Amirkabir Journal of Civil Engineering

Amirkabir J. Civil Eng., 53(5) (2021) 403-406
DOI: 10.22060/ceej.2020.17119.6465

Developing a Decision Tree based on Data Mining Method for Detecting the Influential Parameters on the Power of Flood Destruction

H. Farhadi, A. Esmaeily, M. Najafzadeh*

Department of Surveying Engineering, Graduate University of Advanced Technology, Kerman, Iran.

ABSTRACT: Floods, as one of the natural disasters, cause irreparable damages to the urban infrastructures, agricultural lands, and natural resources. Therefore, access to comprehensive information on influential factors the extent of flood damage can be useful in estimating the extent of the damage. In this way, this study investigates the creation of a database of effective parameters on flood destruction power using a case study of Landsat-7 satellite images with ETM+ sensor and ASTER DEM data using a decision tree. In the current research, environmental parameters such as canopy, natural slope, and slope direction were considered to evaluate flood degradation power in the study area and the decision tree model was created using these criteria. Ultimately, based on these parameters, the number of changed pixels (after the flood) in the study area is 692361 which indicates 62312.49 hectares of degraded land in the study area. According to the findings of the present study, lands with low canopy characteristics, namely normalized differential vegetation index (NDVI) between 0.2 and 0.4, low slope 0 to 45 degrees, and Southern slope direction caused the most damage caused by floods. Also, areas with dense NDVI, high slope, and northern slope orientation have a preventative influence on floods-caused damages. Overall, it can be found that the decision tree, as a data mining method, is capable of yielding better accuracy and quality in determining the effective parameters in estimating flood destruction power by increasing the input variables.

Review History:

Received: Sep. 24, 2019
Revised: Dec. 17, 2019
Accepted: Dec. 17, 2019
Available Online: Jan. 29, 2020

Keywords:

Flood, Decision Tree
Data Mining
Change Detection
Landsat-7

1- Introduction

Floods, causing great damages, have attracted meticulous attention from civil engineers. The rate of flood degradation is contingent upon the time and season of occurrence and consequently causing heavy damage to urban construction, agricultural land, and natural resources. Thus, obtaining pieces of information about the flood and the characterization of destruction power in urban districts can play a key role in estimating the floods-caused damages. For this purpose, the use of data and Remote Sensing techniques is suggested as a key solution. Recently, there have been many activities in flood monitoring using Remote Sensing. In one article, the potential and limitations of open satellite data for flood mapping are presented, in which the potential and limitation of the data are examined in terms of time and climate [1]. Due to the limitations of weather conditions during floods such as cloud cover, employing radar is considered an efficient tool [2]. Agricultural lands receive the greatest impact from floods; therefore, a study has been carried out to investigate the relationship between flood severity and land use in the Kerman basin [3]. The main purpose of this research is to apply using a decision tree to create a database of influential parameters on flood destruction power using a case study of Landsat-7 satellite with ETM+ sensor and ASTER DEM[1] data.

2- Data and Study Area

The data used in this study are ASTER DEM and Landsat-7 satellite imagery with ETM+ Sensor for cities in northwestern Iran.

3- Methodology

Input data includes Landsat-7 satellite images before and after the flood and Digital Elevation Model (DEM) used to detect changes occurring after the flood.

3.1. Pre-Processing Of Aster Dem

Pre-processing of ASTER DEM in the current research includes two main steps: 1-mosaic of DEM scenes and 2-mask Urmia Lake from DEM.

3.2. Pre-Processing Of Landsat-7 Imagery

Satellite imagery often has geometrical and radiometric errors that are needed to correct these errors for accurate and better estimation of the area. Geometrical errors of the image include the displacement error caused by rotation and the change of position of the sensor at the moment of image

*Corresponding author's email: moha.najafzadeh@gmail.com

1 Digital Elevation Model

Table 1. Parameters values in the decision tree

Slop	Low	0 < Slop < 50
	High	50 < Slop < 100
Canopy (*NDVI*)	Low	0.2 < canopy > 0.4
	Medium	0.41 < canopy > 0.6
	High	0.61< canopy > 0.99
Aspect Slop (degree)	Northern	340 º < Aspect < 20 º
	Eastern	70 º < Aspect < 110 º
	Southern	160 º < Aspect < 200 º
	Western	250 º < Aspect < 290 º

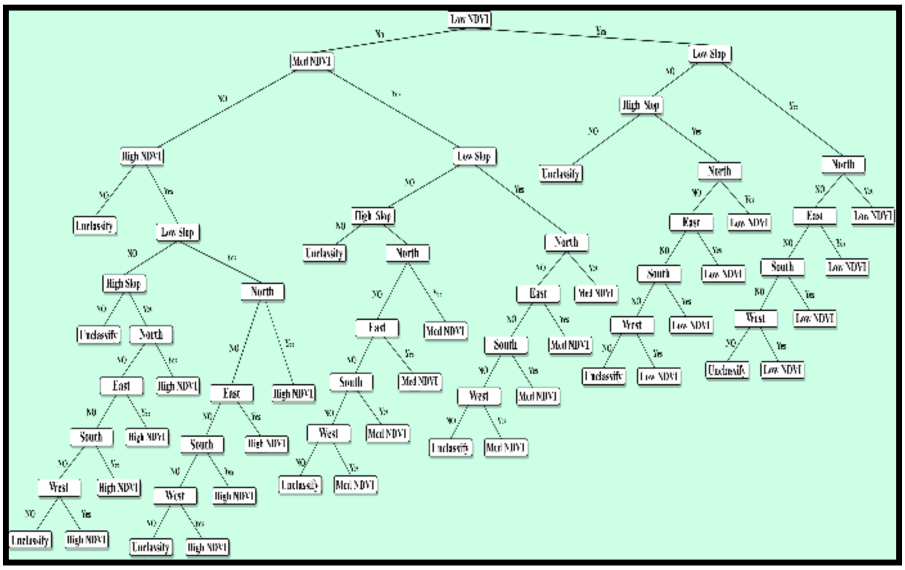

Fig. 1. The overall structure of the decision tree

acquisition, which must be eliminated. Radiometric image errors are generally the errors that occur on every single pixel, and these errors can be both self-sensor and atmospheric at the moment of image acquisition. In the present research, to eliminate this error, the empty pixels were replaced with the mean values of the neighboring pixels. FLAASH[1] method, which is a radiometric correction, is applied for atmospheric correction.

3.3. Generating Canopy Map From Landsat-7

Normalized difference Vegetation Index (*NDVI*) produced using Landsat-7 images, is one of the parameters considered to determine flood damage. *NDVI* is calculated from Eq. (1):

$$NDVI = \frac{NIR - R}{NIR + R} \qquad (1)$$

In Eq. (1), *NDVI* is the normalized differential vegetation index, *NIR* is the near-infrared reflection band, and *R* is the red band reflectance of the Landsat-7 satellite.

1 Fast Line- of- sight Atmospheric Analysis of Spectral Hyper cubes

3.4. Generating Slope And Aspect Map

Other parameters considered in determining the flood destruction potential in the current research are slope and Aspect map.

3.5. Generating Slope And Aspect Map

A decision tree is a multi-step classification that can be used for an image or a set of images. The decision can be served for any attribute in the dataset. The unique capability of this tool in the classification of features is that any data such as 2D and 3D data can be evaluated. This method is used for data mining and detection of changes and finally estimation of flood Destruction. In the present study, environmental parameters such as vegetation, natural slope, and slope direction were considered to evaluate flood degradation power in the understudied area. The code for this tool is written in the ENVI software IDL environment. The decision tree model was developed using the criteria listed in “Table 1” as shown in “Fig. 1”.

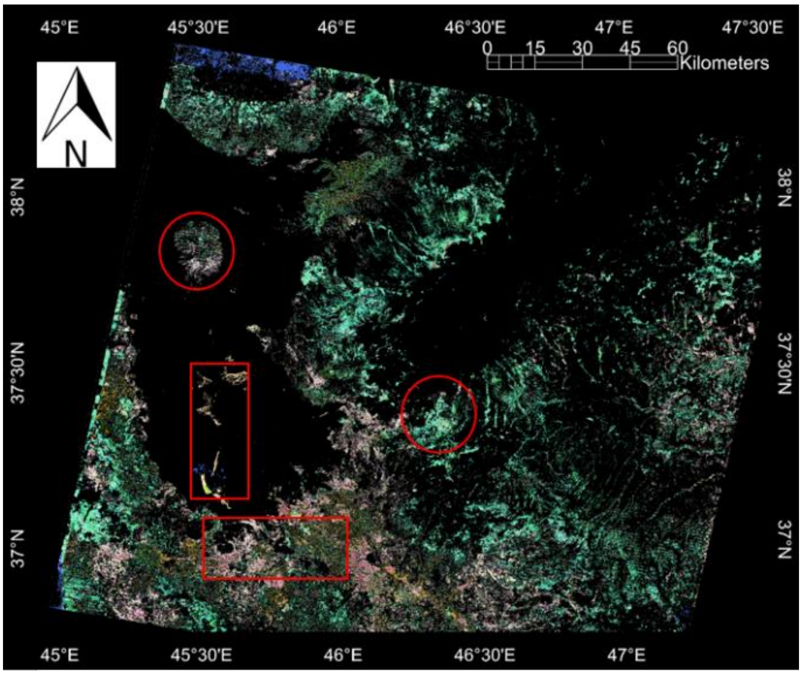

Fig. 2. The results of the occurred changes

Table 2. Statistical characteristics of the amount of flood destruction power according to classification parameters

Canopy	Slop	Aspect slop	Number of pixels changed	Changed percentage %	Area changed
Low	Low	Western	173348	0.4	15601.32 hectare
Low	Low	Eastern	168430	0.36	15158.7 hectare
Medium	Low	Southern	21677	0.05	1950.93 hectare
Medium	Low	Western	13256	0.028	1193.04 hectare
Medium	Low	Eastern	12742	0.027	1146.78 hectare
Low	High	Southern	2432	0.005	218.8 hectare
Low	High	Eastern	380	0.001	34.2 hectare
Low	High	Western	304	0.001	27.36 hectare
Medium	High	Southern	9	0.0001	8100 m^2
Medium	High	Eastern	5	0.0001	4500 m^2
Medium	High	Western	3	0.0001	2700 m^2
Sum			**692361**	**% 1.58**	**62312.49 hectare**

4- Results and Discussion

The result of the changes is shown in “Fig. 2”. The amount of flood destruction power according to the vegetation, slope, and Aspect slop classification parameters is in “Table 2”. In accordance with “Table 2”, areas with low vegetation, low slope, and south, west, and eastern criteria were 0.7%, 0.4%, and 0.36%, respectively, most impacted by floods.

5- Conclusion

In the current investigation, lands with low canopy characteristics, namely normalized differential vegetation index (*NDVI*) between 0.2 and 0.4, low slope 0 to 45 degrees, and Southern slope direction caused the most damage caused by floods. Furthermore, lands with dense *NDVI*, high slope, and northern slope orientation have a preventative influence on floods-caused damages. Finally, it can be found that the decision tree yielded satisfying performance and quality in determining the influential parameters in predicting the destruction power of flood by increasing the input variables.

Based on these parameters, the number of changed pixels (after the flood) in the study area is 692361 which indicates 62312.49 hectares of degraded land in the study area.

Obtaining a large number of parameters requires an efficient hardware system. Based on the findings of this study, it can be concluded that decision tree–based-database provided useful information about natural disasters such as floods and how to deal with them.

References

[1] D. Notti, D. Giordan, F. Caló, A. Pepe, F. Zucca, J. Galve, Potential and limitations of open satellite data for flood mapping, Remote Sensing, 10(11) (2018) 1673.

[2] G. Schumann, Preface: Remote sensing in flood monitoring and management, Multidisciplinary Digital Publishing Institute, 2015.

[3] R. Kazemi, J. Porhemmat, Investigating the relationship between flood intensity and land use in Kerman province basins, in: 13th National Conference on Watershed Management Science and Engineering and 3rd National Conference on Natural and Environmental Conservation, Focusing on Watershed and Protecting Natural Resources and the Environment, Ardebil, 2018.

HOW TO CITE THIS ARTICLE

H. Farhadi, A. Esmaeily, M. Najafzadeh, Developing a Decision Tree based on Data Mining Method for Detecting the Influential Parameters on the Power of Flood Destruction. Amirkabir J. Civil Eng., 53(5) (2021) 403-406

DOI: 10.22060/ceej.2020.17119.6465