1 **Exploring the potential relationship between the occurrence**

2 **of debris flow and landslide**

3 **Zhu Liang[1], Changming Wang[1], Donghe Ma[2] and Kaleem Ullah Jan Khan[1]**

4 [1]College of Construction Engineering, Jilin University, 130000 Changchun, People's Republic of

5 China;

6 [2]China Water Northeastern Investigation, Design and Research Co.Ltd.

7 E-mail:wangcm@jlu.edu.cn

8

9 **Abstract:** The aim of the present study is to explore the potential relationship between debris flow

10 and soil slide by establishing susceptibility zoning maps (SZM) separately with the use of random

11 forest. Longzi County, located in Southeastern Tibet, where historical landslides occurred

12 commonly, was selected as the study area. The work has been carried out with the following steps:

13 **(1)** An inventory map consisting of 448 landslides (399 soil slides and 49 debris flows) was

14 determined; **(2)** Slope units and 11 conditioning factors were prepared for the susceptibility

15 modelling of landslide while watershed units and 12 factors for debris flow; **(3)** SZM were

16 constructed for landslide and debris flow, respectively, with the use of random forest; **(4)** The

17 performance of two models were evaluated by 5-fold cross-validation using relative operating

18 characteristic curve (ROC), area under the curve (AUC) and statistical measures; **(5)** The potential

19 relationship between soil slide and debris flow was explored by the superimposition of two zoning

20 maps; **(6)** Gini index was applied to determined the major factors and analyze the difference

21 between debris flow and soil slide; **(7)** A combined susceptibility map with two kinds of disaster

22 was obtained. Two models had demonstrated great predictive capabilities, of which accuracy and

23 AUC was 87.33%, 0.902 and 85.17%, 0.892, respectively. The loose sources need by the debris

24 flow were not necessarily brought by the landslides although most landslides can be converted

25 into debris flow. The area prone to debris flow did not promote the occurrence of landslide. A

26 susceptibility zoning map composed of two or more natural disasters is comprehensive and

27 significant in this regard, which provides valuable reference for researches of disaster-chain and

28 engineering applications.

29 **Key words:** Landslide; Debris flow; Susceptibility; Random forest; Potential relationship

30

1. Introduction

32 Soil slide and debris flow are two kinds of natural phenomenon mainly occurring in mountainous

33 areas, which pose considerable threats to people, industries, and the environment directly or

34 indirectly. Generally, damages can be decreased to a certain extent by predicting the likely

35 location of future disasters (Pradhan, 2010). Thus, extensive research has been conducted for the

36 prediction and susceptibility assessment of soil slide and debris flow.

37   In geomorphology, a "landslide" is the movement of a mass of rock, debris or earth down a

38 slope, under the influence of gravity (Cruden and Varnes, 1996). According to different variables,

39 landslides can be divided into different types (Varnes, 1978). Debris flow is a specific type of

40 landslide, which can be defined as (Hungr et al. 2013): "Very rapid to extremely rapid surging

41 flow of saturated debris in a steep channel". Generally, slides that occur on a steep slope and

42 become disaggregated as they tumble down can transform into debris flows if they contain

Natural Hazards
and Earth System
Sciences
Discussions
EGU
Open Access

43   sufficient water for saturation (Huang et al., 2020). Therefore, slides may provide sufficient

44   material source for the occurrence of debris flow and most of the slides are accompanied by debris

45   flow. In the past, few scholars have specifically distinguished the slides and debris flow in terms

46   of susceptibility assessment (Alessandro et al., 2015; Guzzetti et al., 2005). In addition, some

47   scholars made separate evaluations of slides and debris flow (Park et al., 2011; Haydar et al.,

48   2016). Some scholars have proposed a coupled model of landslide-debris flow (Chiang et al., 2012;

49   Gomes et al., 2013). However, not every slide has evolved into a debris flow and the material

50   source of the debris flow is not necessary coming from slides. The formation and manifestations

51   of different types of landslides are different, especially debris flow, which is a kind of "wet

52   flow"(Varnes, 1978). In other words, there is no determined connection between debris flow and

53   other types of landslide. Therefore, the potential relationship between debris flow and other types

54   of landslide need further exploration.

55   Besides, the conditioning factors and mapping units involved in the susceptibility assessment

56   different kinds of landslides are not identical. Especially slope and water content are the most

57   critical factors controlling movements of debris flow (Takahashi 2007). Therefore, it is more

58   reasonable to evaluate the susceptibility of different kinds of landslides separately. As an example,

59   one landslide inventory map includes only one type of landslide, as does debris flow.

60   The methods of susceptibility assessment can be broadly classified as qualitative or

61   quantitative (Aleotti et al., 1999). Several methods and approaches have been proposed and tested

62   to ascertain susceptibility, such as physical-based approaches (Carrara et al., 2008), heuristic

63   methods (Blais et al., 2016) and statistically-based approaches (Reichenbach et al., 2018). In

64   addition, new machine learning models, such as neural networks (Park et al.,2013), support vector

65   machines (Colkesen et al.,2016) and random forest (RF) (Zhu et al., 2020a), have also been

66   applied.

67   The Longzi County in Southeastern Tibet is always exposed to slides and debris flow hazard

68   because of climatic and topographic conditions, which is chosen as the study area. The purpose of

69   the present study is to explore the potential relationship between the occurrence of debris flow and

70   soil slide by establishing susceptibility zoning maps separately with the use of random forest. It

71   also provides a reference for the study of landslide-debris flow, a common disaster chain.

72   ## 2. Materials

73   ## 2.1 Study area

74   The study area located in Longzi Township, Longzi County, Southeastern Tibet is bounded by

75   longitudes of 92°15'E and 92°45'E, latitudes of 28°10'N and 28°30'N (Fig.1). It covers an area of

76   about 535 km^2 with a population of more than 6000. The study area belongs to a semi-arid

77   temperate monsoon climate with the annual rainfall of 279 mm, mainly concentrated in May to

78   September. The seismic intensity within the area has a degree of VIII on the modified Mercalli

79   index.

80   The study area belongs to the zone of stratigraphic division of the Northern Himalayan block.

81   The strata is mainly composed of Mesozoic Cretaceous, Jurassic, Triassic, and Cenozoic units.

82   There were three common lithology observed during our field investigation: Siltstone from the

83   Laka Formation (K_1l); Conglomerates from the Weimei Formation (J_3w) and Quaternary slope

84   wash (Q_4^{el+dl}) from the Cenozoic strata.

85   The disasters in the study area mainly consist of rain-fed high frequency debris flows and

Natural Hazards
and Earth System
Sciences
Discussions
Open Access
EGU

86 landslides, which destroyed and flooded roads, bridges, farmlands, villages, etc., causing great

87 economic losses.

2.2 Landslide and debris flow inventory

89 The statistically-based susceptibility models are based on an important assumption: future

90 landslides will be more likely to occur under the conditions which led to the landslides past and

91 present (Varnes, 1984; Furlani and Ninfo, 2015). Therefore, a complete and accurate inventory

92 map is the key for model training and validation. In this study, data comes from historical records,

93 field surveys (**Fig.2 and Fig.3**) and interpretation of Google Earth images carried out in Google

94 Earth pro 7.1(**Fig.4**). Finally, a total of 399 soil slides and 49 debris flow locations were recorded

95 and mapped (**Fig.1**).

2.3 Mapping units

97 The selection of the mapping unit is an important pre-requisite for susceptibility modelling

98 (Guzzetti, 2006). The main mapping units commonly used for landslide and debris flow

99 susceptibility assessment are grid cells (Reichenbach et al., 2018). Despite its popularity and

100 operational advantages, grid-cells have clear drawbacks for susceptibility modelling (Guzzetti et

101 al., 1999). There is no physical relationship between a grid-cell and slope, while slope units can

102 make up for this deficiency. Depending on the landslide type, a slope unit may correspond to an

103 individual slope, an ensemble of adjacent slopes or a small catchment (Reichenbach et al., 2018).

104 The geometry of debris flow is better represented by a polygon or a set of polygons in vector

105 format. In the present study, adjacent slope units were applied to the susceptibility assessment of

106 soil slide. First-order sub-catchments, which is also called watershed unit, was applied to the

107   susceptibility of debris flow (Francesco et al., 2015; Zhu et al., 2020b). Accordingly, the study

108   area was divided into 1003 slope units for the modeling of soil slide or 174 watershed units for

109   debris flow.

2.4 Controlling factors and mapping

111   The selection of evaluation parameters is another key prerequisite to ensure that the model is

112   accurate and reasonable. With reference to previous studies (Ahmed et al., 2016; Xu et al., 2013;

113   Braun et al., 2018), there are differences in the controlling parameters used in soil slide and debris

114   flow susceptibility assessment. The occurrence of debris flow emphasizes the indispensability of

115   provenance, topography and triggering factors. Availability, reliability, and practicality of the

116   factor data were also considered (van Westen et al., 2008). In this paper, 11 controlling factors are

117   selected for the susceptibility assessment of landslide, including distance to fault, distance to road,

118   distance to river, annual rainfall, slope angle, aspect, plan curvature, profile curvature, topographic

119   wetness index, elevation and maximum elevation difference. Besides, a total of 12 controlling

120   factors, including basin area, main channel length, normalized difference vegetation index (NDVI),

121   drainage density, roundness, melton, average gradient of main channel, slope angle, maximum

122   elevation difference, annual rainfall, distance to fault and elevation were selected to fully reflect

123   the characteristics of the watershed for the susceptibility assessment of debris flow. Detailed

124   information on conditioning factors is shown in **Fig.5a~5m**. A brief description of each controlling

125   factor is given below.

126   Aspect, which is frequently used as landslide controlling factor (Dai and Lee, 2002), was

127   reclassified into 8 classes (**Fig. 5g**). Plan curvature and profile curvature were both considered and

128   reclassified into six classes (**Fig. 5b and 5e**). Generally, faults, rivers and roads play a key role in

129   the occurrence of landslides and were reclassified into seven classes using an interval of 1500m

130   (**Fig. 5i~k**). Topographic wetness index was reclassified into five classes (**Fig. 5h**).

131   NDVI reflects the vegetation conditions in the area and was reclassified into 5 classes(**Fig.**

132   **6b**). Drainage density is the ratio of the total drainage length to the watershed area and was

133   reclassified into six classes (**Fig.6 g**). Roundness refers to the ratio of the area of a basin to the

134   area of a circle with the same circumference and was reclassified into six classes (**Fig.6 d**) .

135   Melton ratio refers to the ratio of the degree of undulation in the watershed to the square root of

136   the arithmetic area of the watershed (Melton, 1965), which is reclassified into seven classes (**Fig.**

137   **6a**). Considering the correlation between the two controlling factors, basin area and main channel

138   length are represented by the same graph, which was reclassified into four classes (**Fig.6h**).

139   Average gradient of main channel, which is the ratio of the maximum elevation difference of main

140   channel to its linear length, was reclassified into six classes (**Fig. 6j**).

141   Rainfall is the only triggering factor to be considered for both landslide and debris flow in this

142   paper, which was reclassified into six classes (**Fig. 5a and Fig. 6c**). Slope angle is frequently employed

143   in both landslide and debris flow susceptibility mapping and was reclassified into six classes (**Fig. 5f**

144   **and Fig. 6i**). Maximum elevation difference reflects the kinetic energy condition and is reclassified

145   into 6 classes using an interval of 200m (**Fig. 5c and Fig. 6e**). Elevation was reclassified into five

146   classes (**Fig. 5d and Fig. 6f**), which has also been used by many authors (Ayalew and Yamagishi, 2005;

147   Pourghasemi et al. 2013a, b) .

148   Totally 18 factors are obtained by processing the row data in the ArcGIS 10.2 platform.

149   Morpholigical and topographic related factors were derived from the DEM with a resolution of 30

150   × 30 m. Geological related factors were extracted from 1:50000 geological maps. Rainfall is one

151   of the most important external factors inducing landslides and debris flow, which was determined

152   by ordinary kriging interpolation in ArcGIS by collecting data of 11 precipitation stations near the

153   area under study as a reference.

3. Methods

3.1 Sampling strategy and performance assessment

156   Statistical models for landslide susceptibility zoning reconstruct the relationships between

157   dependent and independent variables using training sets, and verify these relationships using

158   validation sets (Guzzetti et al., 2006a,b), which usually implies the partitioning of the inventory in

159   subsets. The sampling strategy affects the results of the susceptibility map (Yilmaz, 2010). Based

160   on temporal, spatial or random criteria, the partition of landslide inventories can be made (Chung

161   and Fabbri, 2003) and the most applied one is a one-time random selection (Reichenbach et al., 2018).

162   However, there is a need for a more reliable estimation of the model performance. The ability of

163   the models to classify independent test data was elaborated using a k-fold cross validation

164   procedure (k=5 in this paper) (James et al., 2013).

165   The computation of the area under the curve (AUC) is the most popular metrics to estimate

166   the quality of model , which has been applied for ROC curves(Green and Swets, 1966). It is one

167   of the most commonly used indicators. Three statistical metrics as accuracy, sensitivity, and

168   specificity are generally applied to assess the performance of the landslide susceptibility models

169   (Tien Bui et al. 2016).

170
$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Natural Hazards
and Earth System
Sciences

Discussions

171
$$Sensitivity = \frac{TP}{TP + FN}$$
$$Specificity = \frac{TN}{FP + TN}$$

172   (1)

173   where True Positives (TP), i.e., cells predicted unstable and observed unstable, True Negatives

174   (TN), i.e., cells predicted stable and observed stable, False Positives (FP), i.e., cells predicted

175   unstable but observed stable and False Negatives (FN), i.e., cells predicted stable but observed

176   unstable.

3.2 Random Forests

178   Random forest (RF) is a powerful ensemble-learning method and was first introduced by Breiman

179   (2001). RF uses the bagging technique (bootstrap aggregation) to select, at each node of the tree,

180   random samples of variables and observations as the training data set for model calibration.

181   Unselected cases (out of bag) are used to calculate the error of the model (OOB Error). The

182   increase in OOB error is proportional to the importance of the predictive variable (Breiman and

183   Cutler 2004). There are no restrictions on the types of variables, either numerical or categorical.

184   RF has the ability to reduce errors caused by unbalanced data, which is suitable for susceptibility

185   assessment.

186   In order to obtain reliable results of non-parametric models, their respective

187   hyperparameters must be optimized before application (Schratz et al., 2019). Scikit-learn package

188   (Pedregosa et al.,2011) in the programming software python version 3.7 was used for the

189   modeling. The number of trees   and the number of predictive variables used to split the nodes are

190   two user-defined parameters required to grow a random forest (Ahmed et al.,2016). The involved

191   parameters for modeling utilized in this study were shown in **Table 1.** Gini index (the larger the

192   value of the obtained result, the greater the contribution to the occurrence of landslide)

193   (Breiman,2001) was applied to analyze the major conditioning factors for both soil slide and

194   debris flow.

4. Results and verification

4.1 Landslide susceptibility mapping results

197   The predictive accuracy, ROC curves and AUC values of the RF model using training data were

198   showed in **Table 2** and **Fig. 7**. The RF model ensured a satisfactory performance of for classifying

199   landslides with sensitivity value of 91.62%. In terms of the classification of non-landslides zones,

200   specificity value also reached 89.06. An AUC equals to 1 indicates perfect prediction accuracy

201   (Vorpahl et al., 2012). The RF model had great performance in terms of AUC, with value of 0.976.

202   Standard error (St.), confidence interval (CI) at 95% and significance (Sig.) were applied as three

203   evaluation statistics. All these results indicated a reasonable goodness-of-fit for models with the

204   training dataset, for which the values were reasonably small.

205   Verifying the generalization ability of the model is a key step in prediction models as shown

206   in **Table 3** and **Fig. 7**. Accordingly, the values of sensitivity and specificity were 88.69% and

207   86.05%, respectively. The model also achieved a great performance in terms of AUC with value of

208   0.902. In comparison with the training model, the accuracy and AUC values have slightly

209   decreased, but still perform well.

210   The landslide susceptibility map was reclassified into five classes: very low (0~0.2), low

211   (0.2~0.4), moderate (0.4~0.6), high (0.6~0.8), very high (0.8~1) by using the equal spacing

212   method (**Fig.8**). The map should satisfy two spatial effective rules: (1) The existing disaster points

213   should belong to the high-susceptibility class and **(2)** The high-susceptibility class should cover

214   only small areas (Bui et al. 2012). The number of units belonging to very high class reached 179,

215   accounting for 17% (**Fig.9**). Disaster points were mostly in the dark (red or orange) areas. The

216   units belonging to moderate class accounted for the smallest proportion, at 13% (**Fig.9**).

217   The controlling factors with significant effects were selected and normalized as shown in

218   **Table 2**. The weight values of slope angle, distance to fault, plan curvature and topographic wetness

219   index was 0.21, 0.19, 0.17, 0.13 respectively, which was closely related to the occurrence of

220   landslide. The weight values of distance to road, maximum elevation difference, profile curvature

221   and elevation are less than 0.1 as 0.08, 0.08, 0.06, and 0.05, respectively (**Fig.10**).

222   4.2 Debris flow susceptibility mapping result

223   The debris flow susceptibility model perform well with a very high sensitivity and specificity

224   values as 87.80% and 88.89%, respectively. In terms of accuracy and AUC, the model had also a

225   great prediction performance with the value of 88.57% and 0.967 (**Fig.7**). Three evaluation

226   statistics also indicate a reasonable goodness-of-fit for the model.

227   **Table 3** shows that the values of sensitivity and specificity were 85.71% and 84.62%, which

228   were slightly decreased compared to the training model. However, the model had achieved a great

229   performance in terms of AUC, with value of 0.892.

230   The number of units belonging to very high-class reached to 26, which was accounting for

231   15% while the units belonging to high-class accounted for the smallest proportion at 13%. More

232   than half of the units (58%) belong to on a low or very low-class (**Fig.9**). Disaster points were

Natural Hazards
and Earth System
Sciences
Discussions

Open Access

EGU

233   mostly in the dark (Bright or deep red) areas (**Fig.8**).

234   The weight values of main channel length, roundness and slope angle were 0.25, 0.16, 0.14

235   respectively, which has significant influence on the occurrence of debris flow. The weight values

236   of elevation, maximum elevation difference, melton and basin area are close to 0.1, which are 0.13,

237   0.12, 0.1, and 0.1 respectively(**Fig.10**).

238 **4.3 Analysis and comparison of landslide and debris flow**

239 **susceptibility**

240   It is worth comparing the two susceptibility zoning maps. In terms of prediction accuracy, the

241   values of sensitivity, specificity and AUC of landslide model were slightly higher than that of

242   debris flow. However, both models achieved high predictive performance. Therefore, the landslide

243   and debris flow susceptibility assessment models based on RF are reliable. The purpose of the

244   present study is to explore the potential relationship between landslides and debris flows by

245   establishing the respective susceptibility zoning maps. Figure 11 shows the overlapping areas

246   between debris flow and landslide in high or very high-class of susceptibility zoning map. It can

247   be seen that most of the areas with high or very high-class in the map of debris flow are covered

248   with landslides. However, there are also non-overlapping areas between the two zoning maps.

249   There are 23 watershed units belonging to high-class in the debris flow susceptibility zoning map

250   (**Fig.8**), of which 17 units are covered with high or very high-class slope units in the landslide

251   zoning map (**Table 5**). In addition, there are 4 watershed units covered with low or very low class

252   slope units. In the same way, 19 watershed units belonging to very high-class are covered with

253   high or very high-class slop units and 4 watershed units with low or very low-class slop units. In

254 other words, more than 70% of the high or very high-class watershed units are covered with high

255 or very high-class slope units. However, there are still 30% of watershed units with high or very

256 high-class without the distribution of slope units in corresponding grades. It validated the previous

257 view that most of landslides can be transformed into debris flows. Factor analysis was applied to

258 further analyze the reasons for the difference. 36 watershed units with distribution of high or very

259 high-grade slope units were taken as model 1 and the left 8 watershed units as model 2 (**Table 5**).

260 The KMO (Kaiser-Meyer-Olkin) and significance (Sig.) testing are two statistical parameters

261 which ensured the feasibility before application. The KMO values were 0.766 and 0.643

262 respectively, which indicated that the correlation between variables was obvious and suitable for

263 factor analysis (**Table 6**). In model 1, the cumulative contribution rate of the first three factors (C1,

264 C2 ,C3) reached to 83.6%, while the cumulative contribution rate of the first four factors (F1,

265 F2 ,F3 and F4) reached to 80.5% for model 2 (**Table 7**). According to the correlation coefficient

266 of each common factor (**Table 7**), the first common factor mainly highlighted the information of

267 basin area, main channel length and maximum elevation difference. Similarly, the second and the

268 third common factor highlighted the information of slope angle and elevation and roundness,

269 respectively. The difference between the two models is that the second model has the fourth

270 common factor (**Table 8**), which emphasized the effects of rainfall and distance to the fault. The

271 transformation from a landslide to a debris flow often occurs during heavy rainfall (Takahashi,

272 1978), and the landslides are the source area. But landslides are not the only source of debris flows.

273 The loose material distributed in the basin is not necessarily caused by landslide.

274   In turn, we analyze the distribution of high or very high-class slope units in watershed units.

275 The landslide zoning map was put at the bottom floor and the debris flow zoning map on the top

276   floor (**Fig. 11**). There are 167 slope units belonging to high-class, of which 68 units (accounting

277   for about 40%) are distributed in the area of high or very high-class watershed units in the debris

278   flow zoning map (**Table 9**). Besides, 69 slope units (accounting for about 41%) are distributed in

279   the area of low or very low-class watershed units. Similarly, 53 slope units (accounting for about

280   30%) belonging to very high-class are distributed in the area of high or very high-class watershed

281   units and 88 slope units (accounting for about 50%) in low or very low-class slop units (**Table 9**).

282   Comparing with the extent of the landslide affecting the debris flow, the impact of the debris flow

283   on the landslide is not obvious. It indicated that the area prone to debris flow does not promote the

284   occurrence of landslides.

285   Finally, we took the center of gravity of 1,003 slope units as the potential hazard points and

286   spread them over 174 watershed units. Thus, a combining susceptibility zoination map for

287   landslide and debris flow was obtained (**Fig.11**). The darker the color, the higher the class of

288   susceptibility will be. It can be seen that the susceptibility in the south is generally higher than that

289   in the north, and the area in the southwest is disaster-prone. The northeast and central locations in

290   the area are less likely to be affected by landslides and belong to low-susceptibility areas. Green or

291   yellow dots, which refer to slope units with very low or low- class in the landslide zoning map,

292   mainly distributed in light-colored areas but there are also quite a few green or yellow dots

293   distributed in dark areas, which means that the occurrence of debris flow not necessarily depend

294   on landslides. Blue or black spots are mainly distributed in dark areas but there are also quite a

295   few blue or black spots distributed in dark light areas, which means that landslide is not the only

296   condition for debris flow to occur. Most of the watershed units are distributed with two or more

297   colored dots, which means that there would be multiple slope units with different susceptibility

298   class in the same watershed. According to the combining susceptibility zoning map of landslide

299   and debris flow, the study area can be divided into 4 categories: **(1)** Low or very low-class

300   watershed units coupled with low or very low-class slope units; **(2)** Low or very low-class

301   watershed units coupled with high or very high-class slope units; **(3)** High or very high-class

302   watershed units coupled with low or very low-class slope units; **(4)** High or very high-class

303   watershed units coupled with high or very high-class slope units. We assume that the occurrence

304   of landslides can bring rich sources of debris flow, thereby promoting or aggravating the outbreak

305   of debris flow, that is, forming a landslide-debris flow disaster chain. Therefore, the susceptibility

306   assessment of the landslide-debris flow chain in the study area can be roughly divided into three

307   classes, which are low, moderate and high (**Table 10**).

5. Discussion

5.1 Method used for modeling

Many researchers have used different statistically-based methods to evaluate the susceptibility of landslides or debris flows. Logistic regression and discriminant analysis are the most popular methods to use in traditional multivariate statistical analysis. The performance of new learning machines, such as support vector machines and neural networks, has also been verified. RF, as a newly integrated learning machine, has less application in landslide and debris flow analysis. Actually, RF have powerful data processing capabilities and can simultaneously solve problems such as high-dimensional, unbalanced and data loss, which are common in geological disaster assessment. Most importantly, RF can compare the important differences between features and have ability to reduce errors caused by unbalanced data and, which achieved strong generalization properties (Zhu et al., 2020a).

5.2 Potential relationship between landslide and debris flow

There is a certain similarity in the evaluation of the susceptibility of landslide and debris flow from the concept, the selection of controlling factors and the application of modeling strategies. Therefore, some researchers have neglected the difference between landslide and debris flow i.e to express two different disasters with the same susceptibility zoning map (Ciurleo et al., 2016; Ciurleo et al., 2017; Persichillo et al., 2017;). However, similarity does not always mean consistency. Many researchers have previously conducted studies into the debris flow mobilization from shallow landslide using a coupled methodology. They are interested in the dynamic simulation of debris flow based on the prediction of landslide susceptibility (Wang et al., 2013;

329   Fan et al., 2017). However, not every landslide evolves into a debris flow, which means that the

330   analysis process is highly selective or uncertain. In the same way, the source of the debris flow is

331   not limited to landslide. There, the potential relationship between landslide and debris flow needs

332   to be discussed more reasonably and effectively. In this study, the corresponding influencing

333   factors and mapping units are selected to establish landslide and debris flow susceptibility zoning

334   maps, respectively. The potential relationship between landslide and debris flow is explored in two

335   ways: **1)** Superimposing the high or very high-class susceptibility areas in the two maps; **2)**

336   Transforming the slope units into points and distributed them on the watershed units. The

337   relationship between landslide and debris flow is illustrated by the distribution of slope units of

338   different grades on the watershed units with different prone grades.

5.3 Necessity and feasibility of combining multiple natural
disaster susceptibility zoning maps

339
340

341   Previous studies on susceptibility zoning mapping of disaster have agreed that one disaster

342   corresponds to one map. Multiple disasters may be bred simultaneously in a watershed unit and it

343   will cause some confusion in practical. For example, the probability of a disaster occurring in a

344   watershed is negligible, while it is high of another disaster. Therefore, we need to combine

345   multiple zoning maps at the same time to give a comprehensive evaluation, which is arduous to

346   achieve. On the one hand, the prediction accuracy and error of different zoning maps should be

347   similar or even consistent. On the other hand, the dimensions of the mapping unit should be

348   consistent or complementary. The fact that the appropriate prediction method and mapping units

349   applied to the two disasters makes it possible to merge the two zoning maps. Disaster risk is

Natural Hazards
and Earth System
Sciences
Discussions
Open Access
EGU

350 higher in landslide-debris flow chain, causing significant loss of life and property. Therefore, two

351 natural disasters with potential relationship are simultaneously reflected in the same susceptibility

352 zoning map, which can better guide the implementation of engineering, such as landslide-debris

353 flow disaster chain.

6. Conclusion

355 In this study, susceptibility assessment models for landslide and debris flow are established

356 through RF, respectively and the performance of the models are excellent in terms of accuracy and

357 goodness of fit. The potential relationship between landslide and debris flow is discussed by the

358 superimposition of two zoning maps and the following conclusions can be drawn:

359 (1) The landslide and debris flow susceptibility assessment models based on random forest have

360 great performance of accuracy and goodness-of-fit and have the ability to analyze the relative

361 importance of different impact factors, which is suitable for the evaluation of natural disasters;

362 (2) Although most landslides will be converted into debris flow, the landslides are not necessarily

363 the source of debris flow, and the loose sources carried by the debris flow are not necessarily

364 brought by the landslides;

365 (3) By comparing the extent of the landslide affecting the debris flow, the impact of the debris

366 flow on the landslide is not obvious, which indicates that the area prone to debris flow does not

367 promote the occurrence of landslides;

368 (4) A susceptibility zoning map composed of two or more natural disasters is more

369 comprehensive and significant, which provides valuable reference for researchers and engineering

370 applications.

371

372 **Data availability.** The data used to support the findings of this study are included within the

373 article.

374 **Author contributions.** ZL was responsible for the writing and graphic production of the paper.

375 CMW was responsible for the revision of the paper. DHM was responsible for calculation. KUJK

376 was responsible for the translation.

377 **Competing interests.** The authors declare that they have no conflict of interest.

378 **Special issue statement.** This article is part of the special issue"Resilience to risks in built

379 environments". It is not associated with a conference.

380 # Acknowledgements

383 # References:

384 Ayalew L, Yamagishi H (2005) The Application of GIS-based logistic regression for landslide

385 susceptibility mapping in the Kakuda–Yahiko Mountains, central Japan. Geomorphology

386 65:15–31. doi:10.1016/j. geomorph. 2004.06.010

387 Alessandro Trigila, Carla Iadanza, Carlo Esposito, et al. Comparison of Logistic Regression and

388 Random Forests techniques for shallow landslide susceptibility assessment in Giampilieri

389 (NE Sicily, Italy). 2015, 249:119-136.

390 Ahmed Mohamed Youssef, Hamid Reza Pourghasemi, Zohre Sadat Pourtaghi, et al. Erratum to:

391 Landslide susceptibility mapping using random forest, boosted regression tree, classification

392   and regression tree, and general linear models and comparison of their performance at Wadi

393   Tayyah Basin, Asir Region, Saudi Arabia. 2016, 13(5):1315-1318.

394   Anika Braun, Elias Leonardo Garcia Urquia, Rigoberto Moncada Lopez, et al. Landslide

395   Susceptibility Mapping in Tegucigalpa, Honduras, Using Data Mining Methods. 2018.

396   Breiman, L.,2001. Random Forests. Machine learning, 45(1): 5-32. doi:https://doi.org/10.1023/A:

397   10 1093340

398   Breiman L, Cutler A (2004) http://www.stat.berkeley.edu/users/Breiman/RandomForests/ccpapers.

399   H-Tml

400   Bui DT, Pradhan B, Lofman O, Revhaug I, Dick OB (2012) Landslide susceptibility assessment in

401   the Hoa Binh Province of Vietnam: a comparison of the Levenberg-Marquardt and Bayesian

402   regularized neural networks. Geomorphology. doi:10.1016/j.geomorph.2012.04.023

403   Blais-Stevens A, Behnia P (2016) Debris flow susceptibility mapping using a qualitative heuristic

404   method and flow-R along the Yukon Alaska Highway Corridor, Canada. Nat Hazard Earth

405   Syst Sci 16(2):449–462.

406   Chung, C.F., Fabbri, A.G., 2003. Validation of spatial prediction models for landslide hazard

407   mapping. Nat. Hazards 30,451–472.

408   Carrara A, Crosta G, Frattini P (2008) Comparing models of debris-flow susceptibility in the

409   alpine environment. Geomorphology 94:353-378

410   Chunxiang Wang, Hideaki Marui, Gen Furuya, et al. Two Integrated Models Simulating Dynamic

411   Process of Landslide Using GIS. 2013.

412   Chung, C.F., Fabbri, A.G., 2003. Validation of spatial prediction models for landslide hazard

413   mapping. Nat. Hazards 30, 451–472.

414 Colkesen, I., Sahin, E.K., Kavzoglu, T., 2016. Susceptibility mapping of shallow landslides using

415   kernel-based Gaussian process, support vector machines and logistic regression. African

416   Earth Sciences]–>J. Afr. Earth Sci. 118, 53–64.

417 Cruden, D.M., Varnes, D.J., 1996. Landslide types and processes. In: Turner, A.K., Schuster, R.L.

418   (Eds.), Landslides, Investigation and Mitigation, Special Report 247. Transportation Research

419   Board, Washington D.C., pp. 36–75 ISSN: 0360-859X, ISBN: 030906208X.

420 Dai, F.C., Lee, C.F., 2002. Landslide characteristics and slope instability modelling using GIS,

421   Lantau Island, Hong Kong. Geomorphology 42, 213–228.

422 D. W. Park, N. V. Nikhil, S. R. Lee. Landslide and debris flow susceptibility zonation using

423   TRIGRS for the 2011 Seoul landslide event. 2013, 13(11):2833-2849.

424 Furlani, S., Ninfo, A., 2015. Is the present the key to the future? Earth-Sci. Rev. 142 (C),38–46.

425 Francesco Bregoli, Vicente Medina, Guillaume Chevalier, et al. Debris-flow susceptibility

426   assessment at regional scale: Validation on an alpine environment. 2015, 12(3):437-454.

427 Green, D.M., Swets, J.M., 1966. Signal Detection Theory and Psychophysics. John Wiley and

428   Sons, New York ISBN: 0-471-32420-5.

429 Guzzetti, F., Carrara, A., Cardinali, M., Reichenbach, P., 1999. Landslide hazard evaluation: a

430   review of current techniques and their application in a multi-scale study, Central Italy.

431   Geomorphology 31, 181–216.

432 Guzzetti. F, M. Galli, P. Reichenbach, et al. Landslide hazard assessment in the Collazzone area,

433   Umbria, Central Italy. 2006, 6(1):115.

434 Guzzetti, F., Galli, M., Reichenbach, P., Ardizzone, F., Cardinali, M., 2006a. Landslide hazard

435   assessment in the Collazzone area, Umbria, central Italy. Nat. Hazard. Earth Syst. Sci. 6,

Natural Hazards
and Earth System
Sciences
Discussions

436   115–131.

437   Guzzetti, F., Reichenbach, P., Ardizzone, F., Cardinali, M., Galli, M., 2006b. Estimating the

438   quality of landslide susceptibility models. Geomorphology 81, 166–184.

439   Gomes, R. A. T., Guimaraes, R. F., Carvalho Júnior, O. A., Fernandes, N. F., and Amaral Jr., E. V.:

440   Combining spatial models for shallow landslides and debris flows prediction, Remote Sens.,

441   5,2219–2237, 2013

442   Hungr, O., Leroueil, S., Picarelli, L., 2013. The Varnes classification of landslide types, an update.

443   Landslides 11 (2), 167–194.

444   Huang Xiaohu, Guo Fei, Deng miaolin, Yi Wu, Huang Haifeng.Understanding the deformation

445   mechanism and threshold reservoir level of the floating weight-reducing landslide in the

446   Three Gorges Reservoir Area, China. Landslides DOI 10.1007/s10346-020-01435-1.

447   Haydar Y. Hussin, Veronica Zumpano, Paola Reichenbach, et al. Different landslide sampling

448   strategies in a grid-based bi-variate statistical susceptibility model. 2016, 253:508-523.

449   James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An Introduction to Statistical Learning.

450   Springer, New York, p. 441.

451   Lee, S., Dan, N.T., 2005. Probabilistic landslide susceptibility mapping on the Lai Chau province

452   of Vietnam: focus on the relationship between tectonic fractures and landslides.

453   Environmental Geology 48, 778–787.

454   Linfeng Fan, Peter Lehmann, Brian McArdell, et al. Linking rainfall-induced landslides with

455   debris flows runout patterns towards catchment scale hazard assessment. 2017, 280:1-15.

456   Liqiang Tong, Wensheng Qi, Guoying An,Chunling Liu. Remote sensing survey of major

457   geological disasters in the Himalayas[J].Journal of engineering geology,2019,27(03):496.

458   Melton M A．The Geomorphic and Paleoclimatic Significance of Alluvial Deposits in Southern

459   Arizona: A Reply［J］. The Journal of Geology，1965，73(1) : 102 －106．

460   Mariantonietta Ciurleo, Michele Calvello, Leonardo Cascini. Susceptibility zoning of shallow

461   landslides in fine grained soils by statistical methods. 2016, 139:250-264.

462   Mariantonietta Ciurleo, Leonardo Cascini, Michele Calvello. A comparison of statistical and

463   deterministic methods for shallow landslide susceptibility zoning in clayey soils. 2017,

464   223:71-81.

465   Maria Giuseppina Persichillo, Massimiliano Bordoni, Claudia Meisina, et al. Shallow landslides

466   susceptibility assessment in different environments. 2017, 8(2):748-771.

467   P. Aleotti, R. Chowdhury. Landslide hazard assessment: summary review and new perspectives.

468   1999, 58(1):21-44.

469   Pradhan, B., 2010. Landslide susceptibility mapping of a catchment area using frequency ratio,

470   fuzzy logic and multivariate logistic regression approaches. J. Indian Soc. Remote Sens. 38

471   (2), 301–320.

472   Pedregosa,F., Varoquaux,G., Gramfort, A., et al., 2011. Scikit-Learn: Machine Learning in Python.

473   Journal of machine learning research, 12(10): 2825-2830.

474   P. Vorpahl, H. Elsenbeer, M. Märker, and B. Schröder, "How can statistical models help to

475   determine driving factors of landslides?" Ecol. Model., vol. 239, pp. 27–39, 2012

476   Pourghasemi HR, Moradi HR, Fatemi Aghda SM (2013a) Landslide susceptibility mapping by

477   binary logistic regression, analytical hierarchy process, and statistical index models and

478   assessment of their performances. Nat Hazards 69:749–779. doi:10.1007/s11069-013-0728-5

479   Pourghasemi HR, Pradhan B, Gokceoglu C, Mohammadi M, Moradi HR (2013b) Application of

480   weights-of-evidence and certainty factor models and their comparison in landslide

481   susceptibility mapping at Haraz watershed, Iran. Arab J Geosci 6(7):2351–2365.

482   doi:10.1007/s12517-012-0532-7

483   Reichenbach,P.,Rossi,M.,Malamud,B.D.,et al.,2018.A Review of Statistically-Based Landslide

484   Susceptibility Models. Earth-Science Reviews, 180(5): 60-91.

485   doi:https://doi.org/10.1016/j.earscirev. 2018. 03. 001

486   Shou-Hao Chiang, Kang-Tsung Chang, Alessandro C. Mondini, et al. Simulation of event-based

487   landslides and debris flows at watershed level. 2011, 138(1):306-318.

488   Takahashi, T., 1978. Mechanical characteristics of debris flow. Journal of the Hydraulics Division

489   104, 1153–1169.

490   Takahashi T(2007) Debris Flow:Mechanics, Prediction and Countermeasures. Taylor & Francis,

491   448 p. ISBN 978-0-415-43552-9

492   Varnes DJ (1978) Slope movement types and processes, in Schuster, R.L., and Krizek, R.J., eds.,

493   Landslides: Analysis and control, National Research Council, Washington, D.C.,

494   Transportation Re- search Board, National Academy Press, Special Report 176, p. 11–33

495   Varnes, D.J., IAEG Commission on Landslides and other Mass-Movements, 1984. Landslide

496   Hazard Zonation: A Review of Principles and Practice. The UNESCO Press, Paris (63 pp).

497   van Westen CJ, Castellanos E, Kuriakose SL (2008) Spatial data for landslide susceptibility,

498   hazard, and vulnerability assessment: an overview. Eng Geol 102(3–4):112–131

499   Xu WB, Yu WJ, et al. (2013) Debris flow susceptibility assessment by GIS and information value

500   model in a large-scale region, Sichuan Province (China).Natural Hazards 65(3):1379-1392.

501   Yilmaz, I., 2010. The effect of the sampling strategies on the landslide susceptibility mapping by

502   conditional probability and artificial neural networks. Environ. Earth Sci. 60,505–519.

503   Zhu Liang, Changming Wang, , Zhi-Min Zhang and Kaleem-Ullah-Jan Khan. A comparison of

504   statistical and machine learning methods for debris flow susceptibility mapping. Stoch

505   Environ Res Risk Assess (2020a) https://doi.org/10.1007/s00477-020-01851-8

506   Zhu Liang, Changming Wang, Songling Han, Kaleem Ullah Jan Khan, and Yiao Liu.

507   Classification and susceptibility assessment of debris flow based on a semi-quantitative

508   method combination of the fuzzy C-means algorithm, factor analysis and efficacy

509   coefficient.Nat.   Hazards   Earth   Syst.   Sci.,   20,   1287–1304,   2020b

510   https://doi.org/10.5194/nhess-20-1287-2020

511   **Table 1** The optimized parameters of RF

Methods	Parameters
RF	Number of iterations, 100; number of execution slots, 10; 1oob_score = true; percentage of bag size, 0.382; max_features, sqrt; n_estimators, 500

512   **Table 2** Models' performance using training dataset

Metrics	Landslide	Debris flow
TP（%）	88.71	87.80
TN（%）	91.89	88.89
FP（%）	11.29	12.20
FN（%）	8.11	11.11
Sensitivity（%）	91.62	88.77
Specificity（%）	89.06	87.93
Accuracy（%）	90.65	88.57

AUC	0.976	0.967

513 **Table 3** Models' performance using verification dataset

Metrics	Landslide	Debris flow
TP（%）	85.56	85.71
TN（%）	89.09	84.62
FP（%）	14.44	14.29
FN（%）	10.91	15.38
Sensitivity（%）	88.69	84.79
Specificity（%）	86.05	85.55
Accuracy（%）	87.33	85.17
AUC	0.902	0.892

514 **Table 3** Variables importance assigned for landslide

Test group	Slope angle	Distance to fault	Plan curvature	Topographic wetness index	Distance to road	Maximum elevation difference	Profile curvature	Elevation
Landslide	0.21	0.19	0.17	0.13	0.08	0.07	0.06	0.05

515 **Table 4** Variables importance assigned for debris flow

Test group	Main channel length	Roundness	Slope angle	Elevation	Maximum elevation difference	Melton	Basin area
Debris flow	0.25	0.16	0.14	0.13	0.12	0.1	0.1

516 **Table 5** The overlap number of debris flow and landslide height and very high-class mapping units

Landslide ⟋ Debris flow	Very low	Low	High	Very high
High	3/23	1/23	5/23	12/23
Very high	2/26	2/26	8/26	11/26

517 **Table 6** Statistical parameters of the two models

Statistical parameters ⟋ Model	Model 1	Mode 2
KMO	0.766	0.643
Sig.	0.001	0.003

518 **Table 7** The correlation coefficients between common factors and primitive variables

Factor	F1	F2	F3
NDVI	0.386	-0.336	-0.621
Basin area	0.897	-0.007	0.041
Main channel length	0.984	0.046	-0.023
Slop angle	-0.223	0.829	0.455
Maximum elevation difference	0.744	0.66	0.011
Rainfall	-0.768	0.33	0.201
Average gradient of main channel	-0.753	0.544	0.106
Drainage density	-0.844	0.06	0.015
Roundness	0.331	0.14	0.818
Elevation	0.133	0.846	0.382
Distance to fault	-0.16	0.211	0.421

Melton	-0.625	0.737	0.149
Contribution rate (%)	41.2	24.7	16.7
Accumulative contribution (%)	41.2	65.9	83.6

519   **Table 8** The correlation coefficients between common factors and primitive variables

Factor	C1	C2	C3	C4
NDVI	0.042	-0.079	-0.279	-0.813
Basin area	0.802	-0.344	0.057	0.009
Main channel length	0.885	0.126	-0.196	0.227
Slop angle	0.009	0.748	0.58	-0.057
Maximum elevation difference	0.801	0.434	-0.128	0.144
Rainfall	0.197	-0.076	-0.487	0.637
Average gradient of main channel	-0.744	0.205	0.15	-0.23
Drainage density	-0.776	-0.176	-0.267	0.117
Roundness	-0.014	0.022	0.896	-0.002
Elevation	0.34	0.746	0.25	0.326
Distance to fault	0.31	0.289	-0.344	0.757
Melton	-0.182	0.932	-0.192	0.061
Contribution rate (%)	29.2	20.3	15.2	15.8
Accumulative contribution (%)	29.2	49.5	64.7	80.5

520   **Table 9** The overlap number of landslide and debris flow height and very-high class mapping units

Debris flow / Landslide	Very low	Low	High	Very high
High	36/167	33/167	25/167	43/167
Very high	48/179	40/179	25/179	28/179

521 **Table 10** Comprehensive evaluation of landslide-debris flow susceptibility

Debris flow / Landslide	Low or Very low	High or Very high
Low or Very low	Low	Moderate
High or Very high	Moderate	High
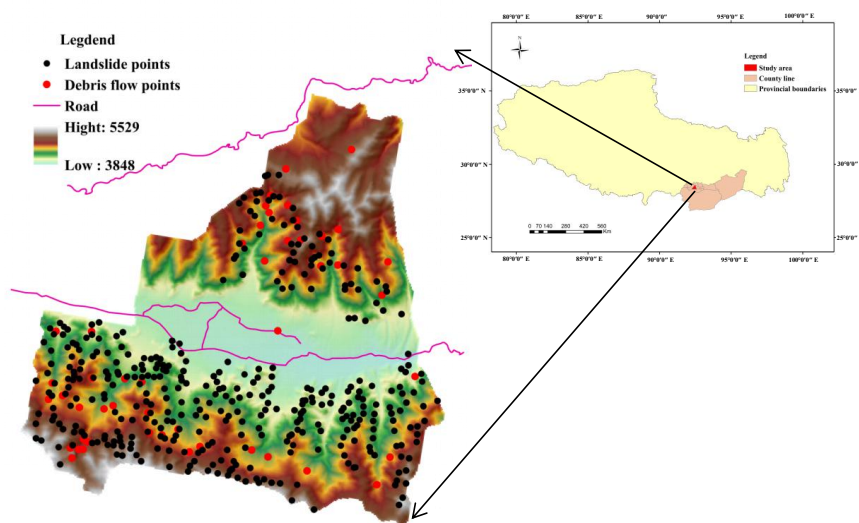
522

523 **Fig.1.** Location map of the study area showing landslide and debris flow inventory.

Natural Hazards
and Earth System
Sciences

Discussions

Open Access

EGU

524

525
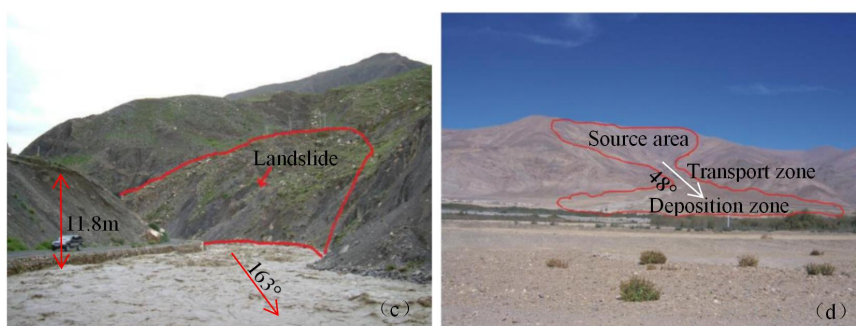
Fig.2. Photos of landslide or debris flow: **(a)** Lunba landslide in a tributary; **(b)** Zhenqiong landslide in

527   Jiayu village; **(c)** Debris flow in Misha Township; **(d)** Debris flow in Lelong Village.

528
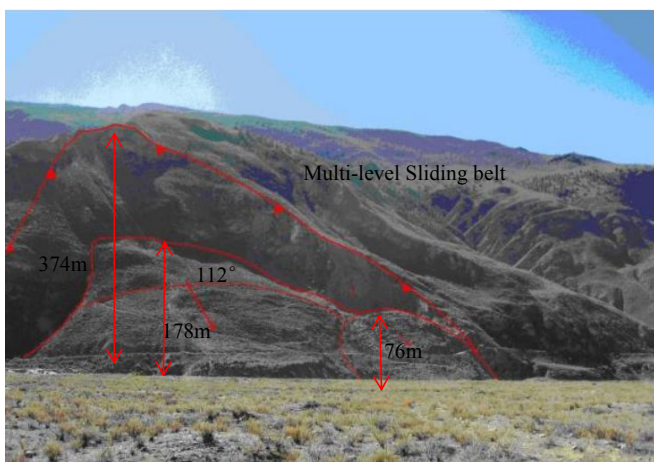
529   **Fig.3.** Multistage landslide in Xiongqu village
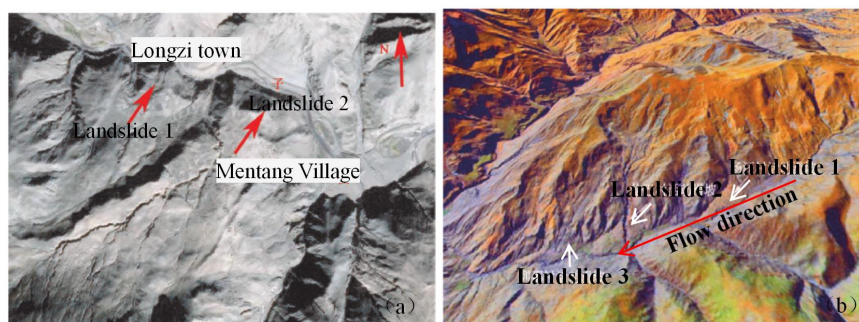
Fig.4. Stereo remote sensing map of landslides in Longzi Township (Tong et al., 2019): （a）Landslides

in Longzi town; **(b)** Landslides in Malu town.

533
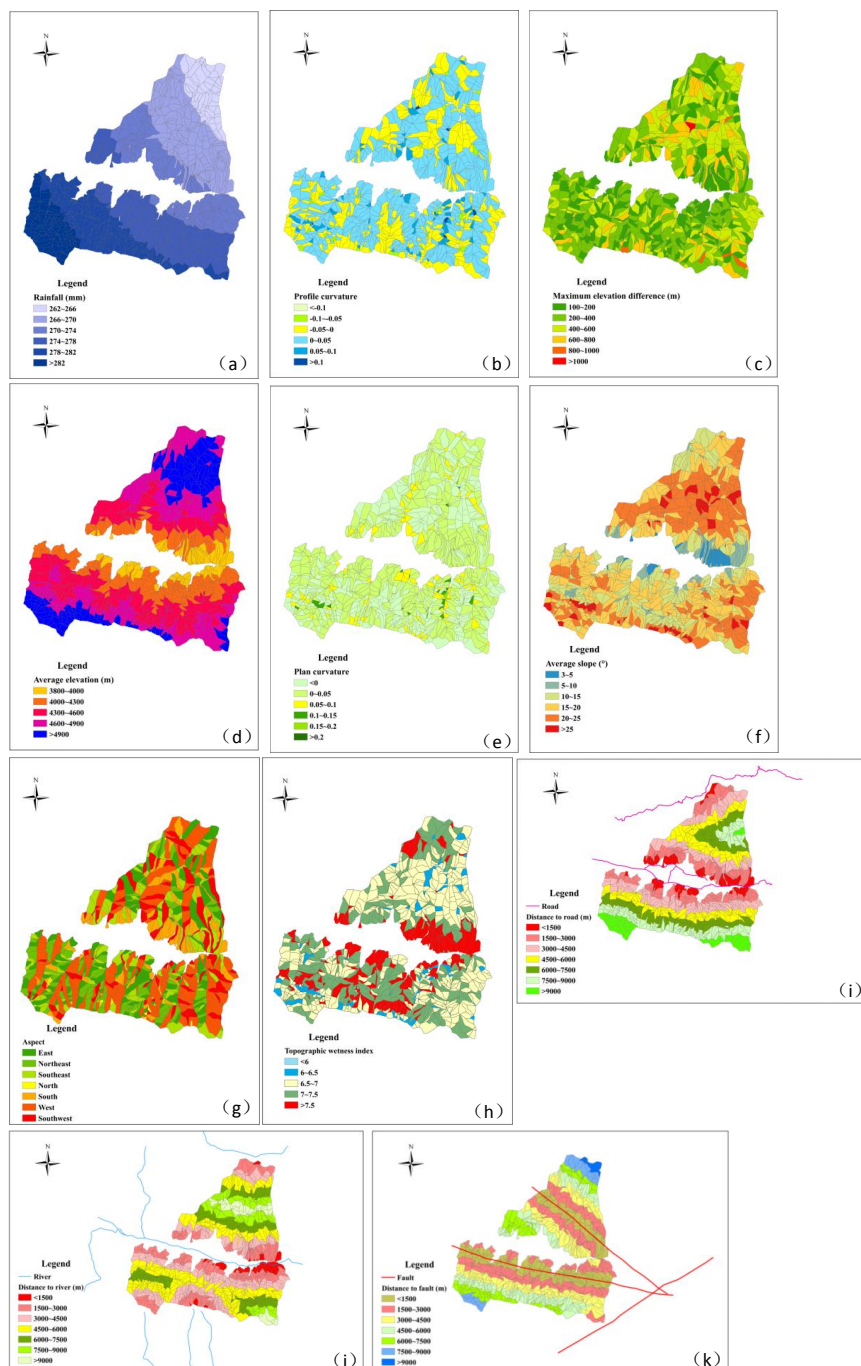
Fig.5. Study area thematic maps for landslide:（**a**）Rainfall;（**b**）Profile curvature;（**c**）Maximum

elevation difference;（**d**）Average elevation;（**e**）Plan curvature;（**f**）Average slope;（**g**）Aspect;

536　　　（**h**）Wetness;（**i**）Distance to road;（**j**）Distance to river;（**k**）Distance to fault.
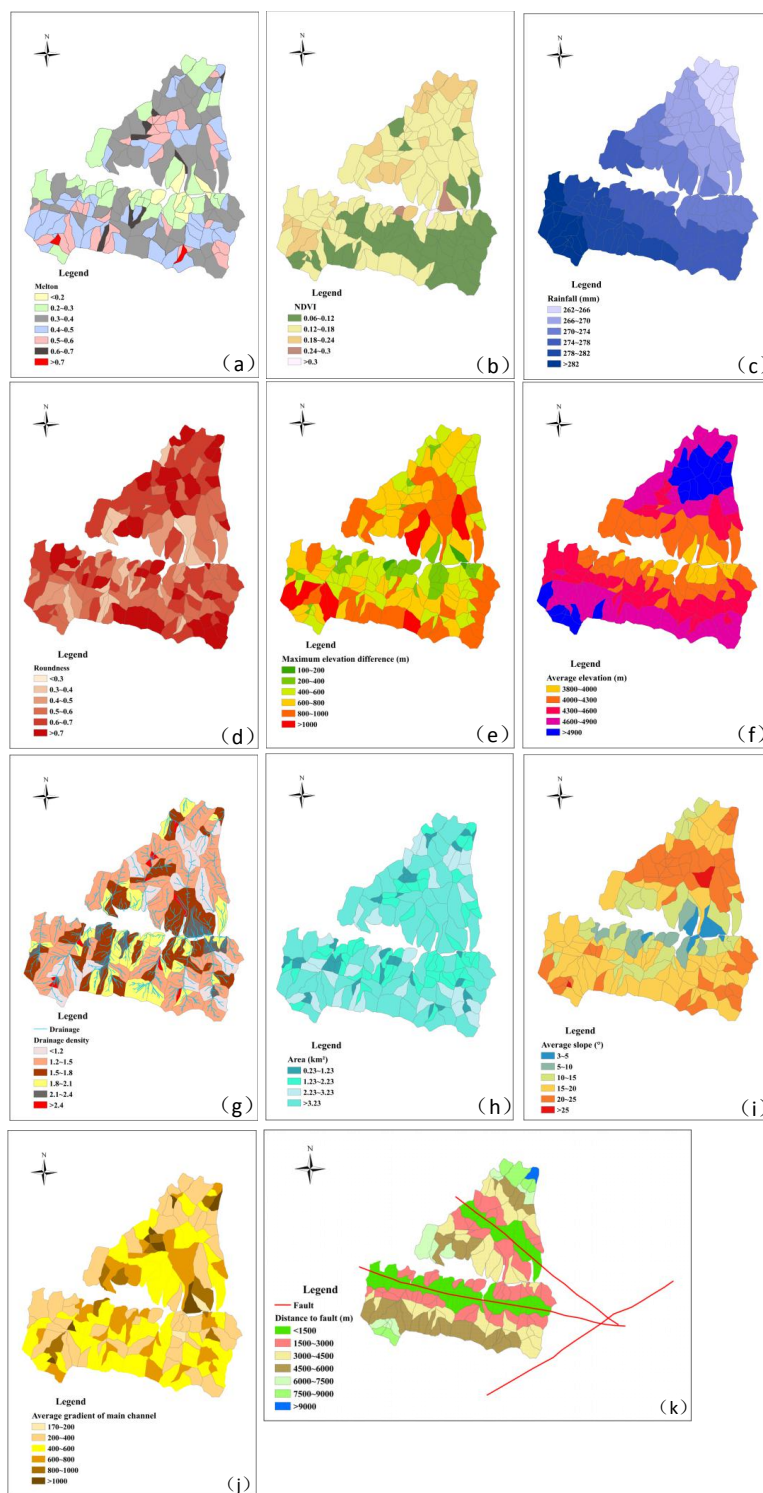
538 **Fig.6.** Study area thematic maps for debris flow: （**a**）Melton; （**b**）NDVI; （**c**）Rainfall; （**d**）Roundness;

539 （**e**）Maximum elevation difference; （**f**）Average elevation; （**g**）Drainage density; （**h**）Area; （**i**）

540 Average slope; （**j**）Average gradient of main channel; （**k**）Distance to fault.


541
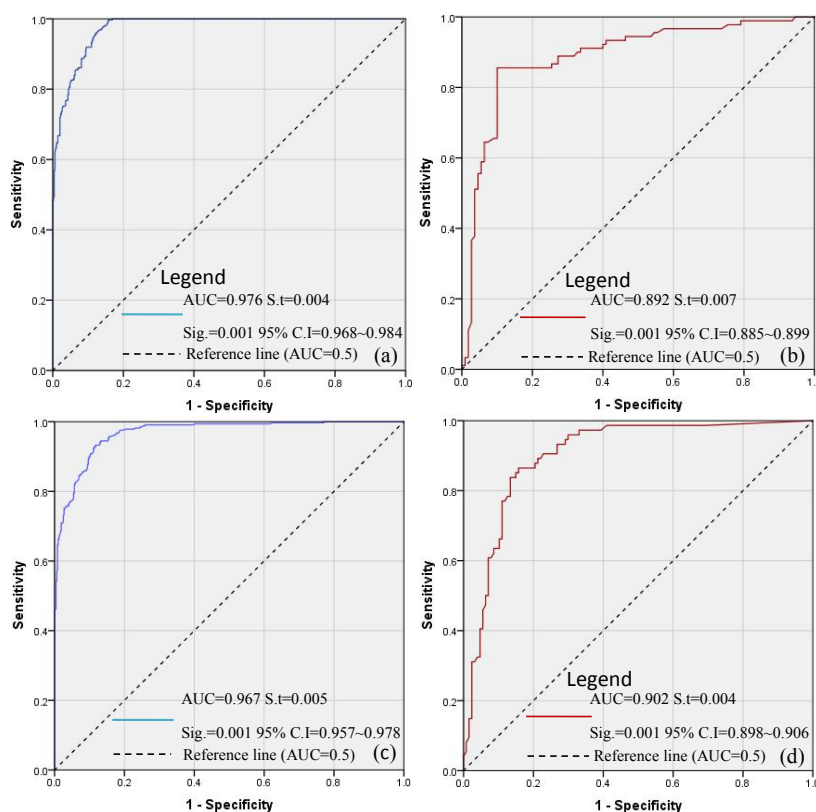
542 **Fig. 7.** Analysis of ROC curve for the two susceptibility maps: **(a)** Success rate curve of landslide using

543 the training dataset; **(b)** Prediction rate curve of landslide using the validation dataset; **(c)** Success rate

544 curve of debris flow using the training dataset; **(d)** Prediction rate curve of debris flow using the

545 validation dataset.
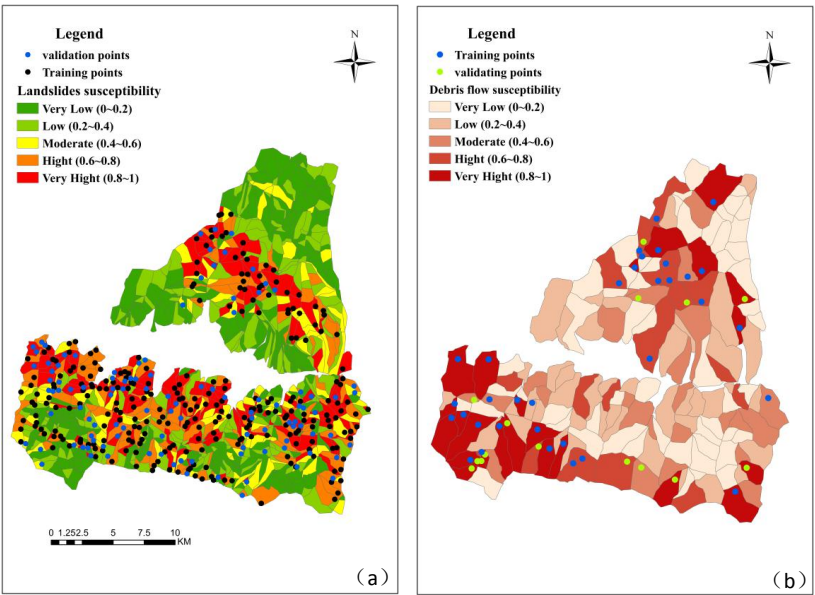
Fig.8. Susceptibility maps:（**a**）Landslide susceptibility zoning map;（**b**）Debris flow susceptibility

zoning map.
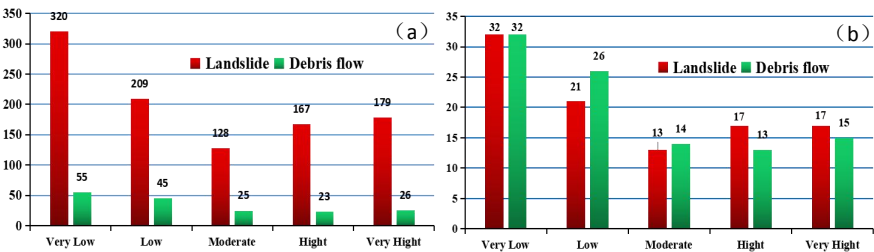
Fig. 9. Numbers and percentage of units in different susceptibility classes for landslide and debris flow:

(a) Numbers of units in different susceptibility classes for landslide and debris flow; **(b)** Percentages of

different susceptibility classes for landslide and debris flow.
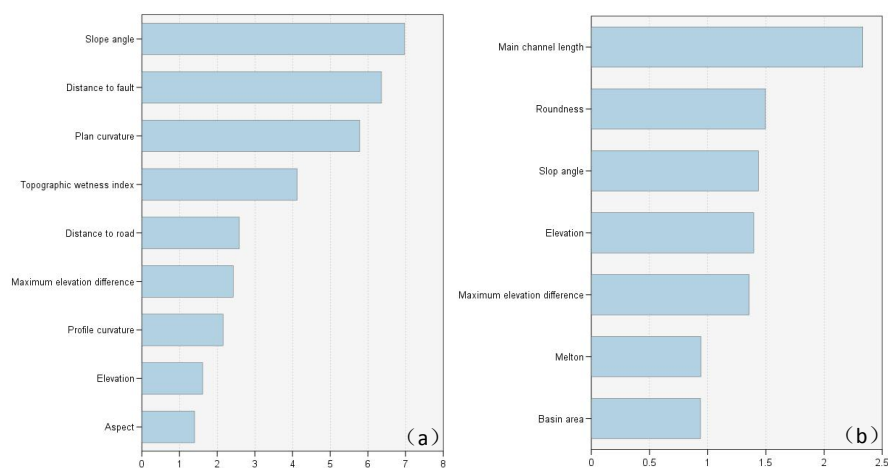
553

554 **Fig.10.** Parametric importance graphics obtained from RF model: **(a)** Parametric importance graphics

555 of landslide; **(b)** Parametric importance graphics of debris flow.

556

557   **Fig.11.** Landslide-debris flow susceptibility maps: **(a)** Height and very high-class watershed units with

558   high or very high slope units; **(b)** High or very high-class watershed units with low or very low slope

559   units; **(c)** High or very high-class slope units with high or very high-class watershed units; **(d)** Mapping

560   units.

561

562

563

564