Coseismic landslides triggered by the 8th August 2017 Ms 7.0 Jiuzhaigou earthquake (Sichuan, China): factors controlling their spatial distribution and implications for the seismogenic blind fault identification

Xuanmei Fana\*, Gianvito Scaringia, Qiang Xua\*, Weiwei Zhana,b, Lanxin Daia, Yusheng Lia, Xiangjun Peia, Qin Yanga, Runqiu Huanga

*a The State Key Laboratory of Geohazards Prevention and Geoenvironment Protection (SKLGP), Chengdu University of Technology, Chengdu, Sichuan, China, 610059*

b *Glenn Department of Civil Engineering, Clemson University, Clemson, SC, USA, 29634*

*\* Corresponding to Qiang Xu (xuqiang\_68@126.com) and Xuanmei Fan (fxm\_cdut@qq.com)*

*Present address: Chengdu University of Technology, Chengdu, Sichuan, China, 610059*

**Online Resource 1: Susceptibility assessment**

The analysis of the spatial distribution pattern of the coseismic landslides triggered by the Jiuzhaigou earthquake was done in GIS environment, through space statistical analysis functions. The raster layers of different factors were prepared and classified, the relative occurrence probability of the different classes for each factor were calculated. A binary logistic regression model was adopted to evaluate the geohazards susceptibility:

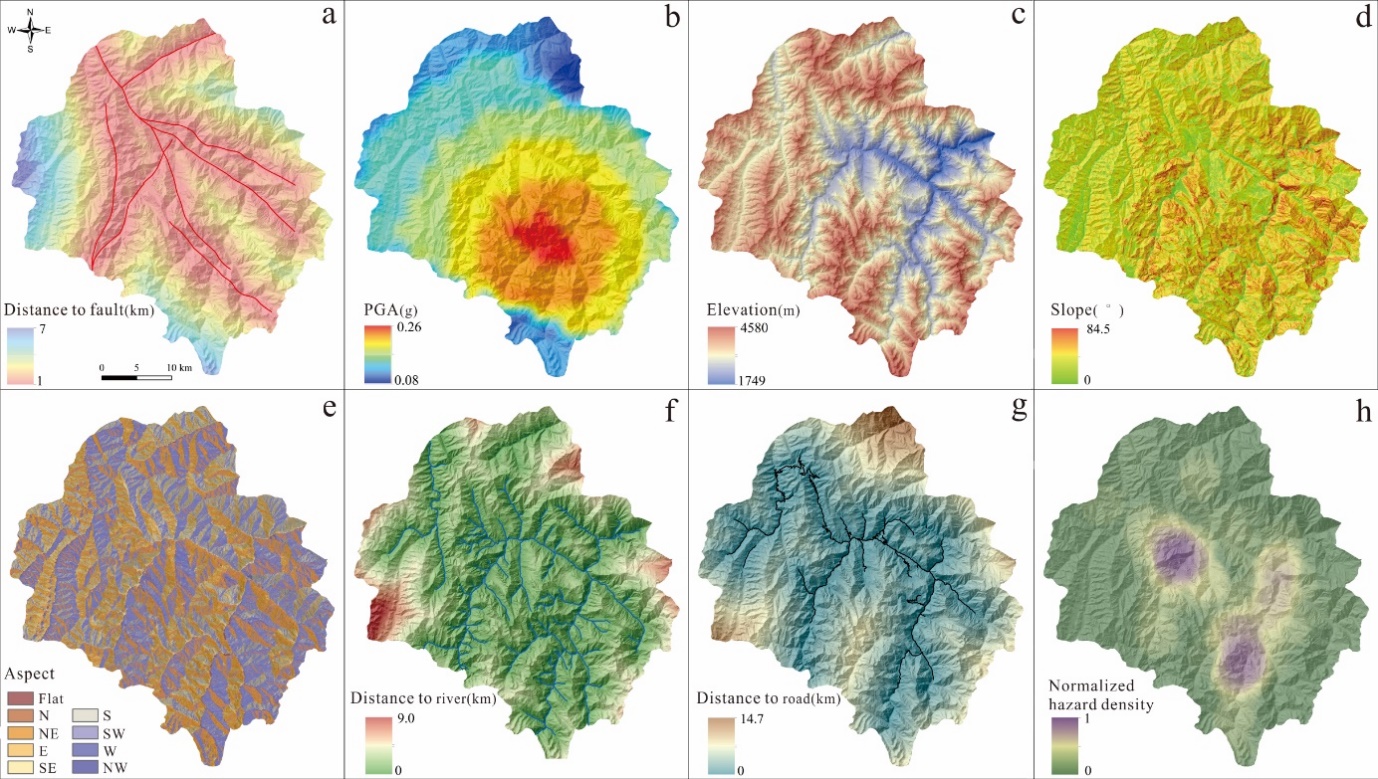
where *P* is the hazard occurrence probability, *α* is an empirical constant and the *βi* are the regression coefficients [calculate](http://dict.youdao.com/w/calculate/#keyfrom=E2Ctranslation)d by the logistic regression for each of the *xi* controlling factors. In the model, two possible values of the dependent variable *P* are possible, i.e. 1 or 0, representing whether the geohazard occurred or not.

Before the construction of the model, each factor had to be normalised, as they are expressed in different units and value ranges. The normalisation process was done through the following equations:

where, is the ratio of landslide area that belongs to class *j* of factor *i* to the total landslide area;is the ratio of area that belongs to class *j* of factor *i* to the total area in the study area; *i* is the number of factors (*i* = 1，2……，8)；*j* is the class number within each factor (*j* = 1，2，…，m) ; is the resulting index value; is the resulting normalised value of each class within each factor. The standardization results of each class of each factor are shown in Tab. S1 and in Fig. S1.

*Tab. S1 Classification and normalisation of the controlling factors.*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Factor | Class | Sij | Zij | Index value | Standard value |
| Slope angle (°) | <10 | 0.0405 | 0.0028 | 0.0700 | 0.0045 |
| 10-20 | 0.0996 | 0.0078 | 0.0781 | 0.0050 |
| 20-30 | 0.1845 | 0.0234 | 0.1266 | 0.0081 |
| 30-40 | 0.3753 | 0.1984 | 0.5285 | 0.0339 |
| 40-50 | 0.2297 | 0.4469 | 1.9458 | 0.1247 |
| 50-60 | 0.0586 | 0.2611 | 4.4578 | 0.2857 |
| 60-70 | 0.0109 | 0.0566 | 5.2046 | 0.3335 |
| >70 | 0.0010 | 0.0031 | 3.1924 | 0.2046 |
| Slope aspect | 0 | 0.0665 | 0.0557 | 0.8363 | 0.0951 |
| N | 0.1432 | 0.1701 | 1.1881 | 0.1351 |
| NE | 0.1471 | 0.2587 | 1.7581 | 0.1999 |
| E | 0.1069 | 0.1687 | 1.5772 | 0.1793 |
| SE | 0.1065 | 0.1229 | 1.1534 | 0.1312 |
| S | 0.1299 | 0.0630 | 0.4851 | 0.0552 |
| SW | 0.1423 | 0.0656 | 0.4606 | 0.0524 |
| W | 0.1059 | 0.0518 | 0.4890 | 0.0556 |
| NW | 0.0516 | 0.0437 | 0.8465 | 0.0963 |
| Elevation（m） | ＜2000 | 0.0028 | 0.0001 | 0.0335 | 0.0127 |
| 2000-2500 | 0.1185 | 0.1027 | 0.8666 | 0.3280 |
| 2500-3500 | 0.6141 | 0.8324 | 1.3555 | 0.5131 |
| 3500-4000 | 0.2212 | 0.0598 | 0.2704 | 0.1024 |
| ＞4000 | 0.0434 | 0.0050 | 0.1157 | 0.0438 |
| Lithology | T | 0.4145 | 0.1333 | 0.3215 | 0.0804 |
| P | 0.1708 | 0.1321 | 0.7736 | 0.1935 |
| N | 0.0025 | 0.0016 | 0.6508 | 0.1628 |
| C | 0.3726 | 0.7204 | 1.9336 | 0.4837 |
| D | 0.0396 | 0.0126 | 0.3177 | 0.0795 |
| PGA（g） | 0.12 | 0.0630 | 0.0054 | 0.0859 | 0.0194 |
| 0.16 | 0.2094 | 0.1332 | 0.6358 | 0.1436 |
| 0.20 | 0.3112 | 0.3771 | 1.2117 | 0.2736 |
| 0.24 | 0.3284 | 0.3619 | 1.1018 | 0.2488 |
| 0.26 | 0.0879 | 0.1225 | 1.3936 | 0.3147 |
| Distance to fault（m） | 0-1000 | 0.4667 | 0.5390 | 1.1549 | 0.3968 |
| 1000-2000 | 0.2575 | 0.3275 | 1.2719 | 0.4370 |
| ＞2000 | 0.2758 | 0.1335 | 0.4840 | 0.1663 |
| Distance to rivers（m） | 0-300 | 0.1697 | 0.1395 | 0.8221 | 0.2395 |
| 300-600 | 0.1500 | 0.2588 | 1.7258 | 0.5028 |
| ＞600 | 0.6803 | 0.6017 | 0.8844 | 0.2577 |



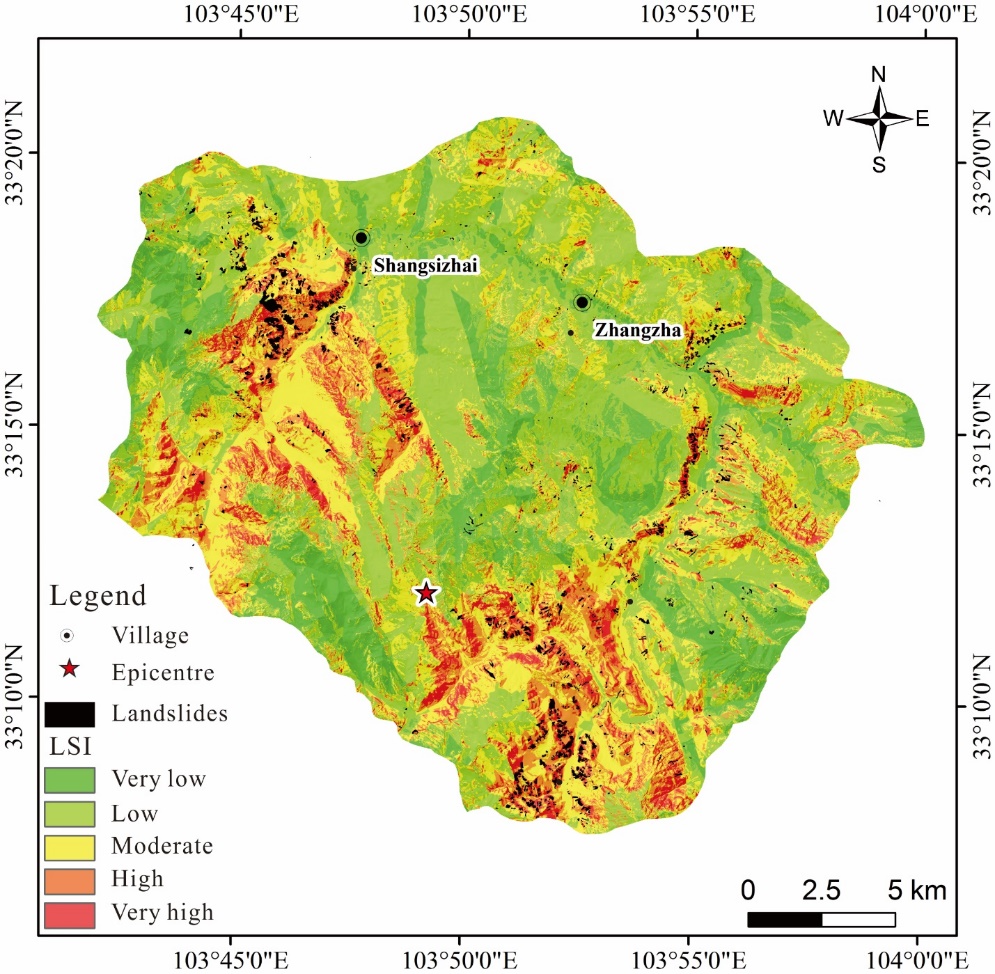
*Fig. S1 Factor maps: distance to possible seismogenic faults (a), PGA (b), elevation (c), slope angle (d), slope aspect (e), distance to rivers (f), distance to roads (g), normalized landslide areal density (h).*

A sample constituting seventy-five percent of the identified coseismic landslides was selected as the training set, while the remaining landslides were used as the validation set. The logistic model was developed through Equation (1), and regression coefficients and weights of each factor were obtained, as reported in Tab. S2.

*Tab. S2 Logistic regression and weight coefficients of the factors considered in the susceptibility assessment.*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Factor | Slope angle | Slope aspect / fault slip direction | Elevation | Lithology | Distance from fault | Distance from rivers | PGA | Slope curvature |
| Regression | 9.356 | 7.984 | 3.569 | 3.354 | 2.596 | 1.963 | 1.499 | -0.993 |
| Weight | 0.2990 | 0.2552 | 0.1141 | 0.1072 | 0.0830 | 0.0619 | 0.0479 | 0.0317 |

Then, the seismic landslide susceptibility map was obtained through the raster math application in ArcGIS, and was reclassified into low, relative low, moderate, relative high, and high susceptibility areas, as shown in Fig. S2. The model accuracy was analysed for both the training sample set and the validation sample set, as reported in Tab. S3.

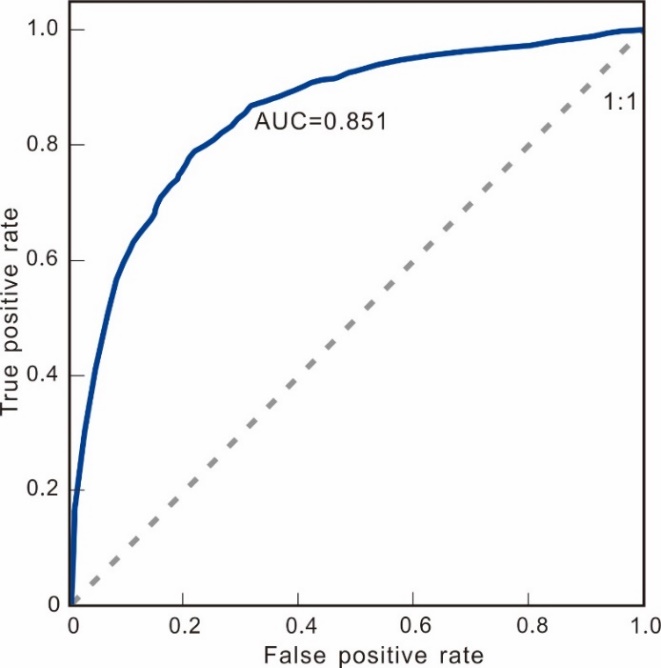


*Fig. S2 Coseismic landslide susceptibility map assuming the blind-fault F1 (see Fig. 2 and Fig. 5 in the main text) to be the seismogenic-fault.*

*Tab. S3 Area and percentage of landslides susceptibility zoning based on the logistic regression model.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Susceptibility | Area of each susceptibility zone /km2 | Area ratio of each zone to total area | Ratio of susceptibility zone to total landslide area in test sample set | Ratio of susceptibility zone to total landslide area in validation sample set |
| High | 19.35 | 4.49% | 42.91% | 44.48% |
| Relative high | 48.43 | 11.24% | 28.58% | 30.94% |
| Moderate | 110.27 | 25.59% | 20.81% | 17.59% |
| Relative low | 208.75 | 48.45% | 6.07% | 5.57% |
| Low | 44.06 | 10.23% | 1.58% | 1.19% |

The area with coseismic landslide susceptibility classified as moderate or higher accounts for 41.32% of the whole study area; 92.35% of the training sample set and 93.24% of the validation sample set fall in these classes, showing a good prediction ability of the model. The ROC (Receiver Operating Characteristic) curve is a widely used tool to estimate the accuracy of susceptibility assessments. The Area Under the Curve, AUC, is defined as the area between the ROC curve and the horizontal axis, and can range from 0.5 to 1. The AUC of the present model is 0.851, as shown in Fig. S3, highlighting a satisfactory prediction precision.



*Fig. S3 ROC curve of the Jiuzhaigou coseismic landslides susceptibility assessment.*