A geographic information system (GIS) is a domain that requires users to use various tools to perform different types of spatial modeling. Bivariate statistical analysis (BSA) assists in hazard modeling. To perform this analysis, several calculations are required and the user has to transfer data from one format to another. Most researchers perform these calculations manually by using Microsoft Excel or other programs. This process is time consuming and carries a degree of uncertainty. The lack of proper tools to implement BSA in a GIS environment prompted this study. In this paper, a user-friendly tool, BSM (Bivariate statistical modeler), for BSA technique is proposed. Three popular BSA techniques such as frequency ratio, weights-of-evidence, and evidential belief function models are applied in the newly proposed ArcMAP tool. This tool is programmed in Python and ~~is~~ created by a simple graphical user interface, which facilitates the improvement of model performance. The proposed tool implements BSA automatically, thus allowing numerous variables to be examined. To validate the capability and accuracy of this program, a pilot test area in Malaysia is selected and all three models are tested by using the proposed program. Area under curve is used to measure the success rate and prediction rate. Results demonstrate that the proposed program executes BSA with reasonable accuracy. The proposed BSA tool can be used in numerous applications, such as natural hazard, mineral potential, hydrological, and other engineering and environmental applications.

Keywords: ArcMAP tool, Bivariate statistical analysis, Geographic information systems

1 Introduction

Techniques to predict a response variable given a set of characteristics are required in several scientific regularities. Numerous applications have been implemented in various areas of geosciences. Bivariate analysis is one of the simplest methods of statistical analysis, and is popular in numerous fields of study. Mathematicians, statisticians, biologists, and hydrologists use this method to perform their analysis. Different types of bivariate statistical analysis (BSA) have been established, for example, frequency ratio (FR), weights-of-evidence (WoE), and evidential belief function (EBF) (Yalcin, 2008). Although each of these methods requires specific mechanisms for calculation, all of these methods operate by using the same concept. Environmental scientists model various natural conditions by using the BSA statistical method. For instance, Ozdemir (2011) employed this technique for the same purpose. The results of the analysis were plotted in ArcGIS after computation in other programs. Mineral potential mapping is also aided by BSA techniques. Carranza (2004) used WoE modeling to map the mineral potential in the administrative province of Abra in northwestern Philippines. Their achievements indicate the plausibility of WoE in the mineral potential mapping of large areas with a small number of mineral prospects. Researchers have applied WoE in mapping mineral potential (Bonham-Carter et al., 1989) and remains popular in this area of research (Carranza et al., 2008).

BSA is in demand in hazard studies because its procedure is simple and efficient. This technique is has been used in natural hazard applications by researchers to predict the spatial distribution of events. Extensive literature on different BSA techniques and their proficiency assessment are also available. BSA techniques can be used as a simple geospatial analysis tool to determine the probabilistic correlation among dependent variables (produced by using the inventory map of a hazard incidence) and independent variables (conditioning factors) containing multi-categorized maps (Oh et al., 2011). In BSA, the overlay of conditioning factors and computation of hazard densities, the significance of each factor, or the particular mixture of factors can be investigated individually. Bivariate statistical analysis functions by using a dependent variable and one conditioning factor. Hence, the significance of each factor is investigated separately (Porwal et al., 2006).

In BSA, each conditioning factor is overlaid with the dependent variable. On the basis of the event density, weights are measured for each class of each factor. By using normalized weights (the correlation between the event density in each class of conditioning factor and the event density of the entire region), each conditioning factor is reclassified and the hazard map is produced. By using the acquired weights, decision rules can be produced on the basis of the knowledge of experts. Conditioning factors can also be combined to generate a map with uniform units, which is then overlaid with the inventory map to provide the density per class. The BSA approach has been used in landslide mapping (Constantin et al., 2011), earthquake studies (Xu et al., 2012b), flood susceptibility mapping (Tehrany et al., 2013), land subsidence (Kim et al., 2006; Lee and Park, 2013), and risk analysis (Hu et al., 2009). Numerous studies have been conducted to exploit the potential application of BSA in the hazard domain.

This research examined the efficiency of statistical analysis, particularly bivariate analysis, in landslide studies in the Cuyahoga River watershed (Nandi and Shakoor, 2010). In another study, FR and WoE were applied in the Sultan Mountains of southwestern Turkey to map areas that are susceptible to landslides (Ozdemir and Altural, 2013). According to Nandi and Shakoor (2010) and Ozdemir and Altural (2013), the BSA model is simple and its input, computation, and outcome procedures are effortlessly understood. The application of EBF in the area of landside studies has been investigated (Lee et al., 2013). Four functions, namely, degree of belief (Bel), degree of disbelief (Dis), degree of uncertainty (Unc) and degree of plausibility (Pls), are calculated separately to determine EBF.

Each of these functions produces valuable information. However, each function requires individual computations with specific formulas. Tien Bui et al. (2012) used EBF and fuzzy logic methods in their research and found that the landslide susceptibility map derived from EBF has the highest prediction ability. They also established the efficiency of BSA in landslide mapping. BSA is also popular in hydrological research. Flood susceptibility maps assist in mitigation strategies. Lee et al. (2012) used the statistical method of FR to produce a map of flood-prone regions in Busan, Korea, in GIS. Tehrany et al. (2013) proposed an ensemble method of FR and LR to detect regions with high flood probability in Kelantan, Malaysia. The conditioning factors were reclassified on the basis of the weights acquired from the FR technique. These factors were entered in LR processing to obtain the MSA result. If the calculation time for these statistics can be reduced, the efficiency of the developed ensemble method will be enhanced. Hence, producing a tool that is capable of performing BSA calculations will help reduce the calculation time of ensemble methods.

The BSA model has been widely used in land subsidence susceptibility mapping. In a study by Lee and Park (2013), the FR model was applied and compared with the machine learning of DT. The BSA is a method that is commonly used in natural hazard investigations. Although this method is not novel, the use of BSA has increased in recent years. RS and GIS have revolutionized the domain of natural hazards (Jebur et al., 2013a; Jebur et al., 2013b). Spatial database consists of different data types that are required to be transferred from one format to another because specific programs accept only specific data formats. Scientists have started to develop new programs in hazard studies because of the vital role of early warning systems in such applications (Osna et al., 2014; Pradhan et al., 2014). GIS is capable to store, analyze and show geographic information. It makes it possible to collect, organize, explore, model and view the spatial data for solving ~~the difficult scheduling and managing~~complex problems (Barreca et al., 2013). Different ~~sorts~~types of spatial data analysis ranges from the simple overlaying of various thematic layers to identify the region to the more complex use of mathematical equations or combined statistical models for the prediction of ~~the~~ natural hazards.~~ In the case that GIS used in land use organization and natural resource managing and safety, is a tool that is able to support scientific study and decision making.~~ The importance of GIS in catastrophic evaluation was proven by many studies related to the GIS tools usages in exploration of various types of data (Steiniger and Hunter, 2013).

~~Currently, created~~For example the existing hydrological GIS-based tools such as ~~for instance~~ Mike SHE and ArcSWAT revealed considerable power in enhancing the accuracy of soil and water evaluations (Lei et al., 2011). These tools are capable of facilitating the modeling and calibration procedure, and decreasing the stages in implementing the models and increasing the precision of the outcomes (Hörmann et al., 2009). The creation of tools that automatically implement susceptibility mapping was applied by Akgun et al. (2012). Akgun et al. (2012) proposed “MamLand,” a program in MATLAB, to create landslide susceptibility mapping by using fuzzy inference system. ArcGIS allows users to produce specific tools for spatial analysis (Stevens et al., 2007). For instance, Pradhan et al. (2014) developed a tool in ArcGIS to apply texture analysis for high-resolution radar data.

Recently, a~~A~~ GIS-based system has been developed by Barreca et al. (2013) to evaluate and process the hazard associated to active faults influencing the eastern and southern flanks of Mt. Etna. The proposed tool was created in ArcGIS which contains ~~of~~ various thematic datasets. It includes spatially-referenced arc-features and associated database. In another paper, Lei et al. (2011), ~~researchers from China,~~ integrated hydrological code EasyDHM and proposed open source MapWindow GIS tool called MWEasyDHM. Their aim was to create the tool by combining modules ~~of~~for preprocessing, modeling, viewing and analysis. MWEasyDHM tool is user friendly, free and proficient which produces selectable multi-functional hydrological analysis. Similarly, a~~A~~ number of GIS tools are programmed by Etherington (2011) in Python environment ~~to be used in~~for landscape genetics researches. Tools are capable to transform files, view genetic relatedness, and calculate landscape associations through least-cost path procedure. The tools are free and available in ArcToolbox. In a separate paper, Roberts et al. (2010) implemented the research to facilitate the advanced analytic methods. ~~Therefore,~~A Marine Geospatial Ecology Tools (MGET) was created in GIS environment which ~~are~~is free, easy to use and efficient tools for the ecologists. The tools were made by integrating different strong programming methods of Python, R, MATLAB, and C++.

The current research aims to reduce the processing time of BSA by introducing an easy-to-use ArcMap tool. On the basis of the aforementioned problem statement regarding the required processing time and difficulties for BSA, a program that is capable of calculating BSA automatically should be developed. Hence, a tool programmed in Python and based on the BSA technique is proposed. This tool automatically extracts the correlation among each class of conditioning factor and event occurrence, reclassifies the factors on the basis of the acquired weights in a GIS environment, and saves each correlation in separate folders. A simple graphical user interface (GUI) improves the model operation because Python knowledge is not required. The entire process can be performed in ArcGIS without any requirement for another program. The proposed tool was tested to generate a landslide susceptibility map of Bukit Antarabangsa, Ulu Klang, Malaysia.

2 Methodology

The procedural and theoretical perspectives of BSA applied in this research include several steps (Fig. 1). In the methodology flowchart, the BSA tool was developed and integrated into ArcGIS. To apply BSA, the conditioning factors should be provided in raster format and classified with the proper scheme by the user. The BSA recognizes the effects of each class of conditioning factor on event occurrence. Hence, this step cannot be eliminated in the BSA process. As a second stage, a dependent variable (training layer) should be constructed by using the inventory map and other resources. This layer should contain a pixel value of one to represent the existence of an event. Once the conditioning factors are classified and the training layers are prepared, FR, WoE, and EBF can be applied automatically. The developed program reclassifies each conditioning factor by using the attained weights and saves them in a separate folder. The group of conditioning factors that have been assessed by BSA are ready to be entered in the raster calculator to derive the corresponding hazard map. The following sub-section represents the overall information on the scheme and functionality of the developed tool.

Fig. 1. About here

2.1 Overall information on scheme and functionality

The program is developed by using ArcGIS and Python for BSA. The tool can be used in ArcGIS 9 and 10 versions. Fig. 2 displays the interface of the tool in GIS toolbox.

Fig. 2. About here

The ArcToolbox provided in this research is used to enter the proposed tool in ArcMap. The user defines the source of the Python files of each model from the properties menu of the script (Fig. 3).

Fig. 3. About here

The program is partitioned into three sections: FR, WoE, and EBF. The theoretical concept and graphic interface of each tool is discussed in the following sections.

2.1.1 FR

The theoretical expression of FR, as well as its usage in landslide susceptibility and flood mapping has been reported in the studies conducted by Yilmaz (2009) and Tehrany et al. (2013) respectively. The FR method has a simple and understandable structure compared with other probabilistic methods. FR is described as the proportion of the region where an event occurred over the entire area; FR is also defined as the proportion of likelihood of an event occurrence to a nonoccurrence for a particular attribute. FR can be calculated by using the following equation:

$$FR = \frac{\frac{N_{pix}(SX_i)}{\sum_{i=1}^{m} SX_i}}{\frac{N_{pix}(X_j)}{\sum_{j=1}^{n} N_{pix}\,(X_j)}} \tag{1}$$

where $N_{pix}(SX_i)$ is the number of pixels, which contain an event in class i of the independent variable; X, $N_{pix}(X_j)$ is the number of pixels and exist in independent variables X_j; m is the number of categoris of the independent variable X_i. Furthermore, n is the total number of independent variables in the whole area (Yilmaz, 2009). Most of the researchers performed these calculations manually by using Microsoft Excel or other programs. Once the weights were obtained, these values were used to reclassify the independent variables by using the spatial analyst tool in ArcGIS. The raster calculator in ArcGIS was used to obtain the final susceptibility map. The proposed tool in ArcMap can apply the FR automatically and reclassify the independent variables on the basis of the gained weights.

The graphic interface of the FR tool consists of one window containing four fields (Fig. 4). Each field is user-defined in ArcGIS. The first field is the input raster, which is related to the desired conditioning factor. The training layer or dependent variable, which is predefined and saved prior to analysis, is selected for the second field. The cell size of the output and its location are specified by the user in the third and fourth fields, respectively. The developed tool has a simple structure, thus providing BSA for each conditioning factor within a few seconds. In manual calculations, this procedure usually requires considerable amount of time to be implemented. The proposed tool reclassifies the analyzed conditioning factor based on the attained weights and saves it in the selected folder by the user.

Fig. 4. About here

2.1.2 WoE

The WoE method is a data-driven technique based on the Bayesian probability framework (Beynon et al., 2000; Neuhäuser and Terhorst, 2007; Porwal et al., 2006). This characteristic provides additional advantages to the proposed tool compared with other statistical methods. To implement WoE, two important parameters of positive weight (W^{+}) and negative weight (W^{-}) are computed (Bonham-Carter et al., 1989). This technique calculates the weight for each independent variable (B) on the basis of the existence or non-existence of the event (A) within the study area (Xu et al., 2012a) by using the following equations:

$$W_i^{+} = \ln \frac{P\{B|A\}}{P\{B|\bar{A}\cancel{A}\}} \tag{2}$$

$$W_i^{-} = \ln \frac{P\{\bar{B}|A\}}{P\{\bar{B}|\bar{A}\}} \tag{3}$$

where P represents the probability, In is the natural $log.$ B , and $\bar{B}$ reveals the existence and nonexistence of the independent variable. A and $\bar{A}$ show the existence and nonexistence of the event. A positive weight (W^{+}) determines the presence of the specific independent variable at the event, and the amount of positive weight represents the positive correlation between the presence of the independent variable and event, respectively. A negative weight (W^{-}) indicates the nonexistence of the independent variable and shows the amount of negative correlation.

The weight contrast is the difference between the two weights of W^{+} and W^{-}:

$$C(C = W^{+} - W^{-}) \qquad (6)$$

The size of the weight contrast demonstrates the spatial relationship between the independent variable and the event. The C value is positive in the case of a positive relationship and is negative in the case of a negative relationship.

The standard deviation of W is calculated as follows:

$$S(C) = \sqrt{S^{2}W^{+} + S^{2}W^{-}} \qquad (7)$$

where $S(W^{+})$ and $S(W^{-})$ are the variance of the positive and negative weights, respectively. These variances can be calculated by using the following equations:

$$S^{2}W^{+} = \frac{1}{N\{B \cap A\}} + \frac{1}{B \cap \bar{A}} \qquad (8)$$

$$S^{2}W^{-} = \frac{1}{N\{\bar{B} \cap A\}} + \frac{1}{\{\bar{B} \cap \bar{A}\}} \qquad (9)$$

By using the proportion of the contrast divided by its standard deviation, the studentized contrast is calculated. The studentized contrast is the final weight that assists the informal test if C is considerably different from zero or if the contrast is probable to be “real.” A complete explanation of the mathematical formulation of this method is accessible in Xu et al. (2012b). Fig. 5 illustrates the user interface of the WoE tool. Each field should be defined similar to FR.

Fig. 5. About here

2.1.3 EBF

Dempster (1967) is an innovator who presented the Dempster–Shafer theory of evidence, which is known as a generalized Bayesian theory of subjective probability. This theory has been used in several fields of study, including environmental and hazard studies (Awasthi and Chauhan, 2011). This theory also has relative flexibility, which is considered its advantage, accepts uncertainty, and is capable of combining beliefs from different sources of evidence. EBF estimates the probability that a hypothesis is true and evaluates how close the evidence is to the truth of that hypothesis. A complex procedure is required to calculate EBF compared with FR. To compute the EBF, four functions (Bel, Dis, Unc, and Pls) should be measured separately (Lee et al., 2013). Individual computation by using specific formulas is required to provide this information.

Assume that a set of independent variables of $\mathrm{C} = (\mathrm{C_i}, \mathrm{i} = 1, 2, 3, \ldots, \mathrm{n})$, which contains mutually exclusive and exhaustive factors of $\mathrm{C_i}$, is used in current research. The function $\mathrm{m}\colon \mathrm{P(C)} \rightarrow [0,1]$ is the basic of the probability assignment.

$$\mathrm{Bel}(\mathrm{C_{ij}}) = \frac{\mathrm{W_{C_{ij}(event)}}}{\sum_{j=1}^{n} \mathrm{W_{C_{ij}(event)}}}, \qquad (6)$$

where C is the frame of discernment and $\mathrm{P(C)}$ is the set of all subsets of C, counting the empty set (Φ) and C itself. Mass function is another name for the mentioned function that satisfies $\mathrm{m}(\Phi) = 0$ and $\sum_{\mathrm{AC}} \mathrm{m}\ (\mathrm{A}) = 1$, where A is any subset of C. The degree in which the evidence support A is calculated by $\mathrm{m(A)}$, which is represented by a belief function (Bel(A)). Suppose that $\mathrm{N(L)}$ and $\mathrm{N(C)}$ are the total number of pixels affected by the event and the total number of pixels in the study area, respectively; $\mathrm{C_{ij}}$ is the j-th class of the independent variable of $\mathrm{C_i}(\mathrm{i} = 1, 2, 3, \ldots, \mathrm{n})$; $\mathrm{N(C_{ij})}$ is the total number of pixels in class $\mathrm{C_{ij}}$; and $\mathrm{N} = (\mathrm{L} \cap \mathrm{C_{ij}})$ is the number of pixels affected by the event in $\mathrm{C_{ij}}$. Therefore, the data-driven measurement of EBF can be calculated by using the following equation (Tien Bui et al., 2012):

where the $\mathrm{C_{ij}}$ is shown by $\mathrm{W_{C_{ij}(event)}}$ and supports the belief that the presence of the event is more than its nonexistence. The detailed mathematical calculation of each function has been discussed in several studies, such as Lee et al. (2013). Fig. 6 represents the interface of the EBF tool, and contains three more fields compared with the two other methods because each EBF function should be applied and saved in a separate folder. Hence, after the selection of the conditioning factor, training layer, and output cell size, the location to save each function should be defined.

Fig. 6. About here

2.2 Code description

The code was designed in python 27 (The default software included with windows 7). In the beginning, the arcpy library is called to check the ~~and the~~ code for ~~checks the~~ spatial extension

~~which should be available~~in order to continue the process. After that, when the user defines the raster, the code calls the raster data as test using the command “GetParameterAsText” which is part of arcpy library. Using same as the previous command, the code will define the output layer for the chosen model. The default path for all the sub process is defined to be in “C” drive because it is the default drive in all the systems. Therefore, the code creates folder calls “"FR_modeler”, "WOE_modeler", or "EBF_modeler" depending on~~f~~ the selected process.

The next stage is to analyze the input layer (e.g. Slope) and “Lookup” command will be applied to prepare the layer for zonal geometry process. The zonal geometry is defined as table to be able to work on the statistic of the output. A filed is added to the attribute of the created table in the previous step namely “zonal” to be used for calculating the percentage of each class of the input layer. A statistics analysis was applied to calculate the sum of all the pixels of the selected layer. Then, a joining process is defined to link the crea~~a~~ted table with the input layer. Subsequently, ~~T~~tabulate area process was ran to calculate the percentage of the occurrence of the independent factor (i.e. Landslide) in each input layer classes. The last step for calculating FR is applied using eq.1. Then, the resulted values is defined as integer and used to reclassify the input layer. The code includes a delete command to delete all the sub process layers and table.

The process of WoE and EBF contains the same process of FR as initial step. However, more statistical analysis and more field are added to calculate the parameters of WoE and EBF which is listed in eq. (2-6). In each selected model, a different folder will be created. The user may overwrite and redo the process as much as required because the command “overwriteOutput” was defined for each code. The three codes is added as appendix 1.

~~2.2~~2.3 Test area and data

Although the developed program can be used in any application that employs BSA, the proficiency of the tool was tested in the hazard domain. To examine the capability and efficiency of the developed program, landslide susceptibility analyses were performed by using the developed ArcMAP tool with three BSA models, namely, FR, WoE, and EBF. The program was tested for the landslide susceptibility mapping of Bukit Antarabangsa, Ulu Klang, Malaysia (Fig. 7).

Fig. 7. About here

A spatial database was constructed and analyzed on the basis of the altitude, aspect, curvature, slope, stream power index, topographic wetness index, distance from the river, distance from the road, and geological layers. Comprehensive overview of the usage of BSA for landslide susceptibility mapping has been reported in numerous studies (Yalcin et al., 2011). Study conducted by Mohammady et al. (2012) provided additional knowledge on the capabilities of these three BSA methods. These previous research compared the three methods of FR, WoE, and EBF and determined the pros and cons of each statistical approach. A total of 47 landslide locations were recorded and a landslide inventory map was prepared. The allocation of the landslide inventory for training and testing was 70% and 30%, respectively (Fig. 7). The training data set (31 landslide locations) was chosen randomly and a dependent layer (landslide layer) was created.

3 Experimental results and discussion

To examine the efficiency of the developed BSM tool, landslide susceptibilities were derived by using all three methods. The correlation among the conditioning factors and landslide occurrence was extracted. The landslide probability index was measured and classified by using the proper scheme. To produce a susceptibility map, the probability index should be partitioned into various classes. The quantile method was applied in the current research because of its reputation in classification. In the quantile classification method, each class has the same number of features. This method has been employed by several researchers, such as Umar et al. (2014) and Papadopoulou-Vrynioti et al. (2014). The method provided appropriate results on the comparison between the created landslide susceptibility map and the spatial distribution of landslide events. The acquired landslide conditioning factors is shown in Fig. 8.

Fig. 8. About here

The derived landslide susceptibility map from WoE shows a different appearance compared with the two other maps. Validation should be performed to determine which map is reliable. The area under curve (AUC) was applied to examine the precision of the derived susceptibility maps (Pérez-Vega et al., 2012). The success rate values were 68%, 63%, and 76% for FR, WoE, and EBF, respectively. Moreover, 71%, 75%, and 80% were the prediction rates for FR, WoE, and EBF, respectively. The EBF represented the highest accuracy compared with other methods in terms of success and prediction rates. The prediction rate value for WoE was high but not as high as EBF. This result is caused by the greater proficiency and capability of EBF compared with WoE. Recognizing the best method for modeling is possible because any comparative study is restricted and the best method for a specific data set is significantly related to the characteristics of that dataset. Fig. 9 illustrates the computed accuracies.

Fig. 9. About here

The design and interface of the developed tool show that the BSA is simple to execute by using the proposed program compared with manual calculation. The derived susceptibility maps and their AUC values suggest that the tool is precise and reliable. Previous research has established that because of the nature of BSA, the obtained results are imprecise compared with machine learning and rule-based methods. Therefore, the measured accuracies are acceptable for these simple statistical methods.

4 Conclusion

To perform hazard studies, several requirements, such as constructing the precise spatial database, obtaining high-resolution imagery, and providing a reliable inventory map, should be fulfilled. Users can be confronted with the insufficiency of appropriate and free tools to perform various analyses. This condition makes such studies complex and in some cases, time consuming. The BSA is one of the fundamental methods in hazard mapping. Hence, developing a tool that manages a large number of factors with an automatic statistical and classification performance is essential. Users commonly have to apply the BSA calculation manually and within separate software. The results have to be entered in a GIS environment and used to reclassify each conditioning factor one after another. The proposed BSM tool can be used to automate the BSA procedure and to facilitate the generation of the probability index. BSM is developed as a tool in ArcGIS, which is capable of performing the three BSA models of FR, WoE, and EBF. This tool can also manage large amounts of conditioning factors with reduced calculation time, thus allowing the replication of various trials. As an example, a significant characteristic of BSM is the reclassification of the conditioning factors on the basis of the acquired weight from BSA. The GUI also allows the application of RF, WoE, and EBF without entering any code from Python, thus helping the user in model operation. The application to landslide susceptibility mapping in Bukit Antarabangsa in Ulu Klang, Malaysia provides significant outcomes. All three methods are applied and landslide susceptibility maps are created. FR, WoE, and EBF acquired success rates of 68%, 63%, and 76%, respectively. AUC values for prediction rates are 71%, 75%, and 80% for FR, WoE, and EBF, respectively. In conclusion, the proposed tool can transform the BSA procedure into a simple and fast technique. This tool can assist scientists in performing statistical analyses for any environment and mathematical application.

References

Akgun, A., Sezer, E. A., Nefeslioglu, H. A., Gokceoglu, C., Pradhan, B.: An easy-to-use MATLAB program (MamLand) for the assessment of landslide susceptibility using a Mamdani fuzzy algorithm, Comput. Geosci., 38, 23-34, 2012.

Awasthi, A., Chauhan, S. S.: Using AHP and Dempster–Shafer theory for evaluating sustainable transport solutions, Environ. Modell. Softw., 26, 787-796, 2011.

Barreca, G., Bonforte, A., Neri, M.: A pilot GIS database of active faults of Mt. Etna (Sicily): A tool for integrated hazard evaluation, J. Volcanol. Geoth. Res., 251, 170-186, 2013.

Beynon, M., Curry, B., Morgan, P.: The Dempster–Shafer theory of evidence: an alternative approach to multicriteria decision modelling, Omega., 28, 37-50, 2000.

Bonham-Carter, G. F., Agterberg, F. P., Wright, D. F.: Weights of evidence modelling: a new approach to mapping mineral potential, Comput. Geol., 89, 171–183, 1989.

Carranza, E. J. M.: Weights of evidence modeling of mineral potential: a case study using small number of prospects, Abra, Philippines, Nat. Resour. Res., 13, 173-187, 2004.

Carranza, E. J. M., Van Ruitenbeek, F., Hecker, C., van der Meijde, M., van der Meer, F. D.: Knowledge-guided data-driven evidential belief modeling of mineral prospectivity in Cabo de Gata, SE Spain, Int. J. Appl. Earth. Obs., 10, 374-387, 2008.

Constantin, M., Bednarik, M., Jurchescu, M. C., Vlaicu, M.: Landslide susceptibility assessment using the bivariate statistical analysis and the index of entropy in the Sibiciu Basin (Romania), Environ. Earth. Sci., 63, 397-406, 2011.

Dempster, A. P.: Upper and lower probabilities induced by a multivalued mapping, Ann. Math. Stat., 325-339, 1967.

Etherington, T. R.: Python based GIS tools for landscape genetics: visualising genetic relatedness and measuring landscape connectivity, Methods. Ecol. Evol., 2, 52-55, 2011.

Hörmann, G., Köplin, N., Cai, Q., Fohrer, N.: Using a simple model as a tool to parameterise the SWAT model of the Xiangxi river in China, Quatern. Int., 208, 116-120, 2009.

Hu, B., Zhou, J., Wang, J., Chen, Z., Wang, D., Xu, S.: Risk assessment of land subsidence at Tianjin coastal area in China, Environ. Earth. Sci., 59, 269-276, 2009.

Jebur, M. N., Pradhan, B., Tehrany, M. S.: Detection of vertical slope movement in highly vegetated tropical area of Gunung pass landslide, Malaysia, using L-band InSAR technique, Geosci. J., 18, 61-68, 2013a.

Jebur, M. N., Pradhan, B., Tehrany, M. S.: Using ALOS PALSAR derived high-resolution DInSAR to detect slow-moving landslides in tropical forest: Cameron Highlands, Malaysia, Geomat. Nat. Hazards. Risk., 1-19, doi:10.1080/19475705.2013.860407, 2013b.

Kim, K. D., Lee, S., Oh, H. J., Choi, J. K., Won, J. S.: Assessment of ground subsidence hazard near an abandoned underground coal mine using GIS, Environ. Geol., 50, 1183-1191, 2006.

Lee, M. J., Kang, J. e., Jeon, S.: Application of frequency ratio model and validation for predictive flooded area susceptibility mapping using GIS, in: IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Munich, 895-898, 2012.

Lee, S., Hwang, J., Park, I.: Application of data-driven evidential belief functions to landslide susceptibility mapping in Jinbu, Korea, Catena., 100, 15-30, 2013.

Lee, S., Park, I.: Application of decision tree model for the ground subsidence hazard mapping near abandoned underground coal mines, J. Environ. Manage., 127, 166-176, 2013.

Lei, X., Wang, Y., Liao, W., Jiang, Y., Tian, Y., Wang, H.: Development of efficient and cost-effective distributed hydrological modeling tool MWEasyDHM based on open-source MapWindow GIS, Comput. Geosci., 37, 1476-1489, 2011.

Mohammady, M., Pourghasemi, H. R., Pradhan, B.: Landslide susceptibility mapping at Golestan Province, Iran: a comparison between frequency ratio, Dempster–Shafer, and weights-of-evidence models, J. Asian. Earth. Sci., 61, 221-236, 2012.

Nandi, A., Shakoor, A.: A GIS-based landslide susceptibility evaluation using bivariate and multivariate statistical analyses, Eng. Geol., 110, 11-20, 2010.

Neuhäuser, B., Terhorst, B.: Landslide susceptibility assessment using “weights-of-evidence” applied to a study area at the Jurassic escarpment (SW-Germany), Geomorphology., 86, 12-24, 2007.

Oh, H. J., Kim, Y. S., Choi, J. K., Park, E., Lee, S.: GIS mapping of regional probabilistic groundwater potential in the area of Pohang City, Korea, J. Hydrol., 399, 158-172, 2011.

Osna, T., Sezer, E. A., Akgun, A.: GeoFIS: An Integrated Tool for the Assessment of Landslide Susceptibility, Comput. Geosci., 66, 20-30, 2014.

Ozdemir, A.: Using a binary logistic regression method and GIS for evaluating and mapping the groundwater spring potential in the Sultan Mountains (Aksehir, Turkey), J. Hydrol., 405, 123-136, 2011.

Ozdemir, A., Altural, T.: A comparative study of frequency ratio, weights of evidence and logistic regression methods for landslide susceptibility mapping: Sultan Mountains, SW Turkey, J. Asian. Earth. Sci., 64, 180-197, 2013.

Pérez-Vega, A., Mas, J. F., Ligmann-Zielinska, A.: Comparing two approaches to land use/cover change modeling and their implications for the assessment of biodiversity loss in a deciduous tropical forest, Environ. Modell. Softw., 29, 11-23, 2012.

Porwal, A., Carranza, E., Hale, M.: Bayesian network classifiers for mineral potential mapping, Comput. Geosci., 32, 1-16, 2006.

Pradhan, B., Hagemann, U., Shafapour Tehrany, M., Prechtel, N.: An easy to use ArcMap based texture analysis program for extraction of flooded areas from TerraSAR-X satellite image, Comput. Geosci., 63, 34-43, 2014.

Roberts, J. J., Best, B. D., Dunn, D. C., Treml, E. A., Halpin, P. N.: Marine Geospatial Ecology Tools: An integrated framework for ecological geoprocessing with ArcGIS, Python, R, MATLAB, and C++, Environ. Modell. Softw., 25, 1197-1207, 2010.

Steiniger, S., Hunter, A. J.: The 2012 free and open source GIS software map–A guide to facilitate research, development, and adoption, Comput. Environ. Urban., 39, 136-150, 2013.

Stevens, D., Dragicevic, S., Rothley, K.: < i> iCity</i>: A GIS–CA modelling tool for urban planning and decision making, Environ. Modell. Softw., 22, 761-773, 2007.

Tehrany, M. S., Pradhan, B., Jebur, M. N.: Spatial prediction of flood susceptible areas using rule based decision tree (DT) and a novel ensemble bivariate and multivariate statistical models in GIS, J. Hydrol., 504, 69-79, 2013.

Tien Bui, D., Pradhan, B., Lofman, O., Revhaug, I., Dick, O. B.: Spatial prediction of landslide hazards in Hoa Binh province (Vietnam): a comparative assessment of the efficacy of evidential belief functions and fuzzy logic models, Catena., 96, 28-40, 2012.

Xu, C., Xu, X., Dai, F., Xiao, J., Tan, X., Yuan, R.: Landslide hazard mapping using GIS and weight of evidence model in Qingshui river watershed of 2008 Wenchuan earthquake struck region, J. Earth. Sci., 23, 97-120, 2012a.

Xu, C., Xu, X., Lee, Y. H., Tan, X., Yu, G., Dai, F.: The 2010 Yushu earthquake triggered landslide hazard mapping using GIS and weight of evidence modeling, Environ. Earth. Sci., 66, 1603-1616, 2012b.

Yalcin, A.: GIS-based landslide susceptibility mapping using analytical hierarchy process and bivariate statistics in Ardesen (Turkey): comparisons of results and confirmations, Catena, 72, 1-12, 2008.

Yalcin, A., Reis, S., Aydinoglu, A., Yomralioglu, T.: A GIS-based comparative study of frequency ratio, analytical hierarchy process, bivariate statistics and logistics regression methods for landslide susceptibility mapping in Trabzon, NE Turkey, Catena., 85, 274-287, 2011.

Yilmaz, I.: Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: a case study from Kat landslides (Tokat—Turkey), Comput. Geosci., 35, 1125-1138, 2009.

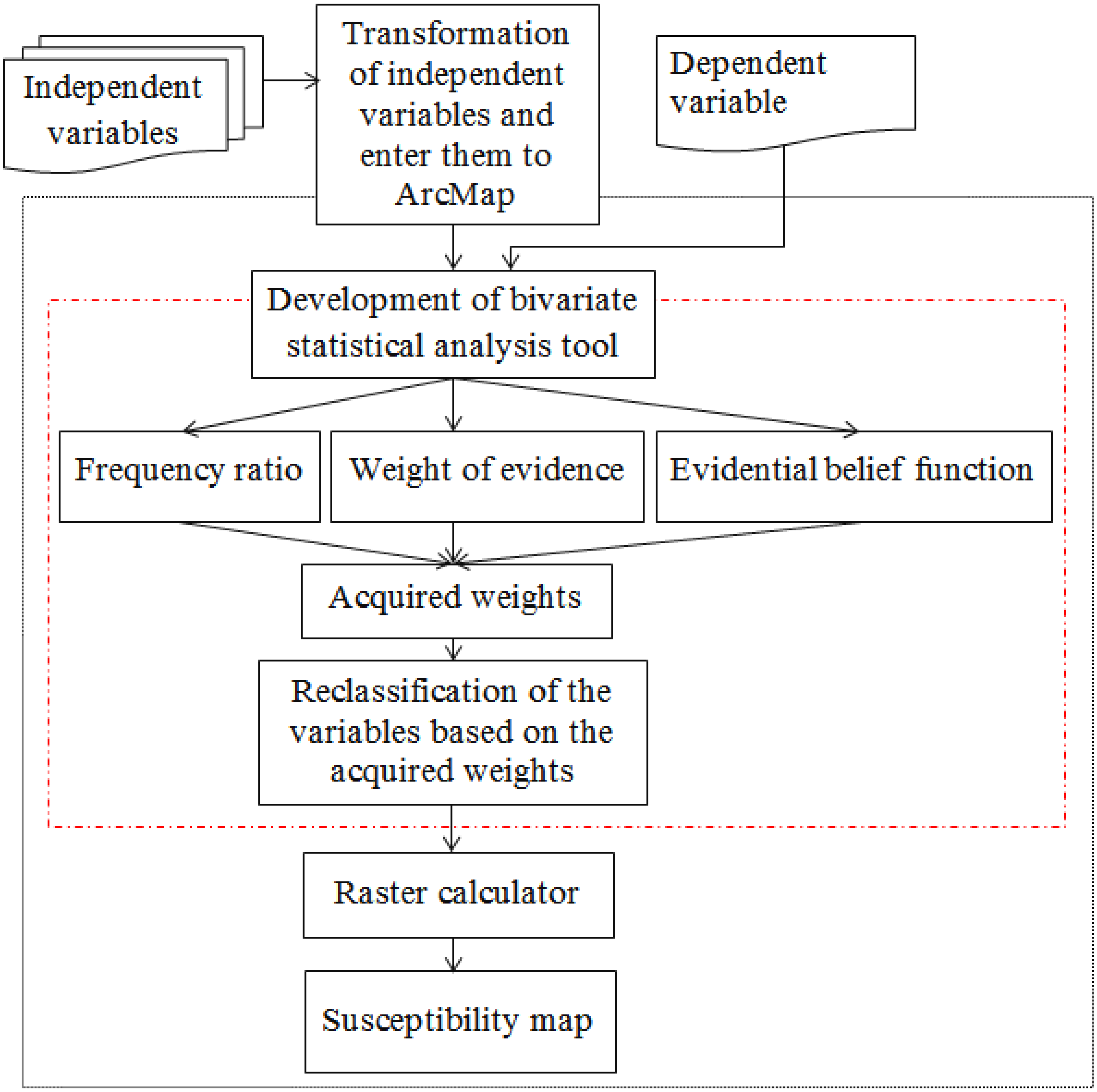

Fig. 1. General design of the methodology and BSA tool.

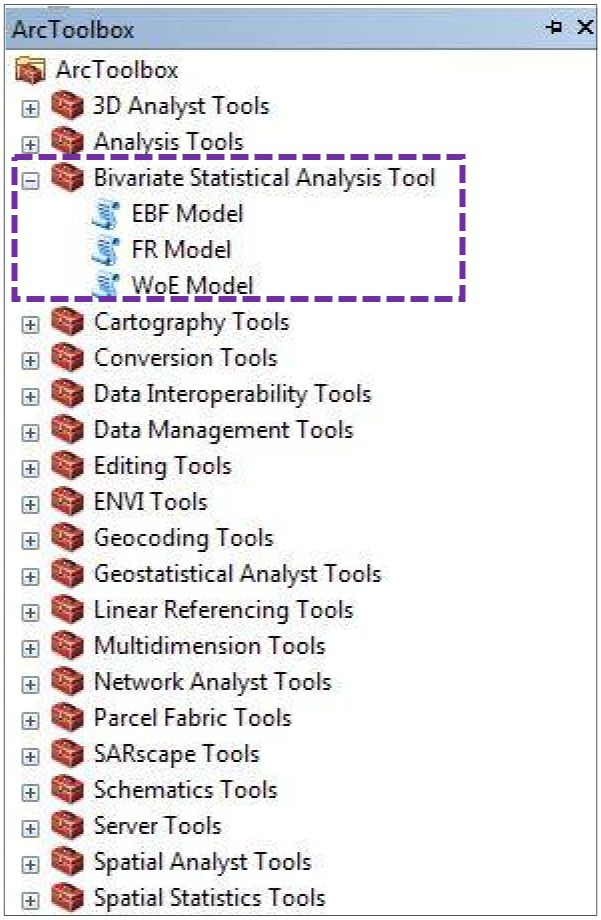

Fig. 2. BSA tool interface.

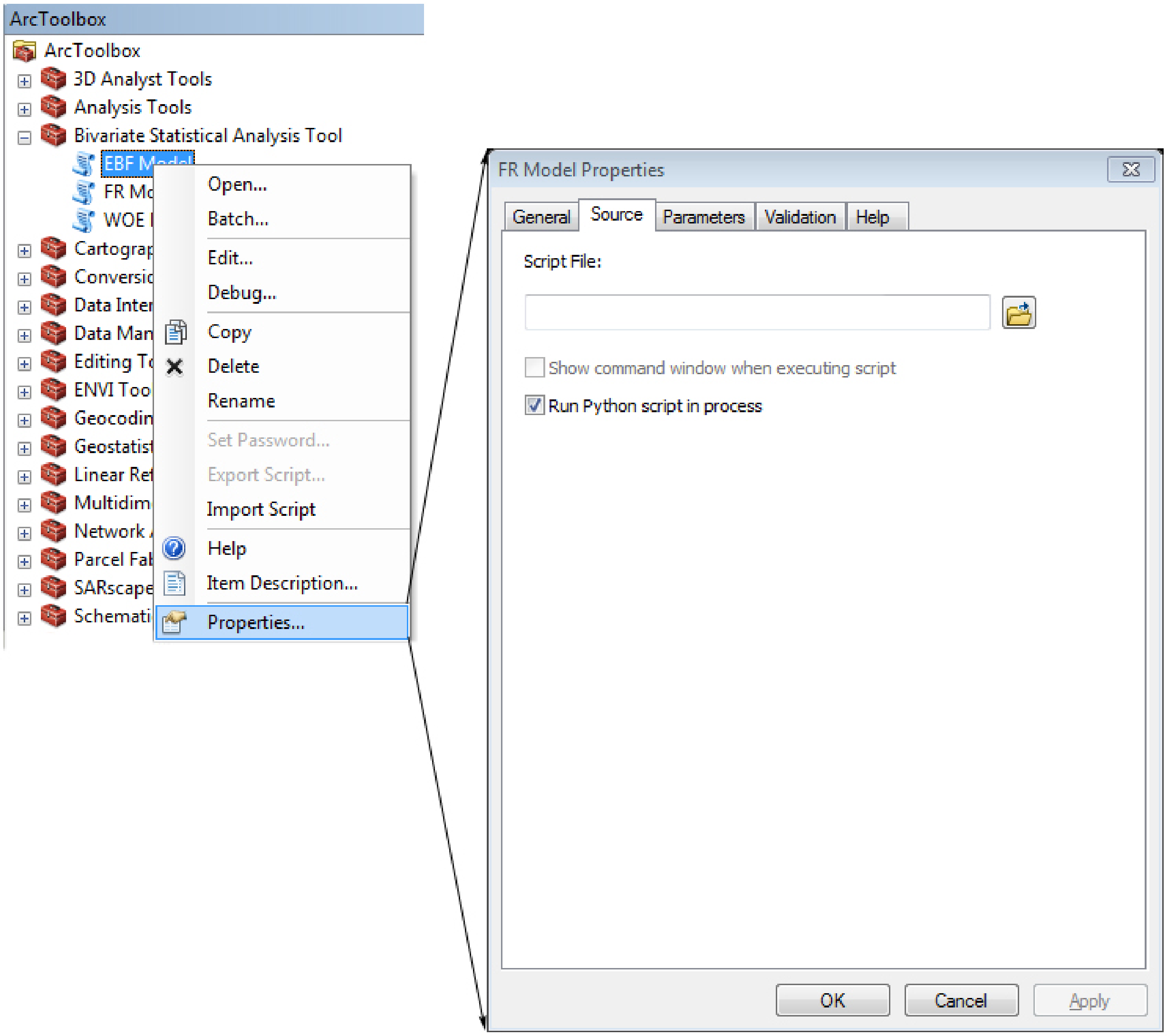

Fig. 3. Procedure to add the BSM tool in ArcGIS.

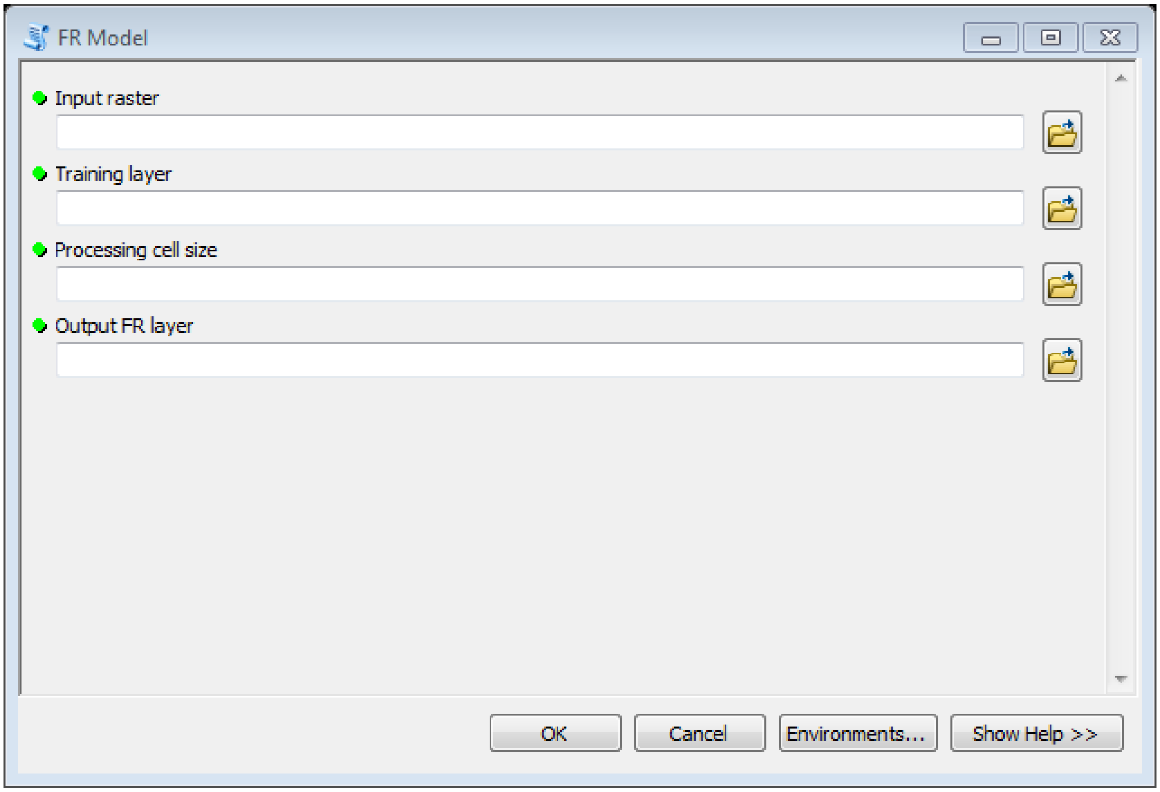

Fig. 4. Graphic user interface of the FR tool.

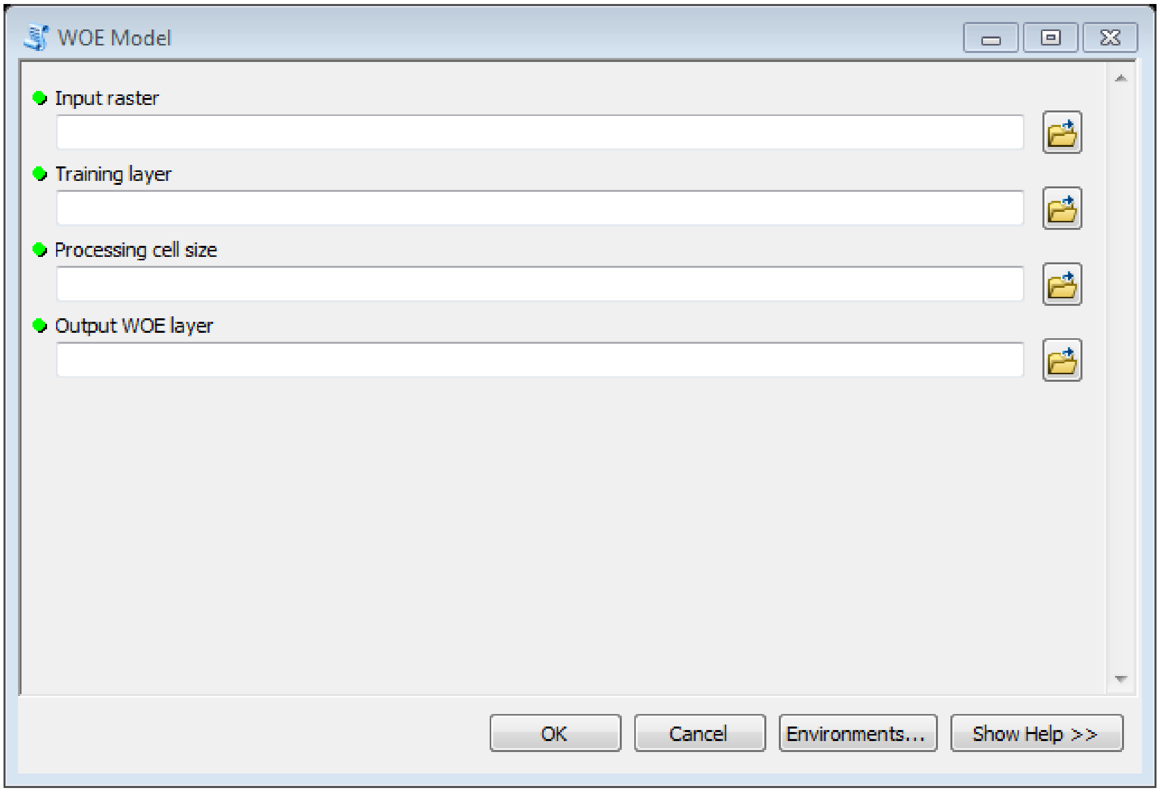

Fig. 5. Graphic interface of the WoE tool.

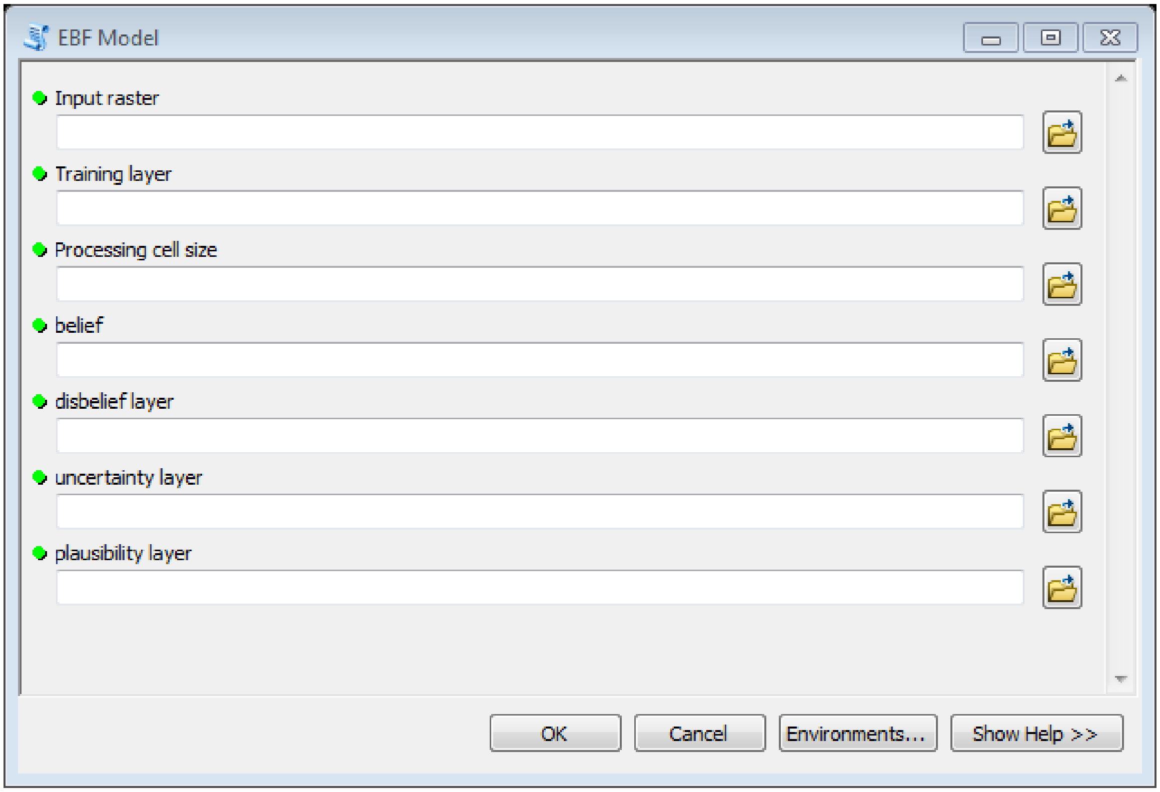

Fig. 6. Graphic interface of the EBF tool.

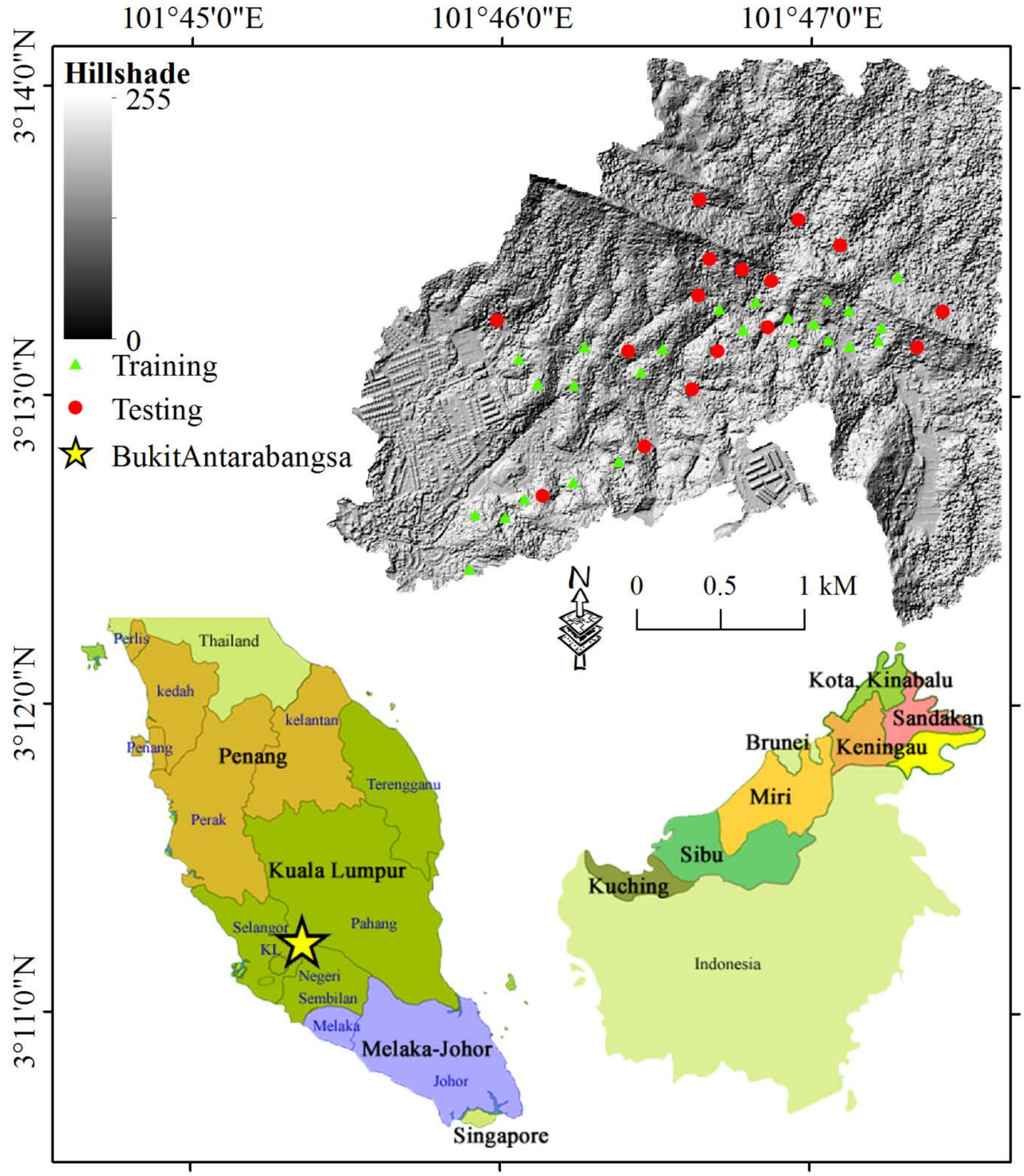

Fig. 7. Location of the pilot study area for testing the proposed ArcMAP tool.

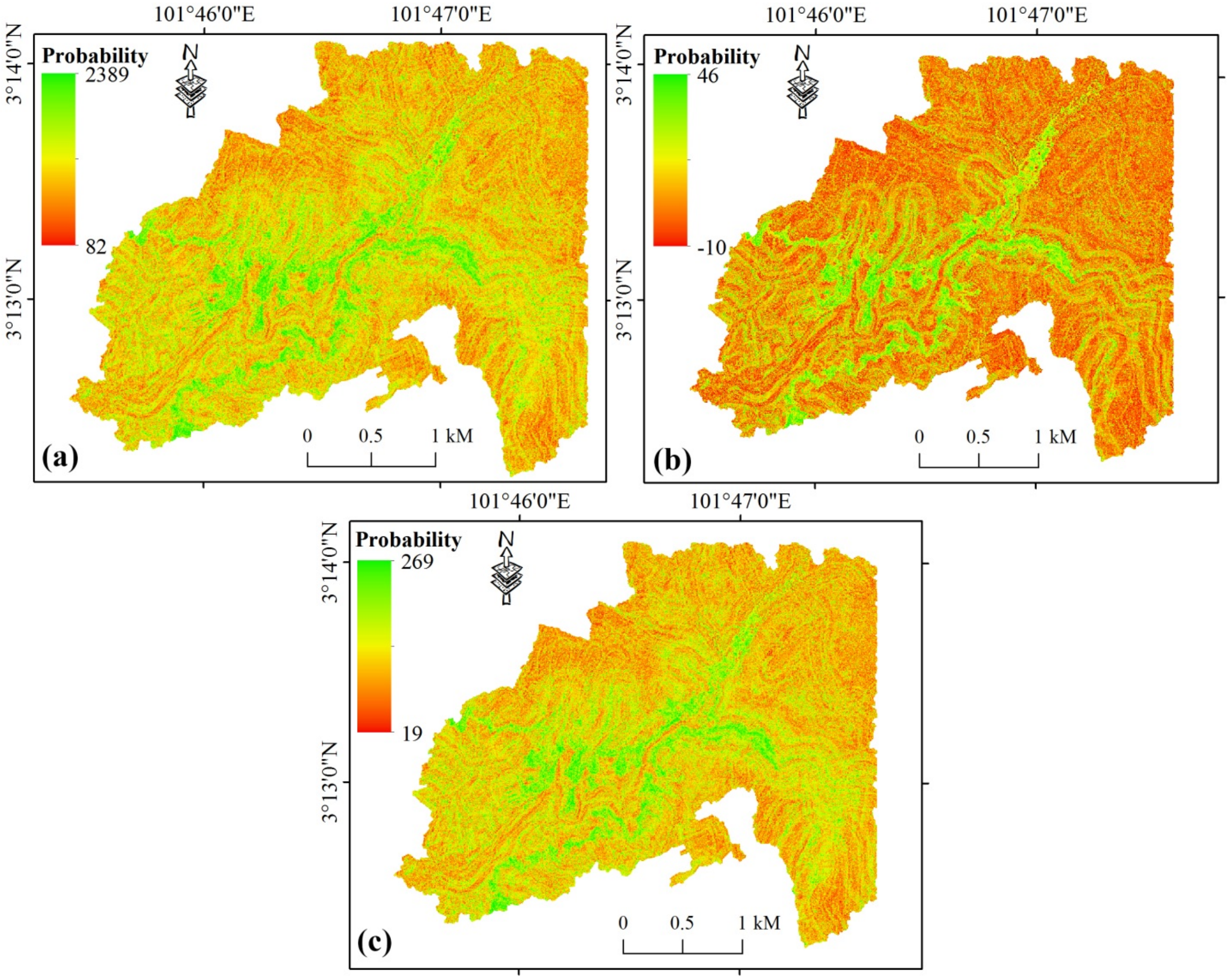

Fig. 8. Landslide susceptibility maps derived from a) FR, b) WoE, and c) EBF.

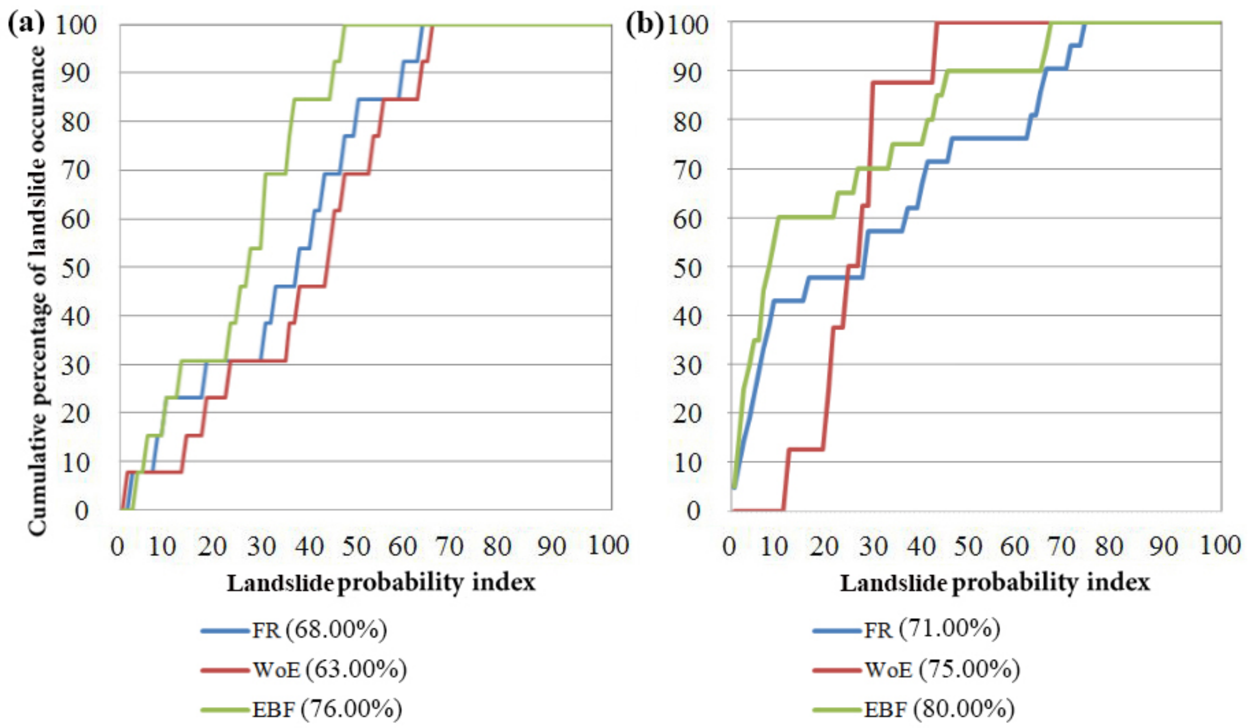

Fig. 9. Graphic representation of the cumulative frequency diagram presenting the cumulative landslide occurrence (%; y-axis) in landslide probability index rank (%; x-axis): a) success rate, and b) prediction rate.

Appendix 1:~~, ~~(the code of the three models coded in python)~~.~~

FR

```
import arcpy
arcpy.CheckOutExtension("spatial")
arcpy.env.overwriteOutput = True
Input_raster = arcpy.GetParameterAsText(0)
Training_layer = arcpy.GetParameterAsText(1)
Processing_cell_size = arcpy.GetParameterAsText(2)
if Processing_cell_size == '#' or not Processing_cell_size:
    Processing_cell_size = "5" # provide a default value if unspecified
Output_FR_layer = arcpy.GetParameterAsText(3)
Folder_Location = "c:\\"
arcpy.CreateFolder_management(Folder_Location, "FR_modeler")
arcpy.env.workspace = "C:\FR_modeler"
arcpy.gp.Lookup_sa(Input_raster, "VALUE", "C:\FR_modeler\lookupp")
arcpy.gp.ZonalGeometryAsTable_sa("lookupp", "VALUE", "c:/fr_modeler/zonal", Processing_cell_size)
arcpy.AddField_management("zonal", "layer_pert", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
arcpy.Statistics_analysis("zonal", "C:\FR_Modeler\ssummary.dbf", "AREA SUM", "")
arcpy.JoinField_management("zonal", "layer_pert", "C:\FR_Modeler\ssummary.dbf", "OID", "SUM_AREA")
arcpy.CalculateField_management("zonal", "layer_pert", "[AREA] / [SUM_AREA]", "VB", "")
arcpy.gp.TabulateArea_sa("lookupp", "VALUE", Training_layer, "VALUE", "c:/fr_modeler/tabulate", Processing_cell_size)
arcpy.AddField_management("tabulate", "layer_l", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
arcpy.Statistics_analysis("tabulate", "C:\FR_modeler\ssssummary.dbf", "VALUE_1 SUM", "")
arcpy.JoinField_management("tabulate", "layer_l", "C:\FR_modeler\ssssummary.dbf", "OID", "SUM_VALUE_")
arcpy.CalculateField_management("tabulate", "layer_l", "[VALUE_1] / [SUM_VALUE_]", "VB", "")
arcpy.JoinField_management("zonal", "VALUE", "tabulate", "VALUE", "")
arcpy.AddField_management("zonal", "fr_layer", "LONG", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
arcpy.CalculateField_management("zonal", "fr_layer", "Int ([layer_l]/ [layer_pert]*100)", "VB", "")
arcpy.CopyRows_management("zonal", "c:/fr_modeler/frresult", "#")
arcpy.gp.ReclassByTable_sa("lookupp","frresult","VALUE","VALUE","FR_LAYER",Output_FR_layer)
```

```
arcpy.Delete_management("C:\FR_modeler\ssssummary.dbf","#")
arcpy.Delete_management("C:\FR_modeler\ssummary.dbf","#")
arcpy.Delete_management("C:\FR_modeler\zonal","#")
arcpy.Delete_management("C:\FR_modeler\frresult","ArcInfoTable")
arcpy.Delete_management("C:/FR_modeler/lookupp","RasterDataset")
arcpy.Delete_management("C:/FR_modeler/tabulate","ArcInfoTable")
arcpy.Delete_management("frresult","#")
```

WoE

```
import arcpy
arcpy.CheckOutExtension("spatial")
arcpy.env.overwriteOutput = True
Input_raster = arcpy.GetParameterAsText(0)
Training_layer = arcpy.GetParameterAsText(1)
Processing_cell_size = arcpy.GetParameterAsText(2)
if Processing_cell_size == '#' or not Processing_cell_size:
    Processing_cell_size = "5" # provide a default value if unspecified
Output_WOE_layer = arcpy.GetParameterAsText(3)
Folder_Location = "c:\\"
arcpy.CreateFolder_management(Folder_Location, "woe_modeler")
arcpy.env.workspace = "C:\woe_modeler"
arcpy.gp.Lookup_sa(Input_raster, "VALUE", "C:\woe_modeler\lookupp")
arcpy.gp.ZonalGeometryAsTable_sa("C:\woe_modeler\lookupp","VALUE","c:/woe_modeler/zonal","5")
arcpy.AddField_management("zonal", "per_cell", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
arcpy.Statistics_analysis("zonal", "C:\woe_modeler\ssummary.dbf", "AREA SUM", "")
arcpy.JoinField_management("zonal", "per_cell", "C:\woe_modeler\ssummary.dbf", "OID", "SUM_AREA")
arcpy.CalculateField_management("zonal", "per_cell", "[AREA] / [SUM_AREA]", "VB", "")
arcpy.gp.TabulateArea_sa("lookupp", "VALUE", Training_layer, "VALUE", "c:/woe_modeler/tabulate", Processing_cell_size)
arcpy.AddField_management("tabulate", "per_depost", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
arcpy.Statistics_analysis("tabulate", "C:\woe_modeler\ssssummary.dbf", "VALUE_1 SUM", "")
```

```
arcpy.JoinField_management("tabulate", "per_depost", "C:\woe_modeler\ssssummary.dbf", "OID", "SUM_VALUE_")
arcpy.AddField_management("tabulate", "deposit", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
arcpy.CalculateField_management("tabulate", "deposit", "[VALUE_1] + 0.0000000001", "VB", "")
arcpy.CalculateField_management("tabulate", "per_depost", "( [deposit] / [SUM_VALUE_] )", "VB", "")
arcpy.JoinField_management("zonal", "VALUE", "tabulate", "VALUE", "")
arcpy.AddField_management("zonal", "fr_layer", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
arcpy.CalculateField_management("zonal", "fr_layer", "[per_depost] / [per_cell]", "VB", "")
arcpy.AddField_management("zonal", "per_non_cl", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
arcpy.CalculateField_management("zonal", "per_non_cl", "1 - [per_cell]", "VB", "")
arcpy.AddField_management("zonal", "pr_non_dep", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
arcpy.CalculateField_management("zonal", "pr_non_dep", "1 - [per_depost]", "VB", "")
arcpy.AddField_management("zonal", "non_fr", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
arcpy.CalculateField_management("zonal", "non_fr", "[pr_non_dep] / [per_non_cl]", "VB", "")
arcpy.AddField_management("zonal", "w_positave", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
arcpy.CalculateField_management("zonal", "w_positave", "Log ( [FR_LAYER] +1 )", "VB", "")
arcpy.AddField_management("zonal", "w_nagative", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
arcpy.CalculateField_management("zonal", "w_nagative", "Log ( [non_fr] + 1 )", "VB", "")
arcpy.AddField_management("zonal", "C", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
arcpy.CalculateField_management("zonal", "C", "[w_positave] - [w_nagative]", "VB", "")
arcpy.AddField_management("zonal", "S2_W_pos", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
arcpy.CalculateField_management("zonal", "S2_W_pos", "( 1 / [deposit] ) + ( 1 / ( [AREA] - [deposit] ))", "VB", "")
arcpy.AddField_management("zonal", "s2_w_nag", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
arcpy.CalculateField_management("zonal", "s2_w_nag", "( 1 / ( [SUM_VALUE_] - [VALUE_12] ) ) + ( 1 / ( [SUM_AREA] - [SUM_VALUE_] - [AREA] - [VALUE_12] ))", "VB", "")
arcpy.AddField_management("zonal", "s_c", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
arcpy.CalculateField_management("zonal", "s_c", "Sqr ( [S2_W_pos] + [s2_w_nag] )", "VB", "")
arcpy.AddField_management("zonal", "woe", "LONG", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
arcpy.CalculateField_management("zonal", "woe", "Int ( [C] / [s_c] )", "VB", "")
arcpy.CopyRows_management("zonal", "c:/woe_modeler/woeresult", "#")
arcpy.gp.ReclassByTable_sa("lookupp", "woeresult", "VALUE", "VALUE", "WOE", Output_WOE_layer, "DATA")
```

```
arcpy.Delete_management("C:\FR_modeler\ssssummary.dbf","#")
arcpy.Delete_management("C:\FR_modeler\ssummary.dbf","#")
arcpy.Delete_management("C:\FR_modeler\zonal","#")
arcpy.Delete_management("C:\FR_modeler\frresult","ArcInfoTable")
arcpy.Delete_management("C:/FR_modeler/lookupp","RasterDataset")
arcpy.Delete_management("C:/FR_modeler/tabulate","ArcInfoTable")
arcpy.Delete_management("frresult","#")
```

EBF

```
import arcpy
arcpy.CheckOutExtension("spatial")
arcpy.env.overwriteOutput = True
Input_raster = arcpy.GetParameterAsText(0)
Training_layer = arcpy.GetParameterAsText(1)
Processing_cell_size = arcpy.GetParameterAsText(2)
if Processing_cell_size == '#' or not Processing_cell_size:
    Processing_cell_size = "5" # provide a default value if unspecified
belief = arcpy.GetParameterAsText(3)
disbelief_layer = arcpy.GetParameterAsText(4)
uncertainty_layer = arcpy.GetParameterAsText(5)
plausibility_layer= arcpy.GetParameterAsText(6)
Folder_Location = "c:\\"
arcpy.CreateFolder_management(Folder_Location, "EBF_modeler")
arcpy.env.workspace = "C:\EBF_modeler"
arcpy.gp.Lookup_sa(Input_raster, "VALUE", "C:\EBF_modeler\lookupp")
arcpy.gp.ZonalGeometryAsTable_sa("lookupp", "VALUE", "c:/EBF_modeler/zonal", Processing_cell_size)
arcpy.AddField_management("zonal", "layer_pert", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
arcpy.Statistics_analysis("zonal", "C:\EBF_Modeler\ssummary.dbf", "AREA SUM", "")
arcpy.JoinField_management("zonal", "layer_pert", "C:\EBF_Modeler\ssummary.dbf", "OID", "SUM_AREA")
```

```
arcpy.CalculateField_management("zonal", "layer_pert", "[AREA] / [SUM_AREA]", "VB", "")
arcpy.gp.TabulateArea_sa("lookupp", "VALUE", Training_layer, "VALUE", "c:/EBF_modeler/tabulate", Processing_cell_size)
arcpy.AddField_management("tabulate", "layer_l", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
arcpy.Statistics_analysis("tabulate", "C:\EBF_modeler\ssssummary.dbf", "VALUE_1 SUM", "")
arcpy.JoinField_management("tabulate", "layer_l", "C:\EBF_modeler\ssssummary.dbf", "OID", "SUM_VALUE_")
arcpy.CalculateField_management("tabulate", "layer_l", "[VALUE_1] / [SUM_VALUE_]", "VB", "")
arcpy.JoinField_management("zonal", "VALUE", "tabulate", "VALUE", "")
arcpy.AddField_management("zonal", "eq1", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
arcpy.CalculateField_management("zonal", "eq1", "[layer_l] / (( [AREA] - [VALUE_12] ) / ( [SUM_AREA] - [SUM_VALUE_]
))", "VB", "")
arcpy.AddField_management("zonal", "eq2", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
arcpy.CalculateField_management("zonal", "eq2", "( ( [SUM_VALUE_] - [VALUE_12] ) / [SUM_VALUE_] ) / ( (
[SUM_AREA] - [SUM_VALUE_] - [AREA] - [VALUE_12] ) / ( [SUM_AREA] - [SUM_VALUE_] ) )", "VB", "")
arcpy.AddField_management("zonal", "belief", "LONG", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
arcpy.Statistics_analysis("zonal", "C:\EBF_modeler\summarryy.dbf", "eq1 SUM", "belief")
arcpy.JoinField_management("zonal", "belief", "C:\EBF_modeler\summarryy.dbf", "belief", "SUM_eq1")
arcpy.AddField_management("zonal", "disbelief", "LONG", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
arcpy.Statistics_analysis("zonal", "C:\EBF_modeler\summaryy.dbf", "eq2 SUM", "disbelief")
arcpy.JoinField_management("zonal", "disbelief", "C:\EBF_modeler\summaryy.dbf", "disbelief", "SUM_eq2")

arcpy.CalculateField_management("zonal", "belief", "Int ( ( [eq1] / [SUM_eq1] ) * 100 )", "VB", "")
arcpy.CalculateField_management("zonal", "disbelief", "Int ( ( [eq2] / [SUM_eq2] ) * 100 )", "VB", "")

arcpy.AddField_management("zonal", "uncertian", "LONG", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
arcpy.AddField_management("zonal", "plusabili", "LONG", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
arcpy.CalculateField_management("zonal", "uncertian", "100 - [belief] - [disbelief]", "VB", "")
arcpy.CalculateField_management("zonal", "plusabili", "100 - [disbelief]", "VB", "")
arcpy.CopyRows_management("zonal", "c:/EBF_modeler/beliefresult", "#")
arcpy.gp.ReclassByTable_sa("lookupp","beliefresult","VALUE","VALUE","BELIEF",belief,"DATA")
arcpy.CopyRows_management("zonal", "c:/EBF_modeler/disbeliefresult", "#")
arcpy.gp.ReclassByTable_sa("lookupp","disbeliefresult","VALUE","VALUE","DISBELIEF",disbelief_layer,"DATA")
```

```
arcpy.CopyRows_management("zonal", "c:/EBF_modeler/uncertiantyresult", "#")
arcpy.gp.ReclassByTable_sa("lookupp","uncertiantyresult","VALUE","VALUE","uncertian",uncertainty_layer,"DATA")
arcpy.CopyRows_management("zonal", "c:/EBF_modeler/plusabilityresult", "#")
arcpy.gp.ReclassByTable_sa("lookupp","plusabilityresult","VALUE","VALUE","plusabili",plausibility_layer,"DATA")
arcpy.Delete_management("C:/EBF_modeler/beliefresult","ArcInfoTable")
arcpy.Delete_management("C:/EBF_modeler/disbeliefresult","ArcInfoTable")
arcpy.Delete_management("C:/EBF_modeler/plusabilityresult","ArcInfoTable")
arcpy.Delete_management("C:/EBF_modeler/ssssummary.dbf","DbaseTable")
arcpy.Delete_management("C:/EBF_modeler/ssummary.dbf","DbaseTable")
arcpy.Delete_management("C:/EBF_modeler/summarryy.dbf","DbaseTable")
arcpy.Delete_management("C:/EBF_modeler/summaryy.dbf","DbaseTable")
arcpy.Delete_management("C:/EBF_modeler/lookupp","RasterDataset")
arcpy.Delete_management("C:/EBF_modeler/tabulate","ArcInfoTable")
arcpy.Delete_management("C:/EBF_modeler/uncertiantyresult","ArcInfoTable")
arcpy.Delete_management("C:/EBF_modeler/zonal","ArcInfoTable")
```