

## DEVELOPING A CELLULAR AUTOMATA MODEL FOR SIMULATING RAINFALL-RUNOFF PROCESS (CASE STUDY: BABOL CATCHMENT)

N. Mahdizadeh Gharakhanlou<sup>1</sup>\*, M. S. Mesgari<sup>1</sup>

<sup>1</sup> Geospatial Information Science Division, Faculty of Geodesy and Geomatics Engineering, and Center of Excellence in Geo-Information Technology, K.N. Toosi University of Technology, Iran – (navid\_mehdizade@email., mesgari@kntu.ac.ir)

**KEY WORDS:** Precipitation, Runoff, Curve number (CN), Antecedent moisture condition (AMC), Hydrologic soil group (HSG), Cellular automata (CA), Geospatial information system (GIS).

### ABSTRACT:

The rainfall-runoff process is one of the central aspects of hydrology. In the rainfall-runoff process, the fast accumulation of rainwater in the catchments and rivers may lead to a flood disaster. To prevent and manage such disasters, it is essential to model and simulate the runoff process. The cellular automata (CA) is a computational model that can represent and simulate complex and dynamic natural phenomena. The main objective of this research is to develop a CA model of the rainfall-runoff process and to use it for simulating and representation of the runoff in the Babol catchment. The study is carried out in two main steps. First, several data manipulation and analyses in ArcGIS 10.3 software are used to prepare the input data of the CA model. In the second step, the CA model is developed and implemented in the NetLogo 6.0.4 platform. The total amount of runoff in each cell during the year 2016 was calculated and presented as the output of the model. A map of the total generated runoff can be an important tool for flood management decision-makers and it can also be used for managing the catchment and assessing the water resources in the study area.

### 1. INTRODUCTION

Understanding the processes of accumulation and transfer of flow to the outlet of the catchments is one of the most important issues in hydrology science. There are several reasons to simulate the rainfall-runoff process in hydrologic models. One of the main advantages of simulation models over the measurement methods in hydrology is the difficulties of measuring the hydrological parameters (Beven 2001). Besides, usually, the purpose of the simulating the rainfall-runoff process is to study the causes of the flood events and to identify the places where the flood is likely to occur.

A hydrological model is a simple representation of the real system that helps to study the function of the catchment in response to various inputs and helps to better understand the hydrological processes (Spruill et al. 2000). The rainfall-runoff process is one of the complicated hydrological processes. Considering the importance of proper runoff estimation in catchment management and water structures design, various methods have been proposed for rainfall-runoff modeling. One of the most recent methods, in this regard, is the cellular automata (CA) model.

In the CA model, the catchment landscape is represented by a network of cells. Then, the interactions between cells are defined according to the rules governing the physiography of the catchment. To have a realistic CA model of the rainfall-runoff process, all real-world data such as terrain and hydrologic characteristics of the catchment should be used. Therefore, it is necessary to collect the relevant spatial data and to use the analytical capabilities of the Geospatial Information System (GIS) to prepare the input data for the CA model.

In general, the CA model could provide a feasible alternative approach to the widely applied physical-based models. The main advantage of the CA model is its ability in simulating a complex

geographic phenomenon using simple local interaction rules. The CA model has been used by various researchers in various scientific and engineering applications. It was first used in hydrology by Murray and Paola (1994) to model the braided rivers. Afterward, many researchers used the CA for hydrology and runoff simulation. Wu et al. (2006) simulated the runoff of the LeAnhe watershed in the northeast of JiangXi Province in China using CA. Parsons and Fonstad (2007) presented a CA model for simulating surface runoff in the Walnut Gulch Experimental Watershed (WGEW) in southeastern Arizona, U.S.A. Rinaldi et al. (2007) developed a CA simulation of the runoff of the Santa Catalina Creek Basin in the center of Buenos Aires province in Argentina. Shao et al. (2015) developed the Runoff Model Based on Cellular Automata (RunCA) to simulate surface runoff at different scales by integrating basic CA rules with fundamental measurable hydraulic properties. As mentioned, there are several studies on the use of CA to simulate the runoff, which all evaluated CA as a suitable method for modeling complex hydrologic processes such as runoff.

In this research, the rainfall-runoff of the Babol catchment is simulated using a CA model. Then, the model is used to determine the amounts of runoff in different parts of the study area. One of the advantages of a rainfall-runoff model is the ability to identify the areas where the runoff occurs.

### 2. MATERIALS AND METHODS

#### 2.1. Study area

The study area is the Babol catchment located in the Mazandaran province of Iran. The area is in the south of the Caspian Sea and covers about 6042 square kilometers (Figure 1).

\* Corresponding author.

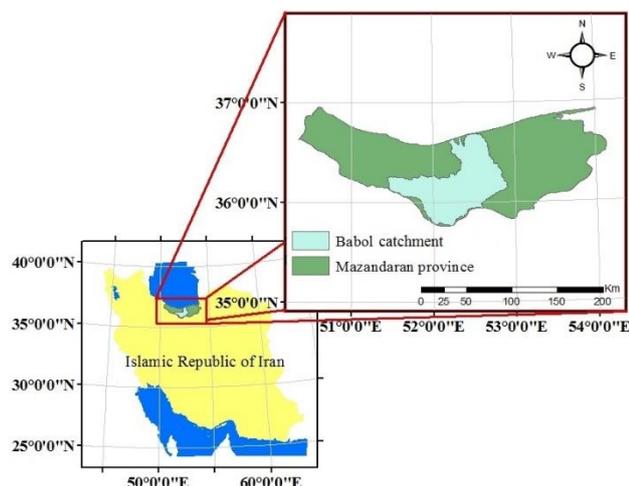


Figure 1. Location of Babol catchment in the Mazandaran province.

## 2.2. The data set of the study

The rainfall-runoff process of any catchment is mainly influenced by its hydrological conditions, geomorphology, topography, and climate condition of the region. The amount and distribution of precipitation, land use, topographic conditions, and soil characteristics of the region are among the main factors affecting the runoff. The data of the study area used to simulate the rainfall-runoff process, along with their sources, are presented in Table 1.

Dataset	Data	Source
Satellite images	Landsat 8 OLI (Operational Land Imager) and TIRS (Thermal Infrared Sensor)	USGS downloaded from its website <a href="https://earthexplore.usgs.gov/">https://earthexplore.usgs.gov/</a>
	ASTER Global Digital Elevation Model (GDEM)	<a href="https://daac.ornl.gov/SOILS/guides/Global_Hydrologic_Soil_Group.html">https://daac.ornl.gov/SOILS/guides/Global_Hydrologic_Soil_Group.html</a>
	Global Hydrologic Soil Groups (HYSGs250m) for Curve Number-Based Runoff Modeling	<a href="https://www.amar.org.ir/english/">https://www.amar.org.ir/english/</a>
Spatial data	Point	Synoptic stations
	Polygon	Land use Babol catchment boundary
Census data	Total monthly precipitation in each synoptic station in Mazandaran province	National Cartographic Center (NCC), Iran Statistical Center of Iran, downloaded from its website <a href="https://www.amar.org.ir/english/">https://www.amar.org.ir/english/</a>

Table 1. The data set used in this research.

## 2.3. Data creation and preparation

In order to have a CA simulation of the rainfall-runoff process, it is necessary to use some data manipulation and data analyses functions to the data so that the data can be entered into the CA model. The input data of the CA model were prepared using ArcGIS 10.3 software. In the remainder of this section, the details of the steps required for the creation and preparation of the data in ArcGIS 10.3 software are described explicitly.

### 2.3.1. Preparation of the precipitation layer

One of the layers that should be created is the precipitation layer. In fact, there is no daily precipitation data in the study area. The only available data is the monthly precipitation data, collected merely for the synoptic stations (Figure 2). Therefore, in order to obtain the continuous precipitation layers in the entire study area, an interpolation method must be applied to the monthly precipitation measures of the 14 synoptic stations in Mazandaran province. To provide continuous layers of precipitation in each month in the study area, several interpolation methods were used. After comparing the values of Root Mean Square Error (RMSE) obtained from different interpolation methods, the Kriging interpolation method was selected and used. In this way, the monthly continuous precipitation layers for the whole study area were obtained in each month. In order to calculate the daily precipitation amounts in the study area, the monthly precipitation amounts are divided by the number of days in each month, and thus, the average daily precipitation amounts are obtained for the study area.

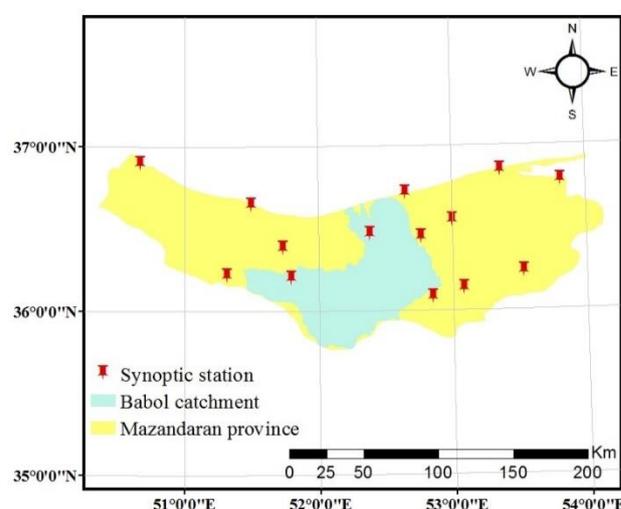


Figure 2. The location of synoptic stations in Mazandaran province.

### 2.3.2. Computation of runoff curve number (CN)

In order to predict the runoff values from the rainfall measures, the National Resources Conservation Service formerly known as the Soil Conservation Service (SCS) has developed a method called the "SCS runoff curve number or CN" (SCS 1985). In order to estimate the runoff in the SCS-CN method, several parameters including antecedent moisture condition (AMC) class, area's hydrologic soil group (HSG), land use, and the hydrologic condition of the area are required.

#### 2.3.2.1. Determination of AMC classes in the study area

Runoff is affected by the AMC, which is the soil moisture before the occurrence of the rainfall. In general, the AMC classes are defined on the basis of the daily precipitation value. In other words, the amounts of precipitation in the previous five days are added up, then the setting of Table 2 is used to determine the class of AMC. Later, the AMC class is used to define the curve number (CN) value. The CN value will be described in the remainder of this paper.

AMC class	Five-days precipitation (mm)	
	Dormant months (mm)	Growing months (mm)
	December to April	May to November
I	<12.7	>35.6
II	12.7-27.9	35.6-53.3
III	>27.9	>53.3

Table 2. Determination of AMC class based on the precipitation of the five days (SCS 1985).

As previously mentioned, since the amount of daily precipitation in the study area was not available; therefore, first, the monthly rainfall in the whole area was calculated by interpolation. Then, by dividing the monthly precipitation into the number of days in each month, the average daily precipitation was obtained. Next, by multiplying the average daily precipitation values by five, the total precipitation of 5 days in the whole area was obtained. Therefore, in this way, the AMC class of the study area was determined for each month.

### 2.3.2.2. Determination of HSGs in the study area

The HSGs is a fundamental component in the CN method of the United States Department of Agriculture (USDA) for estimation of rainfall-runoff. In order to use the soil data in the CN method, the Natural Resource Conservation Services (NRCS) categorized the soils based on the runoff potential into four Global Hydrologic Soil Groups (HSGs) of A, B, C, and D. These four HSGs groups (A, B, C, and D) correspond to soils with low, moderately low, moderately high, and high runoff potential, respectively. In order to determine the HSGs in the study area, the gridded dataset of HSGs (HYSOGs250m) with a resolution of approximately 250m was used (Ross et al. 2018). In Figure 3, the HSGs of the study area are shown.

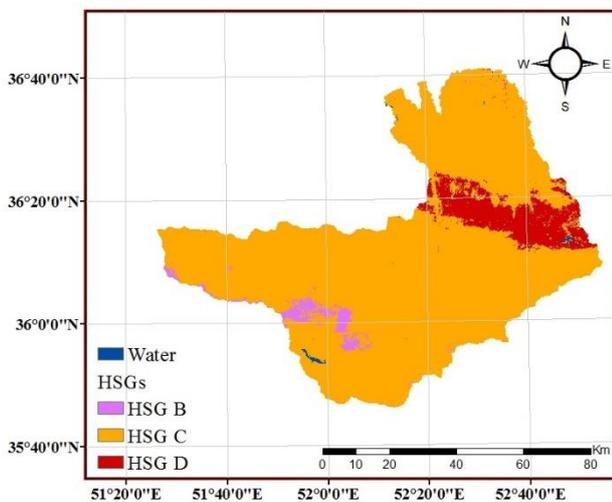


Figure 3. The HSGs of the study area.

### 2.3.2.3. Determination of land use in the study area

In order to assign CN values in the study area, it is necessary to determine the type of land use. For this purpose, the Landsat 8 satellite image was used. The Landsat 8 satellite image was classified into 11 classes of water, urban, rock, forest (poor), forest (fair), forest (good), grassland (poor), grassland (fair), grassland (good), woods with grass, and agriculture. This was carried out in the ArcGIS 10.3 software using supervised classification method. In Figure 4, the classified map of land use in the Babol catchment is shown.

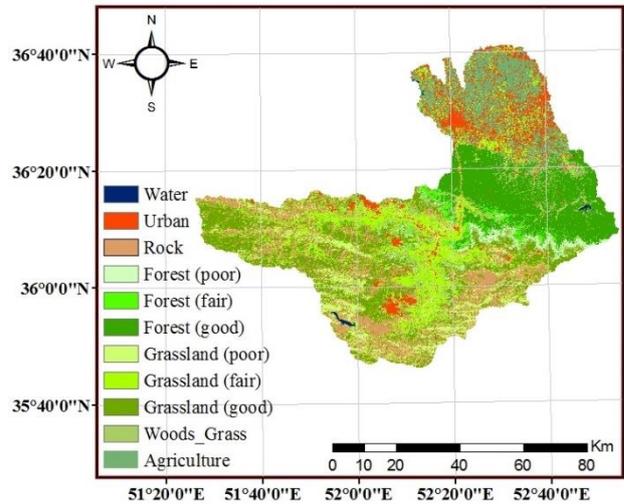


Figure 4. Land use map of Babol catchment.

In order to extract the value of the CN from the tables defined by the SCS, the classified land use map derived from the satellite image was combined with the HSGs map. Then, according to the hydrological conditions of each land use in the study area, the CN values were obtained for the AMC II (normal conditions) in the study area, as shown in Figure 5. It should be noted that the CN values extracted from the tables defined by the SCS are CN values for normal conditions (CN II).

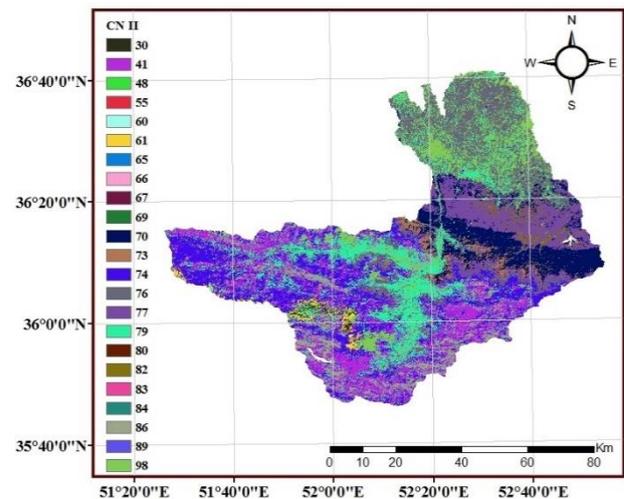


Figure 5. The CN map for AMC II (average soil moisture).

As previously mentioned, the CN values are determined by the AMC class. As shown in Table 2, AMC is grouped into three classes: AMC I: the low soil moisture (dry soil), AMC II: the average soil moisture, and AMC III: the high soil moisture (saturated soil). In order to compute the CN values for dry soil (AMC I) and saturated soil (AMC III), Equations 1 and 2 are used, respectively.

$$CN(I) = \frac{4.2 * CN(II)}{10 - (0.058 * CN(II))} \quad (1)$$

$$CN(III) = \frac{23 * CN(II)}{10 + (0.13 * CN(II))} \quad (2)$$

Where CN(I) = curve number value in class one  
 CN(II) = curve number value in class two

CN(III) = curve number value in class three

Thus, according to the AMC classes obtained for the study area in each month (see section 2.3.2.1), the CN values in each month are computed. It should be noted that according to different classes of AMC in the study area in each month of the year, the CN class will be different in different regions of the study area. In other words, there may be two or three CN classes in the study area in one month of the year.

#### 2.4. Development of the CA model

In this research, in order to develop the CA model for simulating the rainfall-runoff process, the NetLogo 6.0.4 platform is used (Wilensky 1999).

After editing and preparing the data in ArcGIS 10.3, the data is converted to the ASCII format and entered into the NetLogo 6.0.4 platform. In the flowchart of Figure 6, the steps related to the data preparation and data input to the NetLogo 6.0.4 platform are shown.

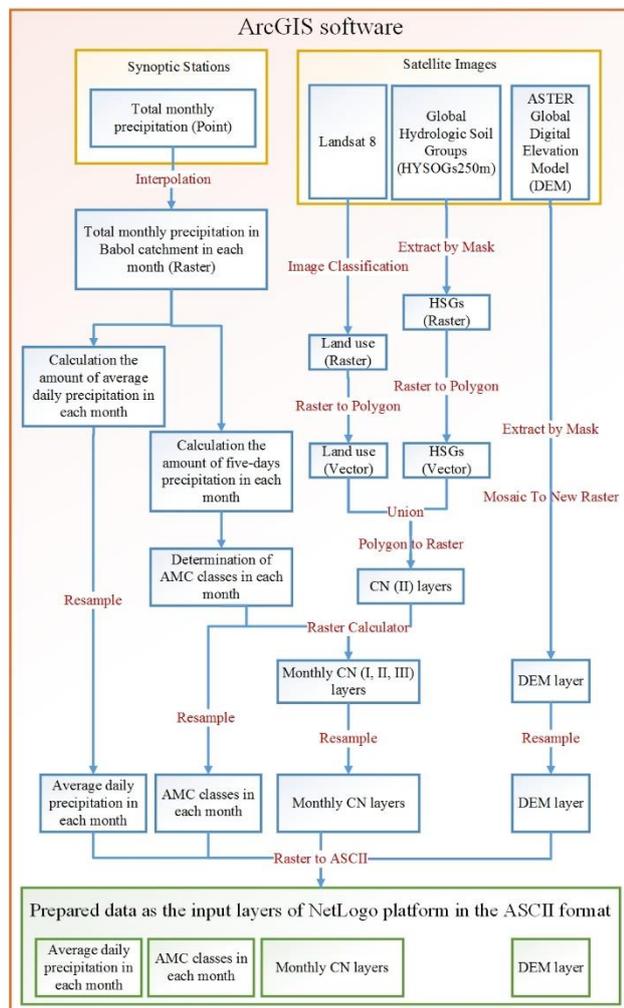


Figure 6. The flowchart of the steps related to data preparation and data input of the NetLogo 6.0.4 platform.

##### 2.4.1. Implementation of the CA Model

After importing the edited and prepared data into the ASCII format, suitable to the NetLogo 6.0.4 platform, the design and implementation of the CA model are done according to the following steps:

- I. Definition of attributes for all cells in the CA model and initializing their values according to the values of input data.
- II. Calculation of the value of potential maximum soil moisture retention after the beginning of runoff ( $S$ ) in each cell by using the CN value of the cell in each month according to Equation (3).

$$S = \frac{25400}{CN} - 254 \quad (3)$$

Where  $S$  = potential maximum soil moisture retention after the beginning of runoff (mm)  
CN = curve number value

- III. Calculation of the value of initial abstractions (amount of water before runoff such as infiltration and interception) in each cell based on the AMC class of each cell and the calculated value of  $S$  in each cell based on Equations 4, 5, and 6.

$$\text{if AMC class is I} \rightarrow I_a = 0.3S \quad (4)$$

$$\text{if AMC class is II} \rightarrow I_a = 0.2S \quad (5)$$

$$\text{if AMC class is III} \rightarrow I_a = 0.1S \quad (6)$$

Where  $I_a$  = the initial abstractions (mm)  
 $S$  = potential maximum soil moisture retention after the beginning of runoff (mm)

- IV. If the amount of average daily precipitation in the cell is greater than the amount of initial abstractions, runoff occurs, in which case the amount of runoff must be calculated.
- V. Calculation of runoff according to Equation 7 in the cell that has the highest elevation in the study area. It should be noted that several cells in the study area may have the same elevation, in which case runoff is firstly calculated in these cells.

$$Q = \frac{(P - I_a)^2}{P - I_a + S} \quad (7)$$

Where  $Q$  = the runoff depth (mm)  
 $P$  = the precipitation value (mm)  
 $I_a$  = the initial abstractions (mm)  
 $S$  = potential maximum soil moisture retention after the beginning of runoff (mm)

- VI. Calculation of the direction of flow in the cells with the highest elevation based on the eight-direction (D8) flow model (Jenson and Domingue 1988). In D8 flow model, there are eight valid output directions relating to the eight adjacent cells into which flow could travel (Figure 7). It should be noted that in the developed CA model, the cells are considered square-shape with the cell-size of 90 meters.

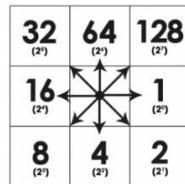


Figure 7. The coding of the direction of flow in the D8 flow model.

- VII. Finding a lower cell that receives the flow of the highest-elevation cell and calculating its total precipitation (the receiving cell is in the Moore neighborhood<sup>1</sup> of the highest-elevation cell). This is achieved by summing up the amount of runoff of the highest-elevation cells with the amount of average daily precipitation in the receiving cell. The value of precipitation in the receiving cell is then replaced by this calculated value.
- VIII. Repeating these processes in cells from the cells with the highest elevation to the cells with the lowest elevation in the study area for one day of each month and calculating the monthly runoff by multiplying the value of calculated daily runoff in the number of days of each month.

#### 2.4.2. Assumptions considered in the research

In this research, in order to simulate the rainfall-runoff process using the CA model, the following two main assumptions are considered:

- The AMC class is considered constant during each month in the study area. Consequently, the amount of CN and the amount of S during each month for the study area are considered constant.
- The amounts of precipitation are considered equal on all the days of each month.

### 3. RESULTS AND DISCUSSIONS

The primary purpose of this study was to develop a CA model for rainfall-runoff simulation in Babol catchment. The amount of runoff generated in each cell per month was calculated by the developed CA model in the NetLogo 6.0.4 platform. In Figure 8, the amount of total runoff generated in each cell during the year 2016 in the Babol catchment is shown.

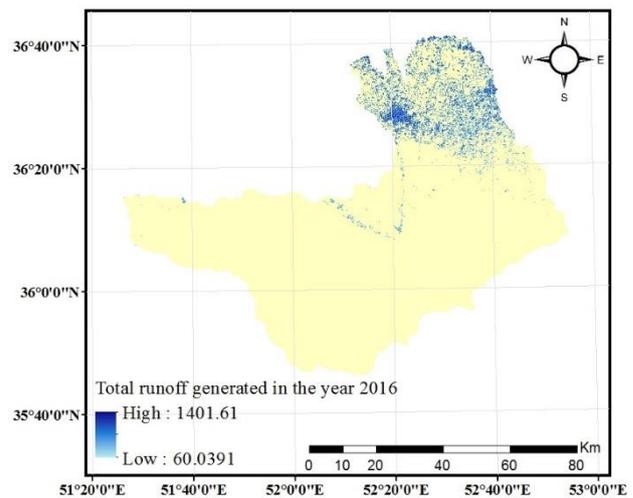


Figure 8. The total runoff generated during the year 2016 in the Babol catchment.

In order to better represent the amount of runoff generated in the study area, the amount of runoff was classified into four groups with equal intervals. In Figure 9, four classes of the runoff amount are shown.

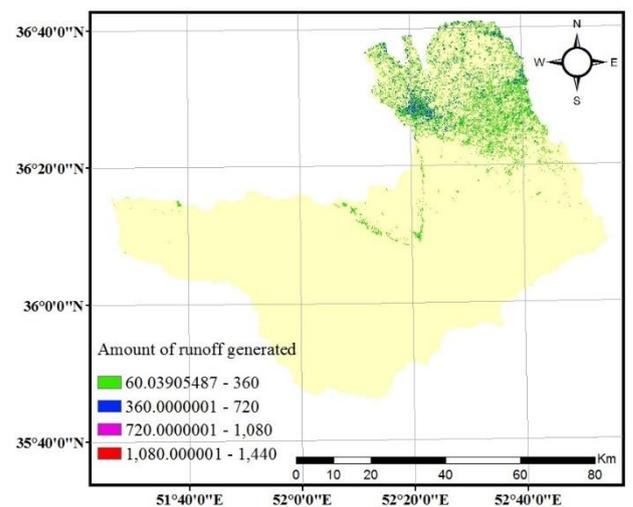


Figure 9. The amount of runoff classified into four groups.

As shown in Figure 9, the higher amounts of runoff generated during the year 2016 in most of the study area are in the range of 60-720 mm. On the other hand, most runoff has occurred in the north of the study area.

By comparing the total runoff map of the catchment, as the output of the research (Figure 8), with the DEM layer of the catchment (Figure 10), it can be concluded that the amount of runoff generated in areas with lower elevation is higher. This is because of the fact that in areas with lower elevation, the water is induced from rainfall is combined with the water flowing from regions with higher elevation, while the only existing water in areas with higher elevation is the rainfall-induced water.

<sup>1</sup> The Moore neighborhood of a cell consists of the eight cells surrounding the cell itself.

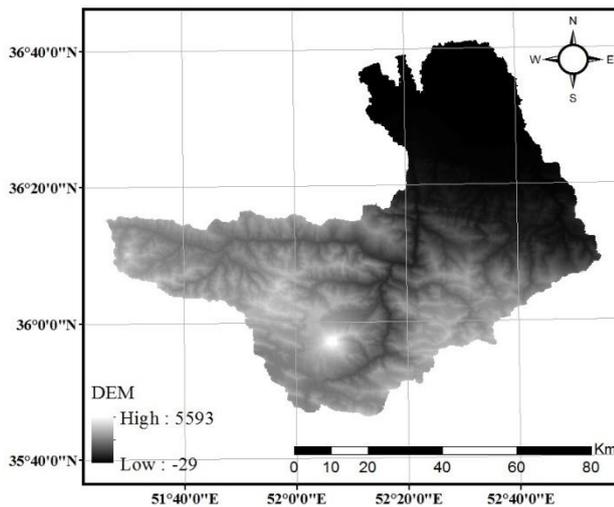


Figure 10. Digital elevation model of the Babol catchment.

On the other hand, by comparing Figures 8 and 10, it is argued that the amount of runoff generated is not affected only by the topographic conditions of the area. Several other factors, such as precipitation amount, land use, hydrological conditions of the area, and HSGs affect the amount of generated runoff. In other words, the amount of runoff generated is influenced by a combination of different factors.

#### 4. CONCLUSIONS

In this research, a model was developed to simulate the rainfall-runoff process in the Babol catchment. The model was developed on the basis of CA, GIS, and basic hydrologic principles. In the developed CA model, the target catchment was partitioned into a series of spatial cells of 90 meters size and the runoff generated in each cell was calculated based on the amount of precipitation, initial abstraction, and the potential maximum soil moisture retention, which was determined from various data. In addition, the direction of flow in each cell was determined by the D8 flow model on the basis of the difference in elevation and the horizontal distance between centers of the central cell and the eight cells of the Moore neighborhood.

In this research, in order to develop CA model to simulate the rainfall-runoff process, the NetLogo 6.0.4 platform was used. The amount of runoff generated in each cell per month of the year 2016 in the Babol catchment was calculated and presented. In addition, the total amount of runoff generated in each cell during the year 2016 was calculated and presented as the output of the research. A map of the total amount of runoff generated during the year 2016 can be an important tool for flood management decision-makers. It can also be used as an effective tool for managing the catchment and assessing water resources in the study area.

Some manipulation and preparation processes were required to be applied on the precipitation data, to overcome the limitations of the available data. First, the daily precipitation data was not available in the study area. Second, the monthly precipitation data was available as points only at synoptic stations of Mazandaran province. In addition, since Tropical Rainfall Measuring Mission (TRMM) data has a very low spatial resolution, therefore it was not suitable for use in this study. Therefore, in this model, the interpolation method was used in order to create continuous precipitation layers.

#### REFERENCES

- Beven, K.J., 2001. *Rainfall-runoff modelling*. Chichester: John Wiley & Sons.
- Jenson, S.K., Domingue, J.O., 1988. Extracting topographic structure from digital elevation data for geographic information system analysis. *Photogrammetric engineering and remote sensing*, 54(11), 1593-1600.
- Murray, A.B., Paola, C., 1994. A cellular model of braided rivers. *Nature*, 371(6492), 54.
- Parsons, J.A., Fonstad, M.A., 2007. A cellular automata model of surface water flow. *Hydrological Processes: An International Journal*, 21(16), 2189-2195.
- Rinaldi, P.R., Dalponte, D.D., Vénere, M.J., Clausse, A., 2007. Cellular automata algorithm for simulation of surface flows in large plains. *Simulation Modelling Practice and Theory*, 15(3), 315-327.
- Ross, C., Prihodko, L., Anchang, J., Kumar, S., Ji, W., Hanan, N., 2018. Global Hydrologic Soil Groups (HYSOGs250m) for Curve Number-Based Runoff Modeling. ORNL DAAC, Oak Ridge, Tennessee, USA.
- Soil Conservation Service (SCS), 1985. National engineering handbook, section 4: hydrology. *US Soil Conservation Service, USDA, Washington, DC*.
- Shao, Q., Weatherley, D., Huang, L., Baumgartl, T., 2015. RunCA: A cellular automata model for simulating surface runoff at different scales. *Journal of Hydrology*, 529, 816-829.
- Spruill, C.A., Workman, S.R., Taraba, J.L., 2000. Simulation of daily and monthly stream discharge from small watersheds using the SWAT model. *Transactions of the ASAE*, 43(6), 1431.
- Wilensky, U., 1999. *Center for connected learning and computer-based modeling*. NetLogo, Northwestern University.
- Wu, H., Chen, X., Zhang, X., 2006. A Cellular Automata based Distributed Model for Simulating Runoff in LeAnhe Watershed. *2006 IEEE International Symposium on Geoscience and Remote Sensing, IEEE*.