Evaluation of Gully Erosion Susceptibility Using Logistic Regression, in Salavat Abad, Kordestan Province

Reza Esmaili*
Department of geography, University of Mazandaran, Iran.
Ronak Shokati
M.Sc of Geomorphology, University of Mazandaran, Iran.

Extended Abstract

Introduction

Gully is one of the forms of water erosion features in many regions of the world. Gully erosion could produce abundant sediment load, reducing the fertility of land and destruction of structures. In literature, for the preparation of landslide susceptibility mapping, several studies have been conducted using logistic regression. But, a few authors focused on gully erosion susceptibility. Susceptibility is defined as the probability of spatial occurrence of a phenomenon on the basis of the relationships between the distribution of the events in the past and a set of predisposing factors. Therefore, statistical methods for quantitative assessment of the place of happening phenomenon are appropriate in future. The studied area is located in Salvatabad of Kordestan province. The regional geology is formed by sedimentary rocks such as marl, limestone, conglomerate and alluvial deposits. The average annual rainfall is about 300 mm and the average annual temperature is 11.3 °C and the climate is cold and dry.

Research Methodology

Gully erosion susceptibility analysis is composed of four stages: (1) gathering data and create a spatial database from relevant factors, (2) assessment of the GES using the relationships between the gully and gully factors affecting their occurrence using logistic regression, (3) construction of GES map with GIS, and (4) validation of the results.

Field surveys and satellite image interpretation allowed identifying gullies. Lithology, land use and land cover, slope, aspect, plan curvature, length-slope factor, stream power index and topographical wetness index were assumed as gully erosion independent variables. The gully areas were converted to points and were split into two groups: training and validation data using a random partition. The occurrence of gullies in relation to the terrain conditions was analyzed using the training data and the final model was applied to the entire region. The validation data was also used to evaluate the prediction of the model. 70% of gully points are selected as training set and the remnant 30% as validation data. From 98000 point, the number of 68600 (70%) for the implementation of model was chosen and the remaining points (29400) was then used as the validation data.

Logistic regression (LR) is a statistical technique that involves one or more independent variables to predict the probability of a binary or categorical response. Logistic regression estimates the probability (P) of the occurrence of a particular event. In this paper, the dependent variable of LR was used as categorical variables and gully data as binary encoded (i.e., the presence or absence of gully). "Tolerance" (TOL) and the ''variance inflation factor'' (VIF) as two indexes were used to detect multicollinearity diagnosis. Then, GES map prepared using the LR method, and the accuracy of model was evaluated by calculating receiver operating characteristics (ROC).

Discussion and Results

The value of TOL and VIF indicated that there was no multicollinearity between independent factors. The model-building process started with 8 effective factors on 17th classes. Then the step regressive analysis in LR was done which showed that 15 classes were statistically significant. The classes of variables with

* Email: r.esmaili@umz.ac.ir

estimated coefficients that have a significance value (Sig.) of less than 0.05 were found. So these variables were accepted as predictor variables. GES map were produced from LR model. The results for gully occurrence showed that marl lithology and poor vegetation cover had the highest gully occurrence. To validate the model, GIS-based map and ROC curves were overlapping. Performance of the ROC curves showed percentage of setting gullies places in the most susceptible areas. Overlay maps in GIS environment were reclassified. The susceptibility values were divided into five categories. Results of validation, based on ROC (receiving operating characteristic) curves, showed that there were satisfactory comparison between the susceptibility map and the location of the existing gullies; therefore, AUC =79% indicated that 79% of the gullies are correctly classified in high and very high susceptibility areas. The GES map obtained from the LR method comprised of the very low, low, moderate, high and very high zones which were 37.7%, 29.5%, 9.9%, 15.4% and 7.5% of the total area, respectively.

Conclusion

In this research, variable lithology, land use, topography (slope, aspect, slope shape, LSF, SPI and TWI) were analyzed with logistic regression and GES map are prepared. Results of validation, based on ROC (receiving operating characteristic) curves confirmed the good accuracy in this area. Also, researches of Luca et al (2011) and Conoscenti et al (2013, 2014) confirmed the usage of LR model in GES.

Keyword: Gully erosion, logistic regression, Salavatabad, Kordestan.