



Landslide Susceptibility Model Validation: A Routine Starting from Landslide Inventory to Susceptibility

Gulseren Dagdelenler, Hakan A. Nefeslioglu, and Candan Gokceoglu

Abstract

The main purpose of the present study is to evaluate the model validation stage of a routine landslide susceptibility mapping. For the purpose, model validation is assessed in three stages; (1) during model data production, (2) during model construction, and (3) during the production of model consequences; landslide susceptibility maps. As the results of these evaluations, it is revealed that training and testing data sets should be separated considering an appropriate separation ratio which is about 80 % and 20 % of the presence (1) data after completion of inventory studies. Correct classification percentages, error matrices, and the Kappa index are suggested to be considered for the training data sets during model construction. Additionally, again the correct classification percentage and the Root Mean Square Error (RMSE) should be considered during this stage for the testing data sets as well. In order to evaluate the spatial performance of the produced landslide susceptibility maps, the use of the Receiver Operating Characteristic (ROC) curves and the Area Under Curve (AUC) statistics is recommended. In the present study, the maximum Kappa index (k) value was calculated to be 0.459 for both the random sampling 1 (Rnd1) in the model 1 and for the random sampling 2 (Rnd2) in the model 2 during the model construction stage. The AUC values were calculated for these random samplings to be 0.781 and 0.790 respectively during the production of the model consequence stage in the study.

Keywords

Landslide susceptibility • Model • Testing • Training • Validation

Introduction

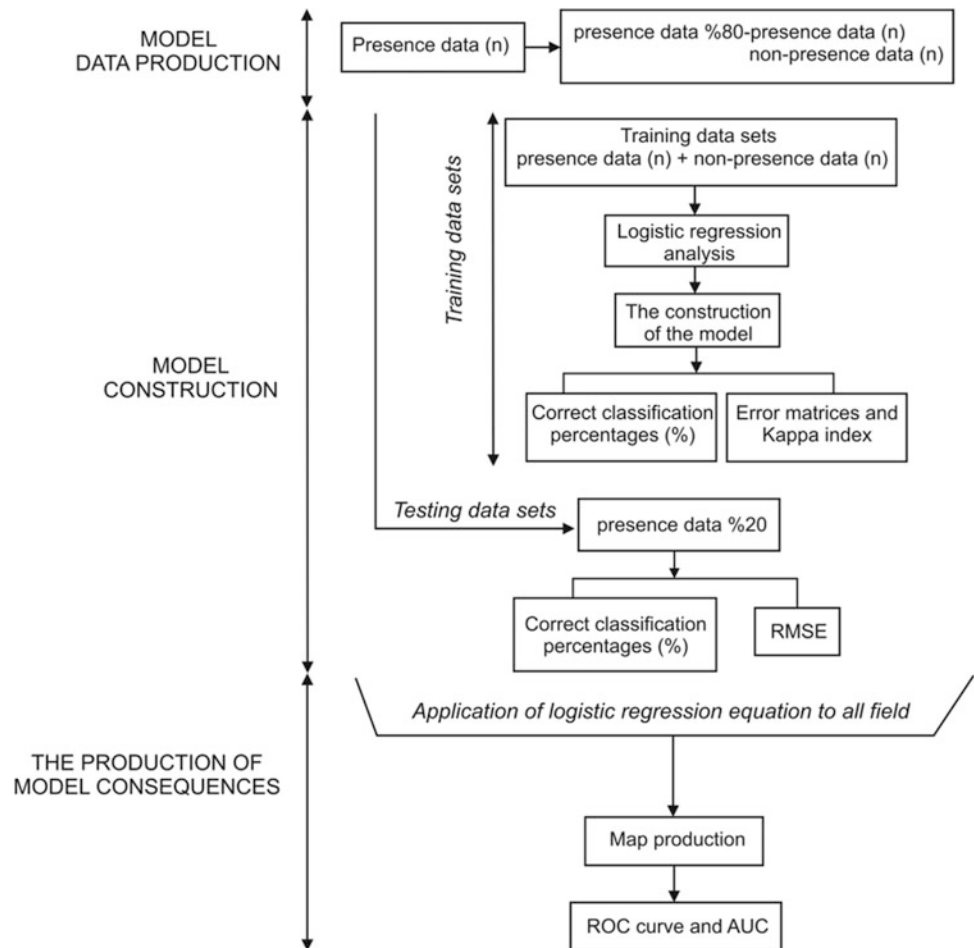
The model validation stage constitutes one of the most important issues in natural hazard modelling studies (Begueria 2006). Obviously, it could be considered that validation of landslide susceptibility models as well as the

maps should also be the most critical topic in landslide susceptibility researches. Although the model validation assessment has a very important role in modelling, many studies in literature do not give necessary importance to model validation, they just allow to be evaluated some basic validation statistics to assess the performance of the models. In landslide susceptibility mapping studies, the performance of the models are checked using different validation statistics by different researchers and that contributes to differences in terms of performance evaluation and the comparisons of the models according to these performances. The main purpose of this study is to represent a procedure for the assessment of the performance of landslide susceptibility mapping. The study proposes a flow chart which evaluates the current validation indices of the models in three stages:

G. Dagdelenler (✉) • C. Gokceoglu
Department of Geological Engineering, Hacettepe University, 06800
Beytepe, Ankara, Turkey
e-mail: gdagdelenler@gmail.com; candan.gokceoglu@gmail.com

H.A. Nefeslioglu
Polatli Technical Sciences Vocational School, Hacettepe University,
06900 Polatli, Ankara, Turkey
e-mail: haneleslioglu@hacettepe.edu.tr

Fig. 1 Procedure for the assessment of the performance of landslide susceptibility mapping



i) the model data production stage, ii) model construction stage and iii) the production of model consequences stage. For the purpose, the landslide susceptibility analyses performed by Dagdelenler (2013) in eastern part of Gallipoli Peninsula (Canakkale, Turkey) were evaluated.

Model Validation

As mentioned, the performance evaluations of the models were performed in three stages given in the flow chart in Fig. 1. These stages are described in detail below.

During Model Data Production

During the landslide susceptibility analyses performed by Dagdelenler (2013), a total of 10 variables (7 continuous and 3 categorical) were used as independent variables and also presence (1) and non-presence (0) data of mapped landslides was used as the dependent variable. 20 % of the presence (1) data was separated as testing and 80 % of the presence (1) data was separated as training data sets for the models

(Nefeslioglu et al. 2011, 2012; Oh and Pradhan 2011; San 2014). This process was carried out three times the models data sets Rnd1, Rnd2 and Rnd3 selected at random were obtained. For each random set, 80 % of the presence (1) data was separated and the training sets were generated by separating also 80 % of non-presence (0) data which equal to 80 % of presence (1) data in number. As it could be clearly realized, this stage constitutes the preparedness for further performance evaluations.

During Model Construction

The landslide susceptibility analyses were performed by applying the logistic regression technique by Dagdelenler (2013). As the results of the logistic regression analyses using training data sets, correct classification percentages (%), error matrices and validation statistics and Kappa index (k) values derived from the error matrices were determined. Correct classification percentages (%) calculated as the results of logistic regression analyses for the models were found to be acceptable and quite similar. They vary between 78 and 79 %.

Table 1 The presentation of true positive

		Observed	
		X ₁ (presence)	X ₀ (non-presence)
Predicted	X' ₁ (presence)	a	b
	X' ₀ (non-presence)	c	d

a; false positive, b; false negative, c; and true negatives, d in an error matrix

An error matrix shows the number of correctly estimated observations for positive and negative cases. In the error matrix in Table 1, the observed and predicted presence (1) and non-presence (0) data sets are represented by a letter (a, b, c, and d). Validation statistics (Table 2) and Kappa index (k) values derived from the error matrices of the models are calculated according to formulations of the validation statistics derived from the error matrix (Begueria 2006). The formula of the Kappa index (k) derived from error matrices was seen in (1,2, and 3). According to the Kappa index classification chart proposed by Landis and Koch (1977), Kappa index value for each model indicates that the model compatibility powers were moderately good.

$$P = \frac{a + d}{N} \quad (1)$$

$$Pe = \frac{(a + b)(a + c) + (c + d)(d + b)}{N^2} \quad (2)$$

$$k = \frac{P - Pe}{1 - Pe} \quad (3)$$

Where,

P = The proportion of observations in agreement; Pe = The proportion in agreement due to chance; k = Kappa index.

RMSE performance index and correct classification percentage (%) values were calculated by using the testing data

sets. Validation indices were derived from the error matrix and Kappa index values (threshold dependent) were specified by using the training data sets. Correct classification percentages for the first, second and third random sets of the landslide body sampling model were calculated as follows: 79.7 %, 80.5 % and 69.8 % respectively. The correct classification percentages for the landslide susceptibility models considered different buffer distances (d = 25 m, d = 50 m, d = 75 m and d = 100 m) in the seed cell samplings (Dagdelenler 2013) vary between 76.6 and 88.5 %. The seed cells obtained by the seed cell sampling strategy (Suzen and Doyuran 2004) are assumed to represent the pre-failure conditions of the landslides for the topographical parameters in particular. The calculated RMSE values for the Model 1 (landslide body samplings) are 0.398, 0.395 and 0.453 respectively and RMSE values vary between 0.334 and 0.422 for the Model 2 (seed cell samplings) at different buffer distances (d = 25 m, d = 50 m, d = 75 m, d = 100 m).

During the Production of Model Consequences

In the third stage of the performance evaluation procedure, the resultant landslide susceptibility maps were analysed by using the ROC curves and the area under ROC curves (AUC). The ROC curve evaluation and the AUC are threshold independent indices which are determined during the production of model consequences. Area under ROC curve value is used as a single threshold independent validation statistics (Begueria 2006). The An AUC value which is close to 1 means the performance of the model is good (Fawcett 2006). The ROC curves of the models were drawn and the AUC values were determined (Table 3). According to the results, the calculated AUC values for the models were found to be close to 1 and were very close to each other (Table 4). These results show that the performances of the models are quite acceptable.

Table 2 The formulation of the validation statistics derived from the error matrix (Begueria 2006)

Validation statistics	The formulation	Description
Efficiency	(a + d)/N	Ratio of correctly classified observations
Incorrect classification rate	(b + c)/N	Rate of correct classified observations
Odds ratio	(a + d)/(b + c)	Ratio of correct and incorrect classification of cases
Positive predictive value	a/(a + b)	Rate of true positives of the total positive estimation values
Negative predictive value	d/(c + d)	Rate of true negatives of the total negative estimation values
Sensitivity	a/(a + c)	Correctly estimated positive cases ratio
Specificity	d/(b + d)	Correctly estimated negative cases ratio
False positive rate	b/(b + d)	False positives ratio of total negatives observations
False negative rate	c/(a + c)	False negatives ratio of total negative observations
Likelihood ratio	Sensitivity/(1-specificity)	The ratio between true positives and false negatives

Table 3 Validation statistical values derived from error matrices and Kappa index (k) for both displacement + accumulation area (Model 1) and seed cell (Model 2) sampling model

Model	Efficiency value	Incorrect classification rate	Odds ratio	Positive predictive value	Negative predictive value	Sensitivity	Specificity	False-positive rate	False-negative rate	Likelihood ratio	Kappa index (k)
Model 1											
Rnd1	0.728	0.272	2.683	0.786	0.670	0.705	0.758	0.242	0.295	2.917	0.456
Rnd2	0.727	0.273	2.665	0.786	0.668	0.703	0.757	0.243	0.297	2.900	0.454
Rnd3	0.730	0.270	2.700	0.790	0.670	0.705	0.761	0.239	0.295	2.953	0.459
Model 2											
Rnd1											
d = 25 m	0.716	0.284	2.520	0.782	0.650	0.691	0.741	0.251	0.309	2.750	0.431
d = 50 m	0.721	0.279	2.580	0.783	0.658	0.696	0.752	0.248	0.304	2.811	0.441
d = 75 m	0.722	0.278	2.600	0.791	0.654	0.695	0.757	0.243	0.305	2.868	0.444
d = 100 m	0.720	0.280	2.578	0.783	0.658	0.696	0.752	0.248	0.304	2.804	0.440
Rnd2											
d = 25 m	0.730	0.270	2.703	0.799	0.661	0.702	0.767	0.233	0.298	3.010	0.459
d = 50 m	0.724	0.276	2.626	0.790	0.659	0.698	0.758	0.242	0.302	2.886	0.448
d = 75 m	0.717	0.283	2.538	0.782	0.653	0.692	0.750	0.250	0.308	2.766	0.434
d = 100 m	0.719	0.281	2.562	0.783	0.656	0.695	0.751	0.249	0.305	2.791	0.438
Rnd3											
d = 25 m	0.720	0.280	2.575	0.793	0.647	0.692	0.758	0.242	0.308	2.859	0.440
d = 50 m	0.720	0.280	2.575	0.785	0.656	0.695	0.753	0.247	0.305	2.811	0.440
d = 75 m	0.721	0.279	2.583	0.785	0.657	0.696	0.753	0.247	0.304	2.821	0.441
d = 100 m	0.718	0.282	2.551	0.781	0.656	0.694	0.749	0.251	0.306	2.771	0.436

Table 4 Area under ROC curve (AUC) values for the models

		Area under curve (AUC)		
		Rnd1	Rnd2	Rnd3
Model 1		0.801	0.802	0.781
Model 2	d = 25 m	0.785	0.790	0.789
	d = 50 m	0.790	0.792	0.792
	d = 75 m	0.787	0.788	0.789
	d = 100 m	0.783	0.785	0.786

Results and Conclusions

The validation indices are evaluated in three stages such as during model data production, model construction, and production of model consequences. In addition, a generalized flow chart for the performance evaluation of the landslide susceptibility models is proposed. According to the flow chart, it could be clearly realized that which validation indices are calculated from which data set and in which stage of the model. In recent huge landslide susceptibility literature, there is vagueness about the validation of the models constructed. This uncertainty starts from the model data production and goes up to production of the model consequences. The common way applied for validation is the evaluation of the ROC curves for whole study area (Ayalew and Yamagishi 2005; Mathew et al. 2007; Pradhan 2010). However, the performance evaluation of the model construction stage is commonly ignored particularly in the studies in which the bivariate statistics, artificial intelligence and data mining techniques are applied (Saito et al. 2009; Yilmaz 2009; Oh and Pradhan 2011; Akgun et al. 2012; Bui and Pradhan 2012; Conforti et al. 2014). Obviously, in order to apply this stage, a pre-processing stage including data production for further evaluations is necessary. The performance evaluation of the model construction stage was suggested to be a separate routine step in model validation for landslide susceptibility analyses in the proposed flow chart. It is commonly desired from a landslide susceptibility model to provide both high prediction capacity for the constructed model and high generalization capacity for the application results in whole study area (Can et al. 2005). Assuming that if all probabilities are calculated to be 1 for whole study area, in this case the spatial performance of the model is found to be 100 %. However, the resultant landslide susceptibility texture is irrational and it could be assumed that there is no generalization capacity for the model for this situation. The probabilities will also be calculated to be 1 for whole areas without landslides so that the calculated AUC values will also be low. In other words, the model prediction capacity is maximum while the generalization capacity is minimum. Hence, it could be clearly realized that the validation indices for the model construction and production

model consequences should be evaluated separately. Therefore, it could be concluded that the proposed methodology in this study enables controlling both prediction and generalization capacities for any landslide susceptibility evaluations appropriately.

According to the results obtained from the sample analyses of the models given in this study, the validation indices are quite close and this also shows that model predictive and generalization capacities could be evaluated to be acceptable.

References

- Akgun A, Sezer EA, Nefeslioglu HA, Gokceoglu C, Pradhan B (2012) An easy to use MATLAB program (MamLand) for the assessment of landslide susceptibility using a Mamdani fuzzy algorithm. *Comput Geosci* 38:23–34
- Ayalew L, Yamagishi H (2005) The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda-Yahiko Mountains, Central Japan. *Geomorphology* 65:15–31
- Beguieria S (2006) Validation and evaluation of predictive models in hazard assessment and risk management. *Nat Hazards* 37:315–329
- Bui D, Pradhan B (2012) Landslide susceptibility mapping at Hoa Binh province (Vietnam) using an adaptive neuro-fuzzy inference system and GIS. *Comput Geosci* 45:199–211
- Can T, Nefeslioglu HA, Gokceoglu C, Sonmez H, Duman TY (2005) Susceptibility assessment of shallow earthflows triggered by heavy rainfall at three subcatchments by logistic regression analyses. *Geomorphology* 72:250–271
- Conforti M, Pascale S, Robustelli G, Sdao F (2014) Evaluation of prediction capability of the artificial neural networks for mapping landslide susceptibility in the Turbolo River catchment (northern Calabria, Italy). *Catena* 113:236–250
- Dagdelenler G (2013) Assessment of sampling and validation strategies in production of landslide susceptibility Maps (Eastern Part of Gallipoli Peninsula). Hacettepe University, Institute of Science, Ankara, Turkey, 186 p. (unpublished)
- Fawcett T (2006) An introduction to ROC analysis. *Pattern Recogn Lett* 27(8):861–874
- Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33:159–174
- Mathew J, Jha VK, Rawat GS (2007) Application of binary logistic regression analysis and its validation for landslide susceptibility mapping in part of Garhwal Himalaya, India. *Int J Remote Sens* 28(10):2257–2275
- Nefeslioglu HA, Gokceoglu C, Sonmez H, Gorum T (2011) Medium-scale hazard mapping for shallow landslide initiation: the Buyukkoy catchment area (Cayeli, Rize, Turkey). *Landslides* 8:459–483

- Nefeslioglu HA, San BT, Gokceoglu C, Duman TY (2012) An assessment on the use of Terra ASTER L3A data in landslide susceptibility mapping. *Int J Appl Earth Observ Geoinform* 14:40–60
- Oh HJ, Pradhan B (2011) Application of a neura-fuzzy model to landslide-susceptibility mapping for shallow landslides in a tropical hilly area. *Comput Geosci* 37:1264–1276
- Pradhan B (2010) Remote sensing and GIS-based landslide hazard analysis and cross-validation using multivariate logistic regression model on three test areas in Malaysia. *Adv Space Res* 45:1244–1256
- Saito H, Nakayama D, Matsuyama H (2009) Comparison of landslide susceptibility based on a decision-tree model and actual landslide occurrence: The Akaishi Mountains, Japan. *Geomorphology* 109:108–121
- San BT (2014) An evaluation of SVM using polygon-based random sampling in landslide susceptibility mapping: The Candir catchment area (western Antalya, Turkey). *Int J Appl Earth Observ Geoinform* 26:399–412
- Suzen ML, Doyuran V (2004) Data driven bivariate landslide susceptibility assessment using geographical information systems: a method and application to Asarsuyu catchment, Turkey. *Eng Geol* 71:303–321
- Yilmaz I (2009) Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: A case study from Kat landslides (Tokat-Turkey). *Comput Geosci* 35:1125–1138