

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

3
4 **Mustafa Neamah Jebur, Biswajeet Pradhan*, Helmi Zulhaidi Mohd Shafri, Zainuddin Md.**
5 **Yusoff, Mahyat Shafapour Tehrani**

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

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

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

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

11
12 **Abstract**

13
14 Modeling and classification difficulties are fundamental issues in natural hazard assessment. A
15 geographic information system (GIS) is a domain that requires users to use various tools to
16 perform different types of spatial modeling. Bivariate statistical analysis (BSA) assists in hazard
17 modeling. To perform this analysis, several calculations are required and the user has to transfer
18 data from one format to another. Most researchers perform these calculations manually by using
19 Microsoft Excel or other programs. This process is time consuming and carries a degree of
20 uncertainty. The lack of proper tools to implement BSA in a GIS environment prompted this
21 study. In this paper, a user-friendly tool, BSM (Bivariate statistical modeler), for BSA technique
22 is proposed. Three popular BSA techniques such as frequency ratio, weights-of-evidence, and
23 evidential belief function models are applied in the newly proposed ArcMAP tool. This tool is
24 programmed in Python and created by a simple graphical user interface, which facilitates the
25 improvement of model performance. The proposed tool implements BSA automatically, thus
26 allowing numerous variables to be examined. To validate the capability and accuracy of this
27 program, a pilot test area in Malaysia is selected and all three models are tested by using the
28 proposed program. Area under curve is used to measure the success rate and prediction rate.
29 Results demonstrate that the proposed program executes BSA with reasonable accuracy. The
30 proposed BSA tool can be used in numerous applications, such as natural hazard, mineral
31 potential, hydrological, and other engineering and environmental applications.

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

33
34 **1 Introduction**

35
36 Techniques to predict a response variable given a set of characteristics are required in several
37 scientific regularities. Numerous applications have been implemented in various areas of
38 geosciences. Bivariate analysis is one of the simplest methods of statistical analysis, and is

39 popular in numerous fields of study. Mathematicians, statisticians, biologists, and hydrologists
40 use this method to perform their analysis. Different types of bivariate statistical analysis (BSA)
41 have been established, for example, frequency ratio (FR), weights-of-evidence (WoE), and
42 evidential belief function (EBF) (Yalcin, 2008). Although each of these methods requires
43 specific mechanisms for calculation, all of these methods operate by using the same concept.
44 Environmental scientists model various natural conditions by using the BSA statistical method.
45 For instance, Ozdemir (2011) employed this technique for the same purpose. The results of the
46 analysis were plotted in ArcGIS after computation in other programs. Mineral potential mapping
47 is also aided by BSA techniques. Carranza (2004) used WoE modeling to map the mineral
48 potential in the administrative province of Abra in northwestern Philippines. Their achievements
49 indicate the plausibility of WoE in the mineral potential mapping of large areas with a small
50 number of mineral prospects. Researchers have applied WoE in mapping mineral potential
51 (Bonham-Carter et al., 1989) and remains popular in this area of research (Carranza et al., 2008).

52
53 BSA is in demand in hazard studies because its procedure is simple and efficient. This technique
54 is has been used in natural hazard applications by researchers to predict the spatial distribution of
55 events. Extensive literature on different BSA techniques and their proficiency assessment are
56 also available. BSA techniques can be used as a simple geospatial analysis tool to determine the
57 probabilistic correlation among dependent variables (produced by using the inventory map of a
58 hazard incidence) and independent variables (conditioning factors) containing multi-categorized
59 maps (Oh et al., 2011). In BSA, the overlay of conditioning factors and computation of hazard
60 densities, the significance of each factor, or the particular mixture of factors can be investigated
61 individually. Bivariate statistical analysis functions by using a dependent variable and one
62 conditioning factor. Hence, the significance of each factor is investigated separately (Porwal et
63 al., 2006).

64
65 In BSA, each conditioning factor is overlaid with the dependent variable. On the basis of the
66 event density, weights are measured for each class of each factor. By using normalized weights
67 (the correlation between the event density in each class of conditioning factor and the event
68 density of the entire region), each conditioning factor is reclassified and the hazard map is
69 produced. By using the acquired weights, decision rules can be produced on the basis of the
70 knowledge of experts. Conditioning factors can also be combined to generate a map with
71 uniform units, which is then overlaid with the inventory map to provide the density per class.
72 The BSA approach has been used in landslide mapping (Constantin et al., 2011), earthquake
73 studies (Xu et al., 2012b), flood susceptibility mapping (Tehrany et al., 2013), land subsidence
74 (Kim et al., 2006; Lee and Park, 2013), and risk analysis (Hu et al., 2009). Numerous studies
75 have been conducted to exploit the potential application of BSA in the hazard domain.

76
77 This research examined the efficiency of statistical analysis, particularly bivariate analysis, in
78 landslide studies in the Cuyahoga River watershed (Nandi and Shakoor, 2010). In another study,

79 FR and WoE were applied in the Sultan Mountains of southwestern Turkey to map areas that are
80 susceptible to landslides (Ozdemir and Altural, 2013). According to Nandi and Shakoor (2010)
81 and Ozdemir and Altural (2013), the BSA model is simple and its input, computation, and
82 outcome procedures are effortlessly understood. The application of EBF in the area of landside
83 studies has been investigated (Lee et al., 2013). Four functions, namely, degree of belief (Bel),
84 degree of disbelief (Dis), degree of uncertainty (Unc) and degree of plausibility (Pls), are
85 calculated separately to determine EBF.

86
87 Each of these functions produces valuable information. However, each function requires
88 individual computations with specific formulas. Tien Bui et al. (2012) used EBF and fuzzy logic
89 methods in their research and found that the landslide susceptibility map derived from EBF has
90 the highest prediction ability. They also established the efficiency of BSA in landslide mapping.
91 BSA is also popular in hydrological research. Flood susceptibility maps assist in mitigation
92 strategies. Lee et al. (2012) used the statistical method of FR to produce a map of flood-prone
93 regions in Busan, Korea, in GIS. Tehrany et al. (2013) proposed an ensemble method of FR and
94 LR to detect regions with high flood probability in Kelantan, Malaysia. The conditioning factors
95 were reclassified on the basis of the weights acquired from the FR technique. These factors were
96 entered in LR processing to obtain the MSA result. If the calculation time for these statistics can
97 be reduced, the efficiency of the developed ensemble method will be enhanced. Hence,
98 producing a tool that is capable of performing BSA calculations will help reduce the calculation
99 time of ensemble methods.

100
101 The BSA model has been widely used in land subsidence susceptibility mapping. In a study by
102 Lee and Park (2013), the FR model was applied and compared with the machine learning of DT.
103 The BSA is a method that is commonly used in natural hazard investigations. Although this
104 method is not novel, the use of BSA has increased in recent years. RS and GIS have
105 revolutionized the domain of natural hazards (Jebur et al., 2013a; Jebur et al., 2013b). Spatial
106 database consists of different data types that are required to be transferred from one format to
107 another because specific programs accept only specific data formats. Scientists have started to
108 develop new programs in hazard studies because of the vital role of early warning systems in
109 such applications (Osna et al., 2014; Pradhan et al., 2014). GIS is capable to store, analyze and
110 show geographic information. It makes it possible to collect, organize, explore, model and view
111 the spatial data for solving complex problems (Barreca et al., 2013). Different types of spatial
112 data analysis ranges from the simple overlaying of various thematic layers to identify the region
113 to the more complex use of mathematical equations or combined statistical models for the
114 prediction of natural hazards. The importance of GIS in catastrophic evaluation was proven by
115 many studies related to the GIS tools usages in exploration of various types of data (Steiniger
116 and Hunter, 2013).

117 For example the existing hydrological GIS-based tools such as Mike SHE and ArcSWAT
118 revealed considerable power in enhancing the accuracy of soil and water evaluations (Lei et al.,

119 2011). These tools are capable of facilitating the modeling and calibration procedure, and
120 decreasing the stages in implementing the models and increasing the precision of the outcomes
121 (Hörmann et al., 2009). The creation of tools that automatically implement susceptibility
122 mapping was applied by Akgun et al. (2012). Akgun et al. (2012) proposed “MamLand,” a
123 program in MATLAB, to create landslide susceptibility mapping by using fuzzy inference
124 system. ArcGIS allows users to produce specific tools for spatial analysis (Stevens et al., 2007).
125 For instance, Pradhan et al. (2014) developed a tool in ArcGIS to apply texture analysis for high-
126 resolution radar data. Recently, a GIS-based system has been developed by Barreca et al. (2013)
127 to evaluate and process the hazard associated to active faults influencing the eastern and southern
128 flanks of Mt. Etna. The proposed tool was created in ArcGIS which contains various thematic
129 datasets. It includes spatially-referenced arc-features and associated database. In another paper,
130 Lei et al. (2011), integrated hydrological code EasyDHM and proposed open source
131 MapWindow GIS tool called MWEasyDHM. Their aim was to create the tool by combining
132 modules for preprocessing, modeling, viewing and analysis. MWEasyDHM tool is user friendly,
133 free and proficient which produces selectable multi-functional hydrological analysis. Similarly, a
134 number of GIS tools are programmed by Etherington (2011) in Python environment for
135 landscape genetics researches. Tools are capable to transform files, view genetic relatedness, and
136 calculate landscape associations through least-cost path procedure. The tools are free and
137 available in ArcToolbox. In a separate paper, Roberts et al. (2010) implemented the research to
138 facilitate the advanced analytic methods. A Marine Geospatial Ecology Tools (MGET) was
139 created in GIS environment which is free, easy to use and efficient tools for the ecologists. The
140 tools were made by integrating different strong programming methods of Python, R, MATLAB,
141 and C++.

142
143 The current research aims to reduce the processing time of BSA by introducing an easy-to-use
144 ArcMap tool. On the basis of the aforementioned problem statement regarding the required
145 processing time and difficulties for BSA, a program that is capable of calculating BSA
146 automatically should be developed. Hence, a tool programmed in Python and based on the BSA
147 technique is proposed. This tool automatically extracts the correlation among each class of
148 conditioning factor and event occurrence, reclassifies the factors on the basis of the acquired
149 weights in a GIS environment, and saves each correlation in separate folders. A simple graphical
150 user interface (GUI) improves the model operation because Python knowledge is not required.
151 The entire process can be performed in ArcGIS without any requirement for another program.
152 The proposed tool was tested to generate a landslide susceptibility map of Bukit Antarabangsa,
153 Ulu Klang, Malaysia.

154 155 **2 Methodology**

156 The procedural and theoretical perspectives of BSA applied in this research include several steps
157 (Fig. 1). In the methodology flowchart, the BSA tool was developed and integrated into ArcGIS.

158 To apply BSA, the conditioning factors should be provided in raster format and classified with
159 the proper scheme by the user. The BSA recognizes the effects of each class of conditioning
160 factor on event occurrence. Hence, this step cannot be eliminated in the BSA process. As a
161 second stage, a dependent variable (training layer) should be constructed by using the inventory
162 map and other resources. This layer should contain a pixel value of one to represent the existence
163 of an event. Once the conditioning factors are classified and the training layers are prepared, FR,
164 WoE, and EBF can be applied automatically. The developed program reclassifies each
165 conditioning factor by using the attained weights and saves them in a separate folder. The group
166 of conditioning factors that have been assessed by BSA are ready to be entered in the raster
167 calculator to derive the corresponding hazard map. The following sub-section represents the
168 overall information on the scheme and functionality of the developed tool.

169 **Fig. 1.** About here

170

171 **2.1 Overall information on scheme and functionality**

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

174 **Fig. 2.** About here

175 The ArcToolbox provided in this research is used to enter the proposed tool in ArcMap. The user
176 defines the source of the Python files of each model from the properties menu of the script (Fig.
177 3).

178 **Fig. 3.** About here

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

181 **2.1.1 FR**

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

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

188 where $N_{pix}(SX_i)$ is the number of pixels, which contain an event in class i of the independent
 189 variable; X , $N_{pix}(X_j)$ is the number of pixels and exist in independent variables X_j ; m is the
 190 number of categoris of the independent variable X_i . Furthermore, n is the total number of
 191 independent variables in the whole area (Yilmaz, 2009). Most of the researchers performed these
 192 calculations manually by using Microsoft Excel or other programs. Once the weights were
 193 obtained, these values were used to reclassify the independent variables by using the spatial
 194 analyst tool in ArcGIS. The raster calculator in ArcGIS was used to obtain the final susceptibility
 195 map. The proposed tool in ArcMap can apply the FR automatically and reclassify the
 196 independent variables on the basis of the gained weights.
 197

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

207 **Fig. 4.** About here

208 2.1.2 WoE

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

$$W_i^+ = \ln \frac{P\{B|A\}}{P\{B|\bar{A}\}} \quad (2)$$

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

216 where P represents the probability, \ln is the natural *log*. B , and \bar{B} reveals the existence and
 217 nonexistence of the independent variable. A and \bar{A} show the existence and nonexistence of the
 218 event. A positive weight (W^+) determines the presence of the specific independent variable at
 219 the event, and the amount of positive weight represents the positive correlation between the

220 presence of the independent variable and event, respectively. A negative weight (W^-) indicates
221 the nonexistence of the independent variable and shows the amount of negative correlation.

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

223

$$C(C = W^+ - W^-) \quad (6)$$

224

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

228

229 The standard deviation of W is calculated as follows:

230

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

231

232 where $S(W^+)$ and $S(W^-)$ are the variance of the positive and negative weights, respectively.

233 These variances can be calculated by using the following equations:

234

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

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

235

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

241 **Fig. 5.** About here

242

243 2.1.3 EBF

244 Dempster (1967) is an innovator who presented the Dempster–Shafer theory of evidence, which
245 is known as a generalized Bayesian theory of subjective probability. This theory has been used in
246 several fields of study, including environmental and hazard studies (Awasthi and Chauhan,
247 2011). This theory also has relative flexibility, which is considered its advantage, accepts
248 uncertainty, and is capable of combining beliefs from different sources of evidence. EBF
249 estimates the probability that a hypothesis is true and evaluates how close the evidence is to the
250 truth of that hypothesis. A complex procedure is required to calculate EBF compared with FR.

251 To compute the EBF, four functions (Bel, Dis, Unc, and Pls) should be measured separately (Lee
252 et al., 2013). Individual computation by using specific formulas is required to provide this
253 information.

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

257

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

258

259 where C is the frame of discernment and $P(C)$ is the set of all subsets of C , counting the empty
260 set (Φ) and C itself. Mass function is another name for the mentioned function that satisfies
261 $m(\Phi) = 0$ and $\sum_{A \subset C} m(A) = 1$, where A is any subset of C . The degree in which the evidence
262 support A is calculated by $m(A)$, which is represented by a belief function ($\text{Bel}(A)$). Suppose that
263 $N(L)$ and $N(C)$ are the total number of pixels affected by the event and the total number of pixels
264 in the study area, respectively; C_{ij} is the j -th class of the independent variable of C_i ($i =$
265 $1, 2, 3, \dots, n$); $N(C_{ij})$ is the total number of pixels in class C_{ij} ; and $N = (L \cap C_{ij})$ is the number of
266 pixels affected by the event in C_{ij} . Therefore, the data-driven measurement of EBF can be
267 calculated by using the following equation (Tien Bui et al., 2012):

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

275 **Fig. 6.** About here

276

277 **2.2 Code description**

278 The code was designed in python 27 (The default software included with windows 7). In the
279 beginning, the arcpy library is called to check the code for spatial extension in order to continue
280 the process. After that, when the user defines the raster, the code calls the raster data as test using
281 the command "GetParameterAsText" which is part of arcpy library. Using same as the previous
282 command, the code will define the output layer for the chosen model. The default path for all the
283 sub process is defined to be in "C" drive because it is the default drive in all the systems.

284 Therefore, the code creates folder calls “FR_modeler”, "WOE_modeler", or "EBF_modeler"
285 depending on the selected process.

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

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

301 **2.3 Test area and data**

302 Although the developed program can be used in any application that employs BSA, the
303 proficiency of the tool was tested in the hazard domain. To examine the capability and efficiency
304 of the developed program, landslide susceptibility analyses were performed by using the
305 developed ArcMAP tool with three BSA models, namely, FR, WoE, and EBF. The program was
306 tested for the landslide susceptibility mapping of Bukit Antarabangsa, Ulu Klang, Malaysia (Fig.
307 7).

308

309 **Fig. 7.** About here

310

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

322

323 **3 Experimental results and discussion**

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

334

335 **Fig. 8.** About here

336

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

348

349 **Fig. 9.** About here

350

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

357

358 **4 Conclusion**

359 To perform hazard studies, several requirements, such as constructing the precise spatial
360 database, obtaining high-resolution imagery, and providing a reliable inventory map, should be

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

382
383

384 **References**

- 385 Akgun, A., Sezer, E. A., Nefeslioglu, H. A., Gokceoglu, C., Pradhan, B.: An easy-to-use MATLAB
386 program (MamLand) for the assessment of landslide susceptibility using a Mamdani fuzzy
387 algorithm, *Comput. Geosci.*, 38, 23-34, 2012.
- 388 Awasthi, A., Chauhan, S. S.: Using AHP and Dempster–Shafer theory for evaluating sustainable transport
389 solutions, *Environ. Modell. Softw.*, 26, 787-796, 2011.
- 390 Barreca, G., Bonforte, A., Neri, M.: A pilot GIS database of active faults of Mt. Etna (Sicily): A tool for
391 integrated hazard evaluation, *J. Volcanol. Geoth. Res.*, 251, 170-186, 2013.
- 392 Beynon, M., Curry, B., Morgan, P.: The Dempster–Shafer theory of evidence: an alternative approach to
393 multicriteria decision modelling, *Omega.*, 28, 37-50, 2000.
- 394 Bonham-Carter, G. F., Agterberg, F. P., Wright, D. F.: Weights of evidence modelling: a new approach to
395 mapping mineral potential, *Comput. Geol.*, 89, 171–183, 1989.
- 396 Carranza, E. J. M.: Weights of evidence modeling of mineral potential: a case study using small number
397 of prospects, Abra, Philippines, *Nat. Resour. Res.*, 13, 173-187, 2004.
- 398 Carranza, E. J. M., Van Ruitenbeek, F., Hecker, C., van der Meijde, M., van der Meer, F. D.: Knowledge-
399 guided data-driven evidential belief modeling of mineral prospectivity in Cabo de Gata, SE
400 Spain, *Int. J. Appl. Earth. Obs.*, 10, 374-387, 2008.
- 401 Constantin, M., Bednarik, M., Jurchescu, M. C., Vlaicu, M.: Landslide susceptibility assessment using the
402 bivariate statistical analysis and the index of entropy in the Sibiciu Basin (Romania), *Environ.*
403 *Earth. Sci.*, 63, 397-406, 2011.

404 Dempster, A. P.: Upper and lower probabilities induced by a multivalued mapping, *Ann. Math. Stat.*, 325-
405 339, 1967.

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

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

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

412 Jebur, M. N., Pradhan, B., Tehrany, M. S.: Detection of vertical slope movement in highly vegetated
413 tropical area of Gunung pass landslide, Malaysia, using L-band InSAR technique, *Geosci. J.*, 18,
414 61-68, 2013a.

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

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

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

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

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

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

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

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

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

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

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

441 Ozdemir, A.: Using a binary logistic regression method and GIS for evaluating and mapping the
442 groundwater spring potential in the Sultan Mountains (Aksehir, Turkey), *J. Hydrol.*, 405, 123-
443 136, 2011.

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

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

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

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

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

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

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

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

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

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

471 Xu, C., Xu, X., Lee, Y. H., Tan, X., Yu, G., Dai, F.: The 2010 Yushu earthquake triggered landslide
472 hazard mapping using GIS and weight of evidence modeling, Environ. Earth. Sci., 66, 1603-1616,
473 2012b.

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

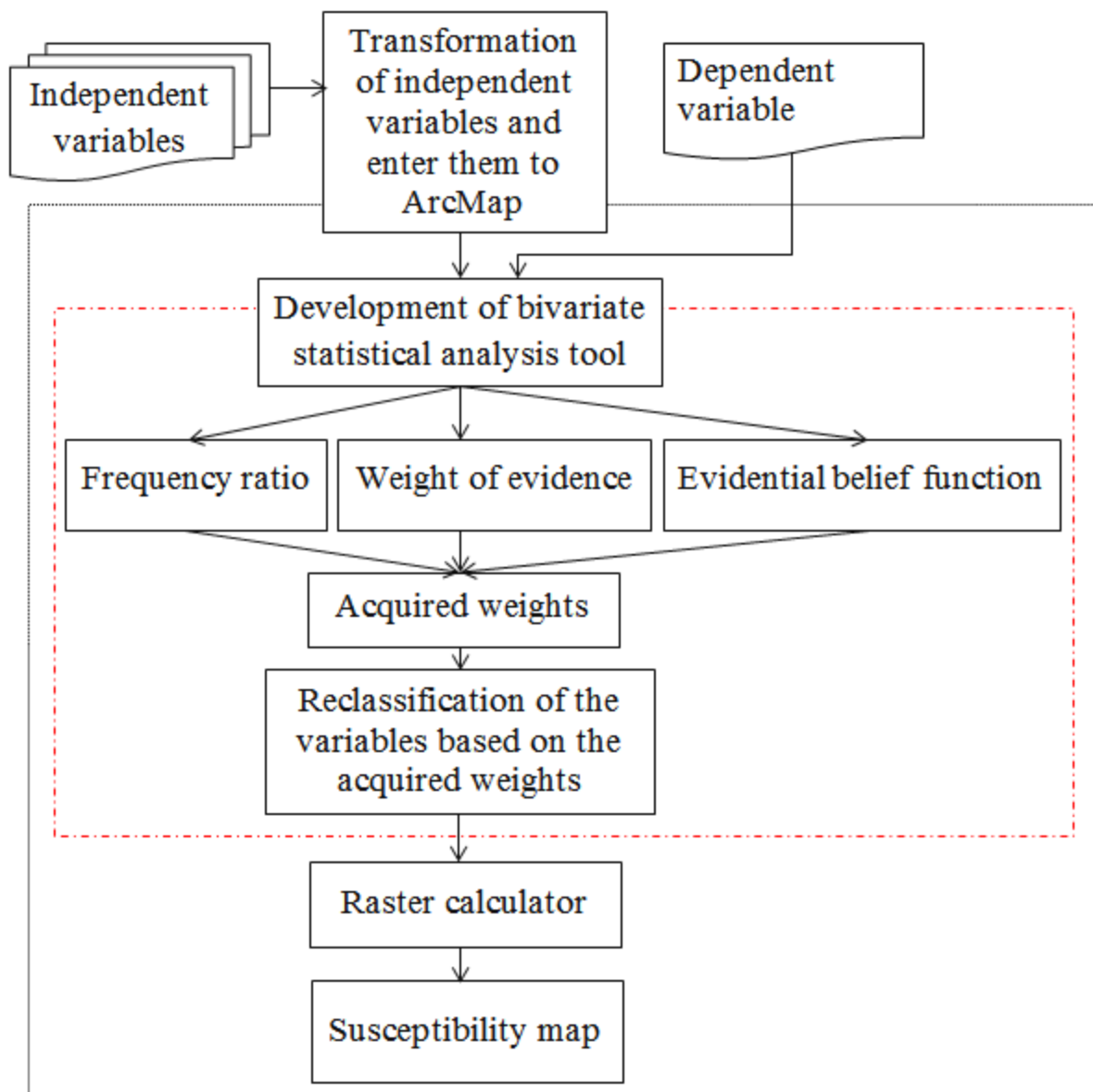
476 Yalcin, A., Reis, S., Aydinoglu, A., Yomralioglu, T.: A GIS-based comparative study of frequency ratio,
477 analytical hierarchy process, bivariate statistics and logistics regression methods for landslide
478 susceptibility mapping in Trabzon, NE Turkey, Catena., 85, 274-287, 2011.

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

482

483

484

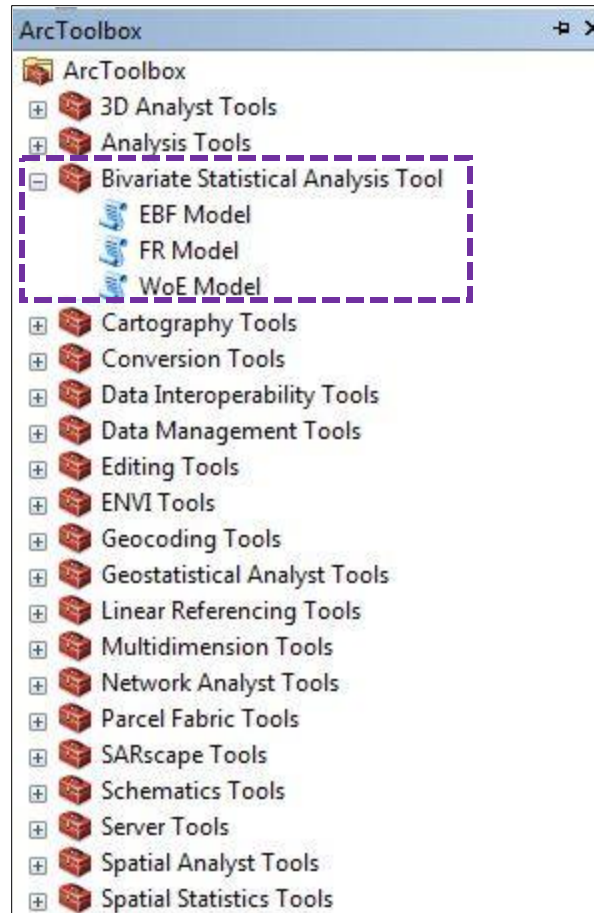


485

486

487

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

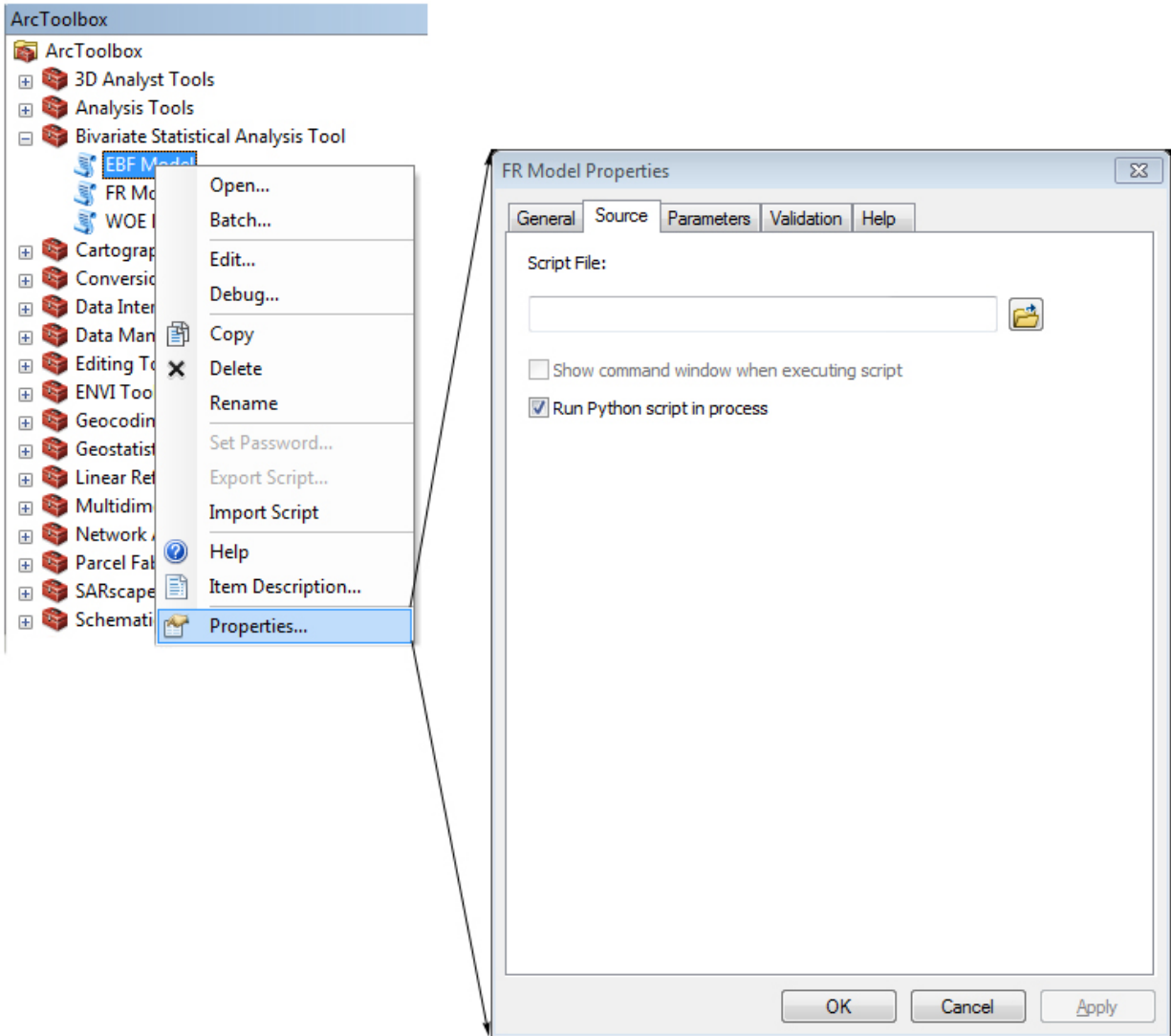


488

489

490

Fig. 2. BSA tool interface.

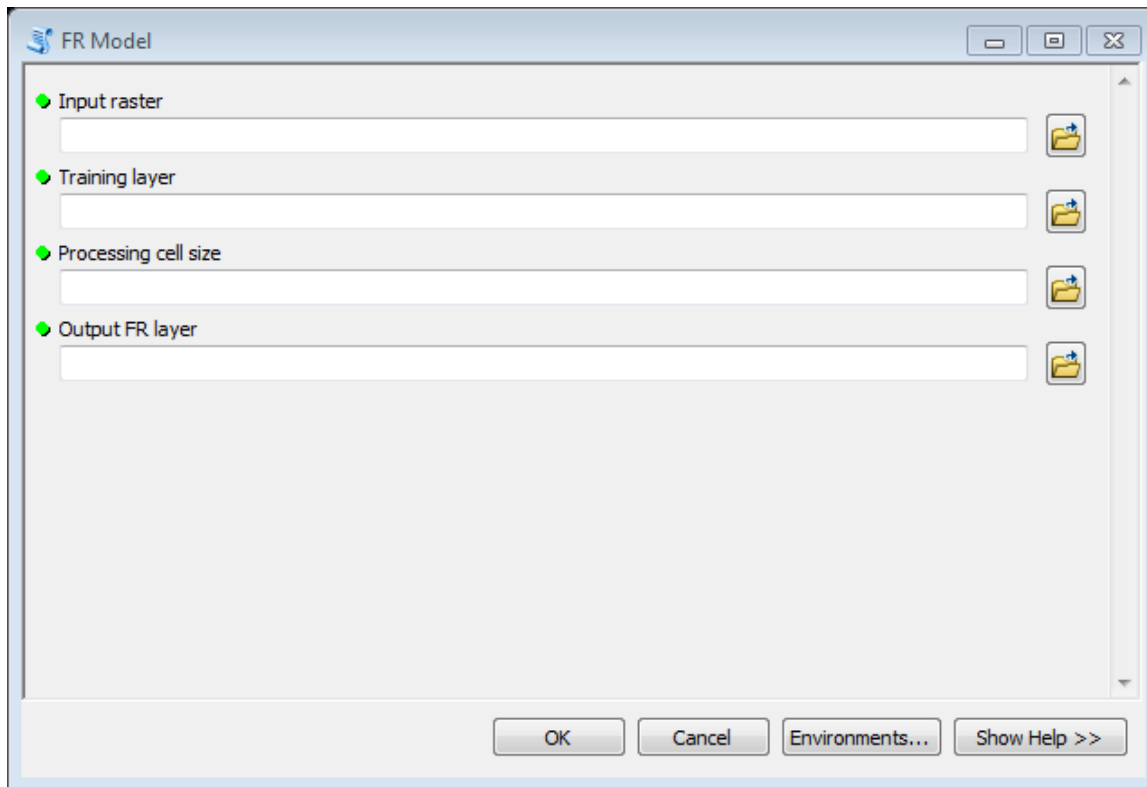


491

492

493

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

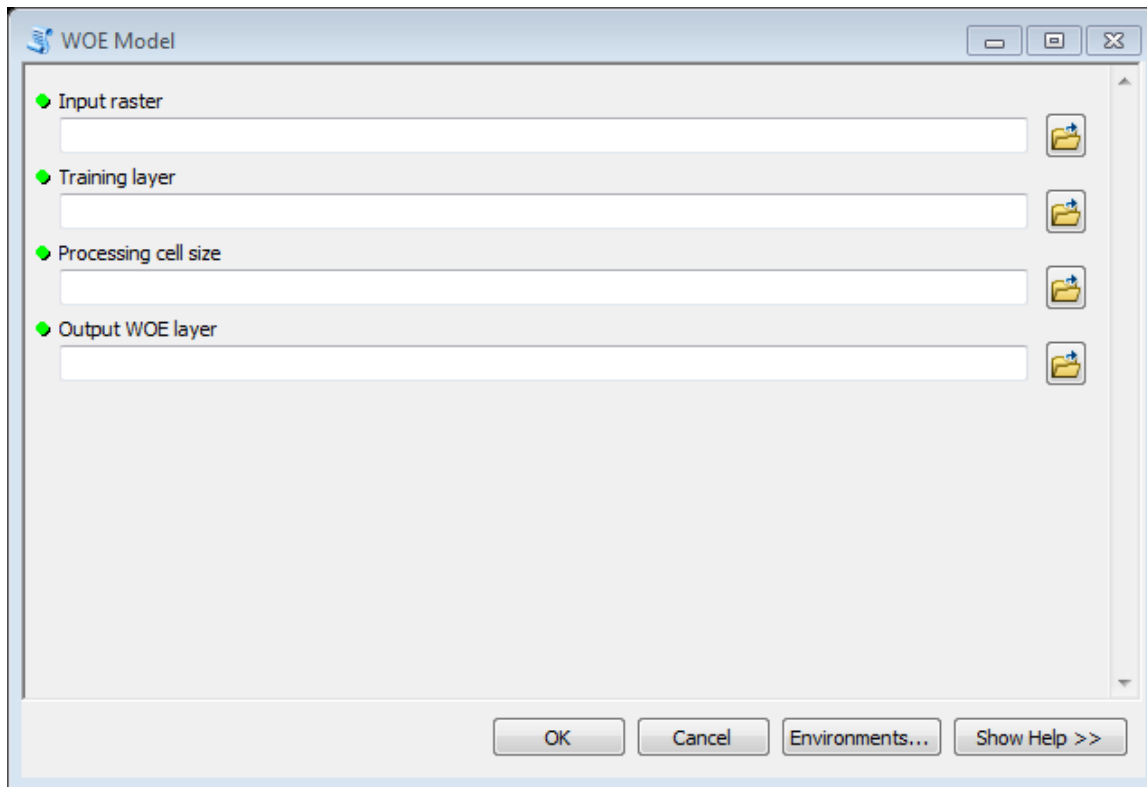


494

495

496

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

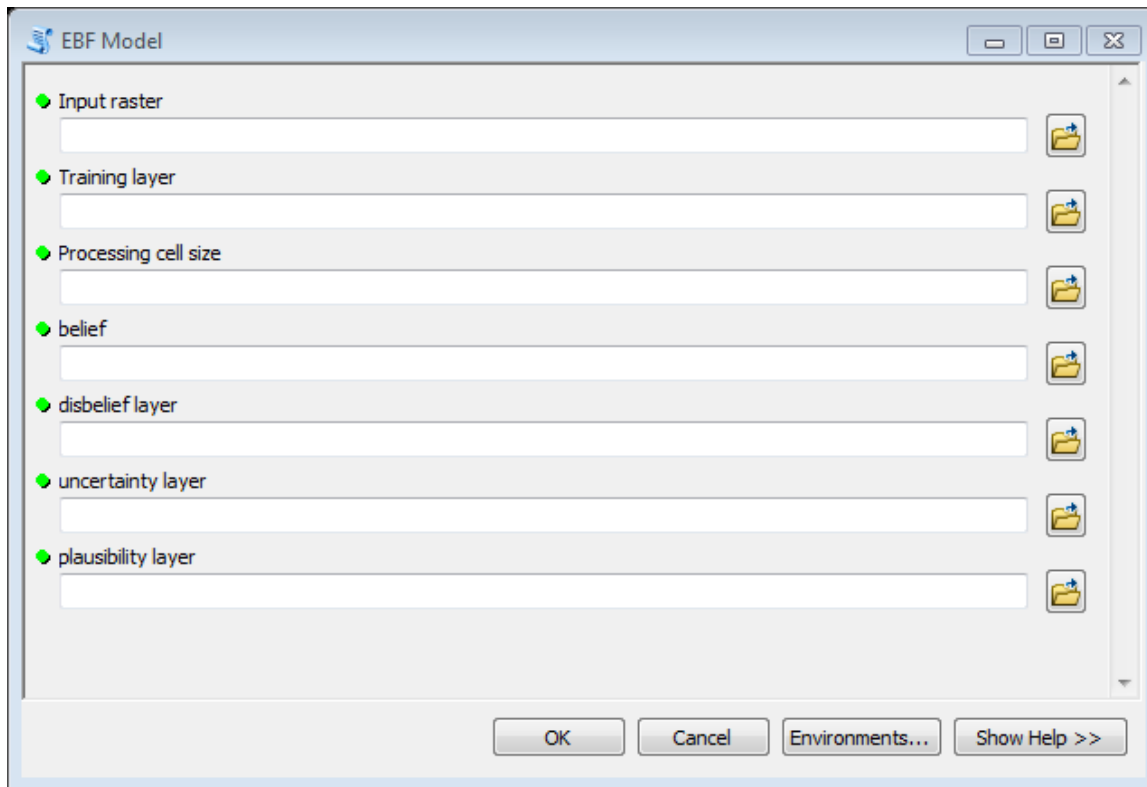


497

498

499

Fig. 5. Graphic interface of the WoE tool.



500

501

502

Fig. 6. Graphic interface of the EBF tool.

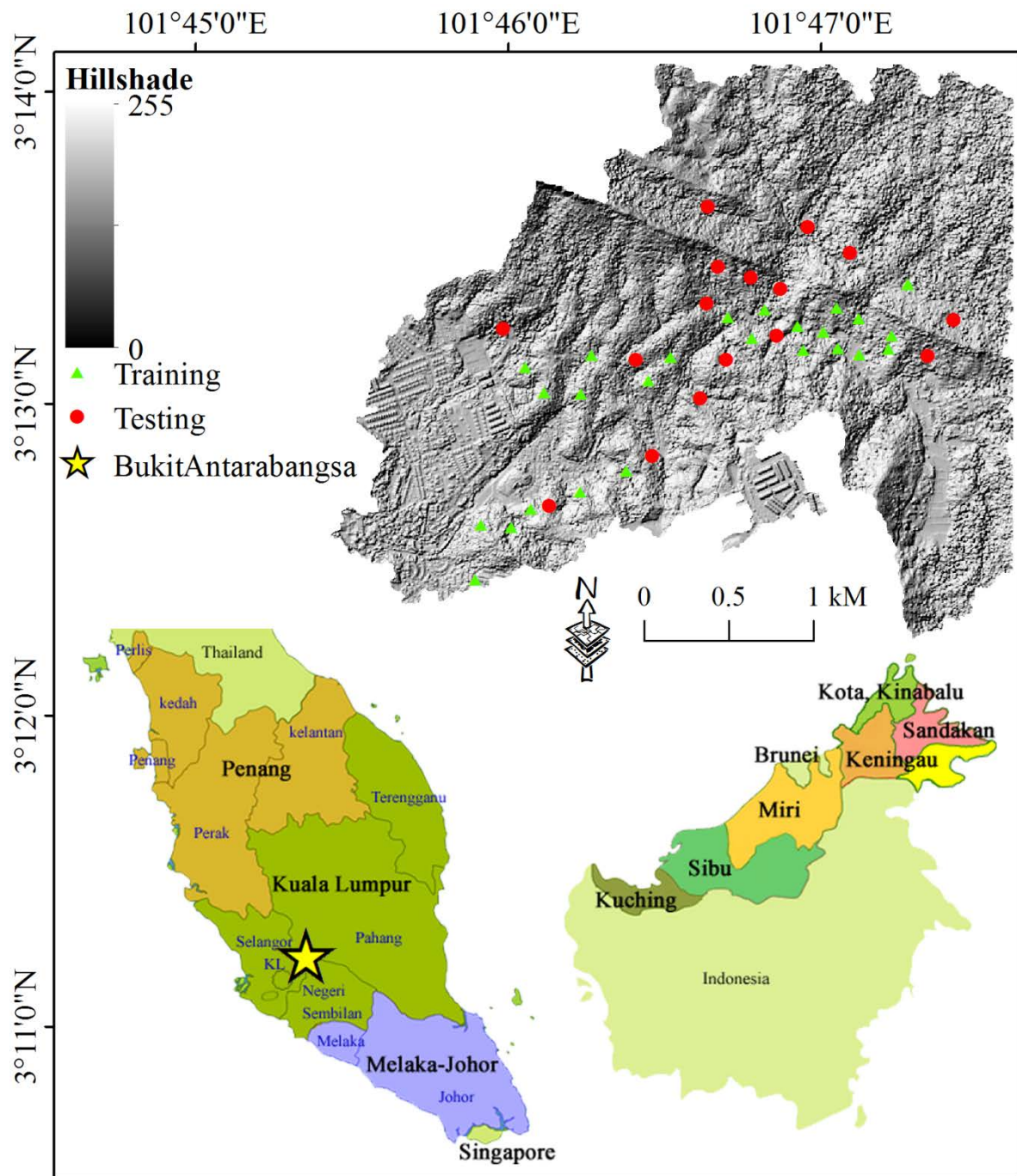
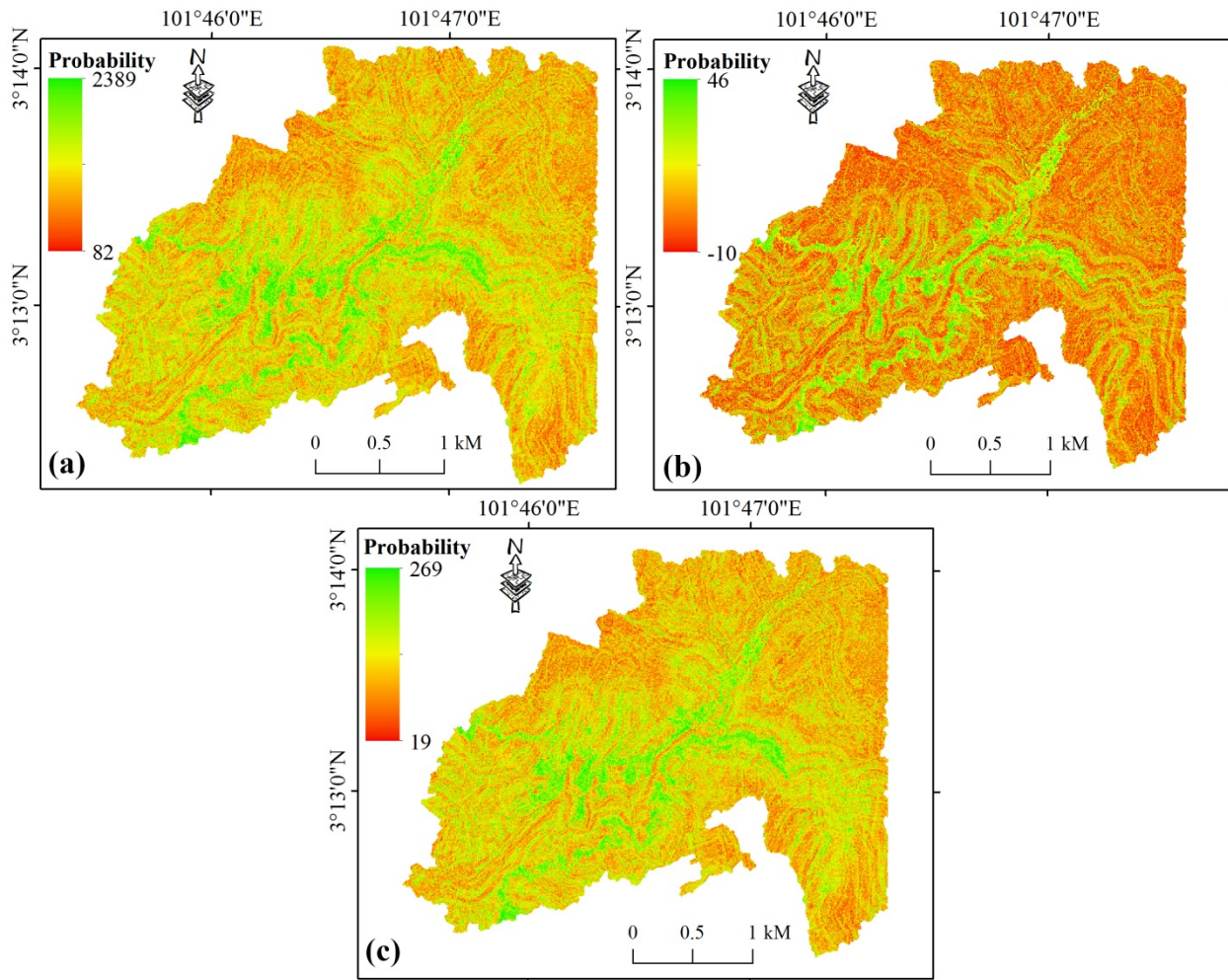


Fig. 7. Location of the pilot study area for testing the proposed ArcMAP tool.

503
504
505

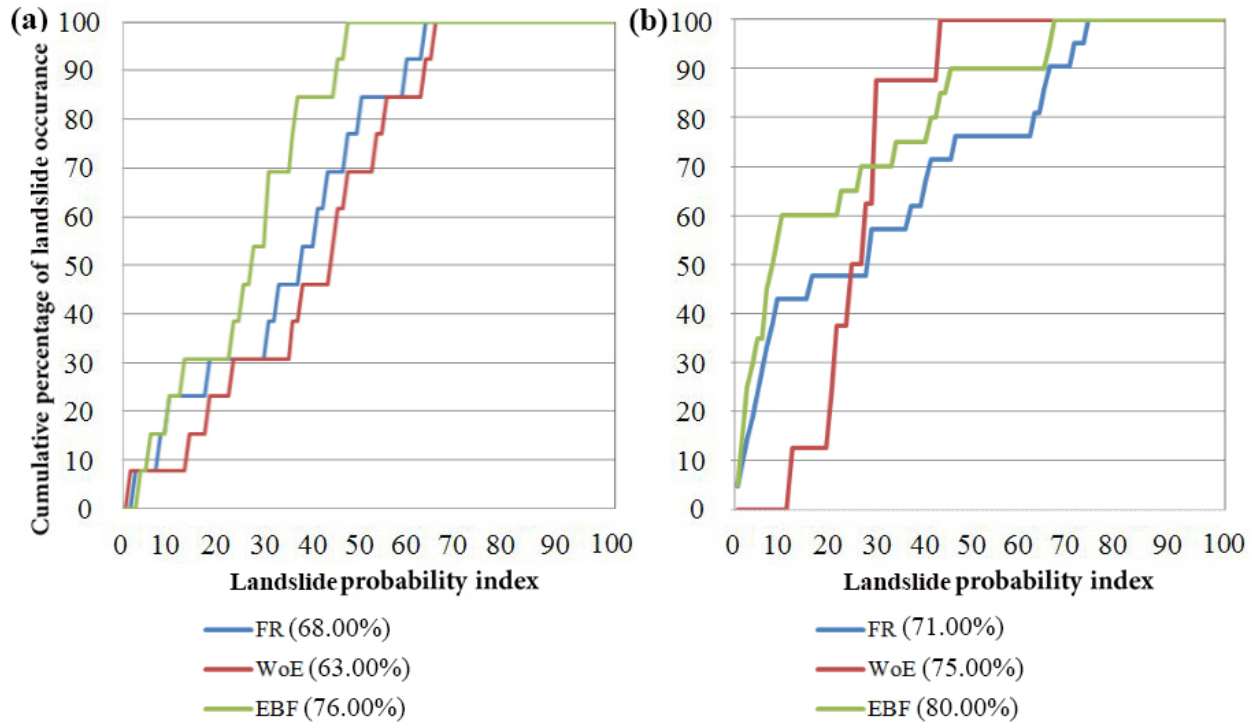


506

507

508

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



509

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

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527 **Appendix 1:**(the code of the three models coded in python)

528 **FR**

529 import arcpy

530 arcpy.CheckOutExtension("spatial")

531 arcpy.env.overwriteOutput = True

532 Input_raster = arcpy.GetParameterAsText(0)

533 Training_layer = arcpy.GetParameterAsText(1)

534 Processing_cell_size = arcpy.GetParameterAsText(2)

535 if Processing_cell_size == '#' or not Processing_cell_size:

536 Processing_cell_size = "5" # provide a default value if unspecified

537 Output_FR_layer = arcpy.GetParameterAsText(3)

538 Folder_Location = "c:\\"

539 arcpy.CreateFolder_management(Folder_Location, "FR_modeler")

540 arcpy.env.workspace = "C:\\FR_modeler"

541 arcpy.gp.Lookup_sa(Input_raster, "VALUE", "C:\\FR_modeler\\lookupp")

542 arcpy.gp.ZonalGeometryAsTable_sa("lookupp", "VALUE", "c:/fr_modeler/zonal", Processing_cell_size)

543 arcpy.AddField_management("zonal", "layer_pert", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")

544 arcpy.Statistics_analysis("zonal", "C:\\FR_Modeler\\ssummary.dbf", "AREA SUM", "")

545 arcpy.JoinField_management("zonal", "layer_pert", "C:\\FR_Modeler\\ssummary.dbf", "OID", "SUM_AREA")

546 arcpy.CalculateField_management("zonal", "layer_pert", "[AREA] / [SUM_AREA]", "VB", "")

547 arcpy.gp.TabulateArea_sa("lookupp", "VALUE", Training_layer, "VALUE", "c:/fr_modeler/tabulate", Processing_cell_size)

548 arcpy.AddField_management("tabulate", "layer_1", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")

549 arcpy.Statistics_analysis("tabulate", "C:\\FR_modeler\\ssssummary.dbf", "VALUE_1 SUM", "")

550 arcpy.JoinField_management("tabulate", "layer_1", "C:\\FR_modeler\\ssssummary.dbf", "OID", "SUM_VALUE_")

551 arcpy.CalculateField_management("tabulate", "layer_1", "[VALUE_1] / [SUM_VALUE_]", "VB", "")

552 arcpy.JoinField_management("zonal", "VALUE", "tabulate", "VALUE", "")

553 arcpy.AddField_management("zonal", "fr_layer", "LONG", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")

554 arcpy.CalculateField_management("zonal", "fr_layer", "Int ([layer_1]/ [layer_pert]*100)", "VB", "")

555 arcpy.CopyRows_management("zonal", "c:/fr_modeler/frresult", "#")

556 arcpy.gp.ReclassByTable_sa("lookupp", "frresult", "VALUE", "VALUE", "FR_LAYER", Output_FR_layer)

```

557 arcpy.Delete_management("C:\FR_modeler\sssummary.dbf", "#")
558 arcpy.Delete_management("C:\FR_modeler\summary.dbf", "#")
559 arcpy.Delete_management("C:\FR_modeler\zonal", "#")
560 arcpy.Delete_management("C:\FR_modeler\frresult", "ArcInfoTable")
561 arcpy.Delete_management("C:/FR_modeler/lookupp", "RasterDataset")
562 arcpy.Delete_management("C:/FR_modeler/tabulate", "ArcInfoTable")
563 arcpy.Delete_management("frresult", "#")
564
565 WoE
566 import arcpy
567 arcpy.CheckOutExtension("spatial")
568 arcpy.env.overwriteOutput = True
569 Input_raster = arcpy.GetParameterAsText(0)
570 Training_layer = arcpy.GetParameterAsText(1)
571 Processing_cell_size = arcpy.GetParameterAsText(2)
572 if Processing_cell_size == '#' or not Processing_cell_size:
573     Processing_cell_size = "5" # provide a default value if unspecified
574 Output_WOE_layer = arcpy.GetParameterAsText(3)
575 Folder_Location = "c:\\"
576 arcpy.CreateFolder_management(Folder_Location, "woe_modeler")
577 arcpy.env.workspace = "C:\woe_modeler"
578 arcpy.gp.Lookup_sa(Input_raster, "VALUE", "C:\woe_modeler\lookupp")
579 arcpy.gp.ZonalGeometryAsTable_sa("C:\woe_modeler\lookupp", "VALUE", "c:/woe_modeler/zonal", "5")
580 arcpy.AddField_management("zonal", "per_cell", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
581 arcpy.Statistics_analysis("zonal", "C:\woe_modeler\sssummary.dbf", "AREA SUM", "")
582 arcpy.JoinField_management("zonal", "per_cell", "C:\woe_modeler\sssummary.dbf", "OID", "SUM_AREA")
583 arcpy.CalculateField_management("zonal", "per_cell", "[AREA] / [SUM_AREA]", "VB", "")
584 arcpy.gp.TabulateArea_sa("lookupp", "VALUE", Training_layer, "VALUE", "c:/woe_modeler/tabulate", Processing_cell_size)
585 arcpy.AddField_management("tabulate", "per_depost", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
586 arcpy.Statistics_analysis("tabulate", "C:\woe_modeler\sssummary.dbf", "VALUE_1 SUM", "")

```



```

587 arcpy.JoinField_management("tabulate", "per_depost", "C:\woe_modeler\ssssummary.dbf", "OID", "SUM_VALUE_")
588 arcpy.AddField_management("tabulate", "deposit", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
589 arcpy.CalculateField_management("tabulate", "deposit", "[VALUE_1] + 0.0000000001", "VB", "")
590 arcpy.CalculateField_management("tabulate", "per_depost", "( [deposit] / [SUM_VALUE_] )", "VB", "")
591 arcpy.JoinField_management("zonal", "VALUE", "tabulate", "VALUE", "")
592 arcpy.AddField_management("zonal", "fr_layer", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
593 arcpy.CalculateField_management("zonal", "fr_layer", "[per_depost] / [per_cell]", "VB", "")
594 arcpy.AddField_management("zonal", "per_non_cl", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
595 arcpy.CalculateField_management("zonal", "per_non_cl", "1 - [per_cell]", "VB", "")
596 arcpy.AddField_management("zonal", "pr_non_dep", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
597 arcpy.CalculateField_management("zonal", "pr_non_dep", "1 - [per_depost]", "VB", "")
598 arcpy.AddField_management("zonal", "non_fr", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
599 arcpy.CalculateField_management("zonal", "non_fr", "[pr_non_dep] / [per_non_cl]", "VB", "")
600 arcpy.AddField_management("zonal", "w_positave", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
601 arcpy.CalculateField_management("zonal", "w_positave", "Log ( [FR_LAYER] + 1 )", "VB", "")
602 arcpy.AddField_management("zonal", "w_nagative", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
603 arcpy.CalculateField_management("zonal", "w_nagative", "Log ( [non_fr] + 1 )", "VB", "")
604 arcpy.AddField_management("zonal", "C", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
605 arcpy.CalculateField_management("zonal", "C", "[w_positave] - [w_nagative]", "VB", "")
606 arcpy.AddField_management("zonal", "S2_W_pos", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
607 arcpy.CalculateField_management("zonal", "S2_W_pos", "( 1 / [deposit] ) + ( 1 / ( [AREA] - [deposit] ) )", "VB", "")
608 arcpy.AddField_management("zonal", "s2_w_nag", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
609 arcpy.CalculateField_management("zonal", "s2_w_nag", "( 1 / ( [SUM_VALUE_] - [VALUE_12] ) ) + ( 1 / ( [SUM_AREA] -
610 [SUM_VALUE_] - [AREA] - [VALUE_12] ) )", "VB", "")
611 arcpy.AddField_management("zonal", "s_c", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
612 arcpy.CalculateField_management("zonal", "s_c", "Sqr ( [S2_W_pos] + [s2_w_nag] )", "VB", "")
613 arcpy.AddField_management("zonal", "woe", "LONG", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
614 arcpy.CalculateField_management("zonal", "woe", "Int ( [C] / [s_c] )", "VB", "")
615 arcpy.CopyRows_management("zonal", "c:/woe_modeler/woeresult", "#")
616 arcpy.gp.ReclassByTable_sa("lookupp", "woeresult", "VALUE", "VALUE", "WOE", Output_WOE_layer, "DATA")

```

```

617 arcpy.Delete_management("C:\FR_modeler\sssummary.dbf", "#")
618 arcpy.Delete_management("C:\FR_modeler\summary.dbf", "#")
619 arcpy.Delete_management("C:\FR_modeler\zonal", "#")
620 arcpy.Delete_management("C:\FR_modeler\frresult", "ArcInfoTable")
621 arcpy.Delete_management("C:\FR_modeler\lookupp", "RasterDataset")
622 arcpy.Delete_management("C:\FR_modeler\tabulate", "ArcInfoTable")
623 arcpy.Delete_management("frresult", "#")
624
625 EBF
626 import arcpy
627 arcpy.CheckOutExtension("spatial")
628 arcpy.env.overwriteOutput = True
629 Input_raster = arcpy.GetParameterAsText(0)
630 Training_layer = arcpy.GetParameterAsText(1)
631 Processing_cell_size = arcpy.GetParameterAsText(2)
632 if Processing_cell_size == '#' or not Processing_cell_size:
633     Processing_cell_size = "5" # provide a default value if unspecified
634 belief = arcpy.GetParameterAsText(3)
635 disbelief_layer = arcpy.GetParameterAsText(4)
636 uncertainty_layer = arcpy.GetParameterAsText(5)
637 plausibility_layer= arcpy.GetParameterAsText(6)
638 Folder_Location = "c:\\"
639 arcpy.CreateFolder_management(Folder_Location, "EBF_modeler")
640 arcpy.env.workspace = "C:\EBF_modeler"
641 arcpy.gp.Lookup_sa(Input_raster, "VALUE", "C:\EBF_modeler\lookupp")
642 arcpy.gp.ZonalGeometryAsTable_sa("lookupp", "VALUE", "c:\EBF_modeler\zonal", Processing_cell_size)
643 arcpy.AddField_management("zonal", "layer_pert", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
644 arcpy.Statistics_analysis("zonal", "C:\EBF_Modeler\sssummary.dbf", "AREA SUM", "")
645 arcpy.JoinField_management("zonal", "layer_pert", "C:\EBF_Modeler\sssummary.dbf", "OID", "SUM_AREA")

```

```

646 arcpy.CalculateField_management("zonal", "layer_pert", "[AREA] / [SUM_AREA]", "VB", "")
647 arcpy.gp.TabulateArea_sa("lookupp", "VALUE", Training_layer, "VALUE", "c:/EBF_modeler/tabulate", Processing_cell_size)
648 arcpy.AddField_management("tabulate", "layer_1", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
649 arcpy.Statistics_analysis("tabulate", "C:\EBF_modeler\ssssummary.dbf", "VALUE_1 SUM", "")
650 arcpy.JoinField_management("tabulate", "layer_1", "C:\EBF_modeler\ssssummary.dbf", "OID", "SUM_VALUE_")
651 arcpy.CalculateField_management("tabulate", "layer_1", "[VALUE_1] / [SUM_VALUE_]", "VB", "")
652 arcpy.JoinField_management("zonal", "VALUE", "tabulate", "VALUE", "")
653 arcpy.AddField_management("zonal", "eq1", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
654 arcpy.CalculateField_management("zonal", "eq1", "[layer_1] / (([AREA] - [VALUE_12]) / ([SUM_AREA] - [SUM_VALUE_]
655 )))", "VB", "")
656 arcpy.AddField_management("zonal", "eq2", "DOUBLE", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
657 arcpy.CalculateField_management("zonal", "eq2", "(( [SUM_VALUE_] - [VALUE_12]) / [SUM_VALUE_] ) / ((
658 [SUM_AREA] - [SUM_VALUE_] - [AREA] - [VALUE_12]) / ([SUM_AREA] - [SUM_VALUE_] ))", "VB", "")
659 arcpy.AddField_management("zonal", "belief", "LONG", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
660 arcpy.Statistics_analysis("zonal", "C:\EBF_modeler\summarryy.dbf", "eq1 SUM", "belief")
661 arcpy.JoinField_management("zonal", "belief", "C:\EBF_modeler\summarryy.dbf", "belief", "SUM_eq1")
662 arcpy.AddField_management("zonal", "disbelief", "LONG", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
663 arcpy.Statistics_analysis("zonal", "C:\EBF_modeler\summarryy.dbf", "eq2 SUM", "disbelief")
664 arcpy.JoinField_management("zonal", "disbelief", "C:\EBF_modeler\summarryy.dbf", "disbelief", "SUM_eq2")
665
666 arcpy.CalculateField_management("zonal", "belief", "Int ( ( [eq1] / [SUM_eq1] ) * 100 )", "VB", "")
667 arcpy.CalculateField_management("zonal", "disbelief", "Int ( ( [eq2] / [SUM_eq2] ) * 100 )", "VB", "")
668
669 arcpy.AddField_management("zonal", "uncertian", "LONG", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
670 arcpy.AddField_management("zonal", "plusabili", "LONG", "", "", "", "", "NULLABLE", "NON_REQUIRED", "")
671 arcpy.CalculateField_management("zonal", "uncertian", "100 - [belief] - [disbelief]", "VB", "")
672 arcpy.CalculateField_management("zonal", "plusabili", "100 - [disbelief]", "VB", "")
673 arcpy.CopyRows_management("zonal", "c:/EBF_modeler/beliefresult", "#")
674 arcpy.gp.ReclassByTable_sa("lookupp", "beliefresult", "VALUE", "VALUE", "BELIEF", belief, "DATA")
675 arcpy.CopyRows_management("zonal", "c:/EBF_modeler/disbeliefresult", "#")
676 arcpy.gp.ReclassByTable_sa("lookupp", "disbeliefresult", "VALUE", "VALUE", "DISBELIEF", disbelief_layer, "DATA")

```

677 arcpy.CopyRows_management("zonal", "c:/EBF_modeler/uncertaintyresult", "#")
678 arcpy.gp.ReclassByTable_sa("lookupp", "uncertaintyresult", "VALUE", "VALUE", "uncertain", "uncertainty_layer", "DATA")
679 arcpy.CopyRows_management("zonal", "c:/EBF_modeler/plusabilityresult", "#")
680 arcpy.gp.ReclassByTable_sa("lookupp", "plusabilityresult", "VALUE", "VALUE", "plusabili", "plusability_layer", "DATA")
681 arcpy.Delete_management("C:/EBF_modeler/beliefresult", "ArcInfoTable")
682 arcpy.Delete_management("C:/EBF_modeler/disbeliefresult", "ArcInfoTable")
683 arcpy.Delete_management("C:/EBF_modeler/plusabilityresult", "ArcInfoTable")
684 arcpy.Delete_management("C:/EBF_modeler/sssummary.dbf", "DbaseTable")
685 arcpy.Delete_management("C:/EBF_modeler/sssummary.dbf", "DbaseTable")
686 arcpy.Delete_management("C:/EBF_modeler/summarryy.dbf", "DbaseTable")
687 arcpy.Delete_management("C:/EBF_modeler/summaryy.dbf", "DbaseTable")
688 arcpy.Delete_management("C:/EBF_modeler/lookupp", "RasterDataset")
689 arcpy.Delete_management("C:/EBF_modeler/tabulate", "ArcInfoTable")
690 arcpy.Delete_management("C:/EBF_modeler/uncertaintyresult", "ArcInfoTable")
691 arcpy.Delete_management("C:/EBF_modeler/zonal", "ArcInfoTable")