



Extension of the Geomorphic Flood Index classifier to predict flood inundation maps for uncalibrated rainfall depths in arid regions

Mohamed A. Hamouda^{1,2} · Ayman G. Awadallah² · Ramadan H. Abdel-Maguid²

Received: 6 September 2023 / Accepted: 22 December 2023 / Published online: 23 January 2024
© The Author(s) 2024

Abstract

Flash floods are a rapid hydrological response that occurs within a short time with rapidly rising water levels and could lead to massive structural, social and economic damages. Therefore, generating flood inundation maps becomes necessary to distinguish areas exposed to floods. Hydrodynamic models are commonly used to generate inundation maps; however, they require high computational power and time, depending on the complexity of the model. For that, researchers developed effective, fast and simplified models. Among the simplified models, the Geomorphic Flood Index (GFI) is one of the most useful classifiers to generate inundation maps. Three main objectives are addressed in this study: (1) extend the GFI classifier to predict flood extent maps for uncalibrated rainfall depths, which will enhance early warning models for better risk assessments of extreme events; (2) enhance the accuracy of the simulated inundation maps using different calibration methods; and (3) investigate the performance of the GFI in various terrains with different resolutions. Three case studies in arid regions in Saudi Arabia were examined with different topographies, using terrains of high resolutions of 1 m and resampled low resolutions, as well as various rainfall depths corresponding to 5–100-yr return periods. The HEC-RAS 2D model was used to generate reference flood inundation maps. The obtained flood extent maps show high similarity compared to the reference maps with accuracy above 80%. Strong relationships between rainfall depths and the threshold GFI parameter were developed which allow producing inundation maps for any rainfall event.

Keywords Flood inundation maps · Geomorphic Flood Index · Flash flood · Calibration methods · Arid region

✉ Ayman G. Awadallah
aawadallah@darcairo.com

¹ Dar Al-Handasah Consulting Firm, Shair and Partners, Cairo, Egypt

² Civil Engineering Department, Fayoum University, Fayoum, Egypt

1 Introduction

The threat posed by natural disasters related to weather and climate to society is growing and leads to massive loss of life and property (Glade and Nadim 2014). From 1970 to 2019, climate disasters accounted for 50% of all recorded disasters, 74% of all reported economic damages and 45% of all reported deaths, where over ninety-one percent of these deaths and injuries occurred in developing countries as reported by the United Nations Country Classification (Mandyam et al. 2022).

Flash floods are common in arid and semiarid regions during rain events because of the scarcity of plants and the low rate of infiltration in mountainous regions. Flash floods are a complex phenomenon that is influenced mainly by regional geology and morphometric characteristics of the watershed (Subyani and Al-Dakheel 2009). Other important factors that influence flash floods are rainfall intensity and duration, surface runoff, evaporation and infiltration rates (Nouh 2006). Flash floods have become a significant concern in urban areas worldwide. This is primarily due to the increase in the population pressure around the floodplains, leading to an increase in urbanization, which reduces the space available for infiltration and exacerbates surface runoff, often exceeding the water-bearing capacity of the catchment (Zaidi 2012). Even in very dry regions like the Arabian Peninsula, flash floods have caused huge number of fatalities (Youssef et al. 2016), billions of dollars of business losses, facilities damages (Al Saud 2010) and unnominated impacts (Kumar 2013). As a result, many countries have implemented early warning systems aiming to minimize the impact of natural hazards on human life and assets (Billa et al. 2004, 2011; Villagrán de León et al. 2006; Pengel et al. 2013).

Due to its multifaceted nature, the problem of warning communities of impending disasters quickly becomes complex. The main difficulty related to warning is to have enough time to predict the impact of floods in addition to estimating the area affected through accurate maps before floods occurrence. This is followed by reporting the prediction to the appropriate authorities, warning affected communities and evacuating those communities (Basha and Rus 2007). Thus, understanding flood risk is essential for controlling its socio-economic and environmental effects.

The purpose of flood inundation maps is to show, at a certain area, the impact of a certain rainfall storm event. These maps support the decision-makers to minimize damages to flooded areas and assess the risks. Models of developing flood hazard maps could be classified into three types: empirical methods, hydrodynamic models and simplified conceptual models (McGrath et al. 2018). Empirical methods can only provide guidance for real-time monitoring and post-event evaluation (Liu et al. 2018). Hydrodynamic models, such as HEC-RAS, deduce inundation maps by solving one-dimensional or two-dimensional equations. These hydrodynamic models take a long running time for specific flood events to deduce high-resolution flood extent maps based on a mesh cell size and a simulation time step (Alipour et al. 2022). On the contrary, simplified conceptual models can generate flood inundation maps in a relatively shorter time (Coulthard et al. 2013).

Developing simplified methods to generate flood maps has been tackled by many researchers. Leopold and Maddock (1953) represented the quantitative equations to determine the natural stream properties such as depth, width, velocity and suspended load based on the discharge using some power functions. In this approach, the discharge at a certain location is considered as a function of its upstream catchment area.

One of the main concerns of simplified models is to determine the starting point of a stream. There are two widely used approaches to tackle this issue: Tarboton et al. (1991)

recommended to consider the channel from the location where the flow accumulation is higher than a specific threshold; while Giannoni et al. (2005) developed a newer method to specify the first point of the channel where the quantity AS^k (A = contributing area, S = local slope, k = drainage basin-related exponent) is higher than a specific threshold.

Various methods were developed based on simplified concepts to generate flood plain maps for areas with lack of high-resolution digital elevation model (DEM) or rain data. One of these methods is generating flood-prone maps using classifiers by developing low-cost geomorphic flood plain algorithms. Many researchers worked to develop methods to define hydrogeomorphic stream networks to simplify the algorithm for generating flood plain maps based on catchment characteristics for a specific flood event (Nardi et al. 2006). Later, Nardi et al. (2008) extended the outcomes and checked the influence of generating floodplain maps in flat areas and tested various treatment methods including DEM treatment methods. Manfreda et al. (2011) developed a simple procedure for detecting areas at risk using only data of catchment topography (DEM) which are globally available. The data contained in a DEM could be summarized by a topographic index which is a good descriptor of the inundation maps, noting that the used DEMs describe the ground surface without taking into account the presence of man-made structures.

Subsequently, Degiorgis et al. (2012) developed a classifier which depends on two main variables chosen among five available ones which could be deduced from the utilized DEM. The first selected variable is the length of the stream that hydrologically links the study element to the nearest element of the drainage network, and the second selected variable is the difference in elevation between the cell under investigation and the end point of the same drainage network.

Different flood descriptors and classifiers were developed and evaluated in terms of their suitability to generate flood plain maps based on DEM and catchments characteristics (Manfreda et al. 2015). After testing many areas, Samela et al. (2016) found that the Geomorphic Flood Index (GFI) is the best classifier that could be used to define flood extent across the utilized DEM. The output inundation maps generated by GFI were compared with FEMA flood plain maps and found that, for getting the optimum value of threshold of the generated inundation maps, it should be first calibrated against a reference flood map obtained from the outcomes of a hydrodynamic model or any other regulatory authority, for an area exceeding 2% of the study area (Samela et al. 2017). A QGIS tool called Geomorphic Flood Area tool (GFA tool) was successfully developed based on the classifier of GFI concept to detect flood-prone areas at the continental scale while reducing computational time and costs, opening up new possibilities for flood risk assessment and management at large scales Samela et al. (2018). The method was improved to predict the depth of the flooding as well, which helps measure how much damage the flood caused (Manfreda and Samela 2019). According to Tavares da Costa et al. (2020), optimal GFI thresholds were also positively correlated with flood extents associated with specific return periods. The developed tool preliminary provides hydrologic and hydraulic models in those regions where there is obvious shortage in rainfall data.

In this paper, we investigate the use of the GFA tool in arid regions where flooding is rare and thus not carving the terrain. Further, we investigate the use of high- and low-resolution terrains, and test the tool in different locations with different topographic characteristics. Thus, the main objectives of this study can be summarized in the following:

- Evaluate the generated flood inundation maps using GFA tool in arid regions comparing with the ones generated using hydrodynamic model results.

- Evaluate the effect of changing GFI parameters and the impact of various rain depths at different return periods and DEM resolutions on the output inundation maps in arid regions.
- Enhance flood inundation maps generated using GFA tool by evaluating different calibration criteria / methods.
- Develop relationships between the output threshold and the rainfall depth to help rapidly obtaining a flood plain map, for an uncalibrated rain event, which could be used in early warning systems.

The paper is organized as follows. The current introduction and literature review are followed by a detailed methodology (Sect. 2) as well as the performance and evaluation criteria. Section 3 presents the study areas. The paper results and their discussion are provided in Sect. 4 and followed by the main conclusions and recommendations for future research (Sect. 5).

2 Methodology

The current section describes the methodology used to generate accurate inundation maps based on the flood descriptor GFI. The parameters of the GFI are calibrated based on generated flood maps using the two-dimensional hydrodynamic HEC-RAS model. The main investigated parameter is the threshold value of the GFI that determines whether a certain location is prone to flooding for that specific rain event. In this paper, we extend the capability of the GFA tool to map uncalibrated rainfall events through establishing relationships to predict the threshold values based on DEM characteristics of the catchment. Moreover, the GFA tool was tested on different digital elevation models to check the impact of DEM resolution on the output flood plain maps. Figure 1 describes how the hydrogeomorphic models and the linear binary classifiers GFI can be used to generate floodplain maps. Hereafter, we present the main components of each models/loop and how we checked the output for uncalibrated rainfall events.

2.1 Development of a hydrodynamic model using HEC-RAS

The River Analysis System (RAS) freely available software, developed by the Hydrologic Engineering Center (HEC-RAS), is a commonly used software that allows to model from 1D steady flow to 2D unsteady flow, dam break analysis, sediment transport, as well as generating floodplain maps.

In this study, two flow regime options were utilized (1) inflows as upstream boundary condition for hydraulic simulation for part of the watercourse and (2) precipitation on the entire catchment area in the model area. In our models, shallow water equations (SWE) were used based on Saint Venant equations (Horváth et al. 2015). This method utilizes finite volumes solution algorithm for more stability and accuracy than the traditional algorithms of finite differences or finite elements. The 2D simulation algorithm starts with initial depth, continuing to the movement of water between cells calculated based on starting depth and cell face rating curve. After that, the new cell volume is calculated based on flux over computational time step. Finally, depth in each cell extracted from volume depth rating curves. This computational algorithm is robust and allows 2D cells to wet and dry. Two-dimensional flow areas can start completely dry and can handle a sudden rush

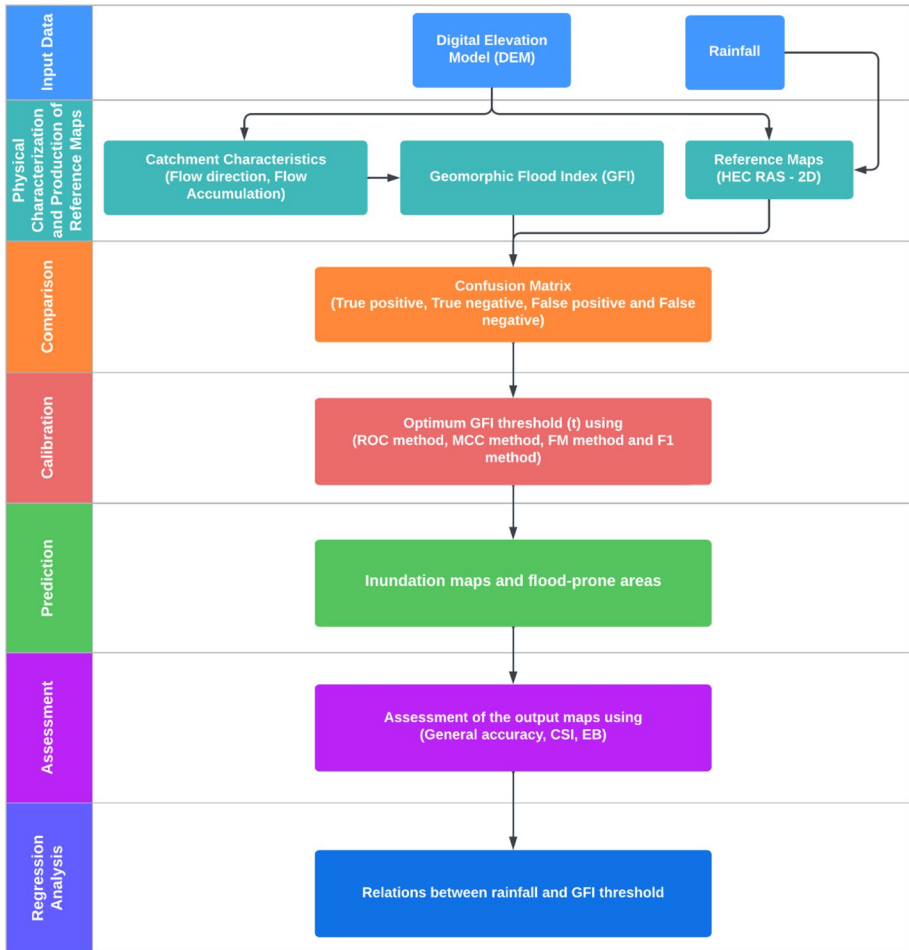


Fig. 1 Process of the geomorphic model

of water into them. The spread of water over a surface is dependent on the roughness of the surface. Manning’s n value is used to define roughness of the main watercourse and its floodplain or on the entire catchment. HEC-RAS outputs are used as the flood maps benchmark.

2.2 Simplified model: the GFA tool

The GFA is a tool that utilizes the GFI to generate inundation maps based on geomorphic indices extracted from terrains and existing inundation maps obtained for a portion of the area of interest. This is accomplished by subdividing the study area to small cells utilizing the DEM and then categorizing each cell inside the area of interest domain into two groups: flooded areas and non-flooded areas (Samela et al. 2017).

The GFA tool requires geomorphic data to calculate the GFI to derive floodplain maps. The geomorphic data could be extracted from terrain. The required input files are:

- DEM with voids and/or local low points (sinks): the original digital elevation model from any source, either SRTM, ALOS, LIDAR or contour map information.
- DEM fill: the raster of DEM is filled using a suitable tool (ArcGIS, QGIS, etc.) by removing local sinks and peaks in the data, which are spurious cells with elevations greater or lower than what would be expected.
- Flow direction: the raster of flow directions from each pixel to its steepest downslope neighbor. Using the D8 direction method by comparing the elevation relative to 8 (eight) surrounding cells, flow directions are determined based on the direction of the steepest slope.
- Flow accumulation: the raster of accumulated flow into each pixel, as determined by accumulating the number of all pixels that flow into each downslope cell.
- Part of the existing floodplain map: to calibrate the GFI classifier and produce maps for the sub-scheme, a reference flood map for a minimum of 2% of the domain is used to estimate floodplain maps for the remaining areas (Samela et al. 2017). The reference floodplain could be obtained from FEMA, NOAA, USGS in the USA or authorities already developed floodplain maps or could be generated from any hydrodynamic models (HEC-RAS, InfoWorks ICM, Urban flow) as it is the case in this research.

2.3 Preprocessing steps

After preparing the input data, the next step in the GFA Tool is to select one of the following methods to identify the starting point of each stream inside the study area.

- AS^k method: The start of the channel is determined where the flow accumulation exceeds a specific threshold (Tarboton et al. 1991).
- The channel starts from the location where the flow accumulation exceeds a specific threshold (Tarboton et al. 1991).

This is needed as the main GFA process (described in the next subsections) realize on the calculations of the catchment area upstream of a certain cell.

2.4 The GFA process

The value of GFI at each cell is estimated as Eq. (1):

$$\text{GFI} = \ln \frac{h_r}{H} = \ln \frac{aA_r^n}{H} = \ln(a) + n \ln(A_r) - \ln(H) \quad (1)$$

where h_r is the difference between the water level and the stream bed level [m], A_r is the contributing area of water runoff to the point of interest [km²], a is a scale factor, n is the exponent (dimensionless) and H [m] is the difference between the elevation of each cell and the elevation of the nearest stream point of the identified path, as shown in Fig. 2. The water depth (WD) could be calculated for every cell of the floodplain areas as follows (Manfreda and Samela 2019):

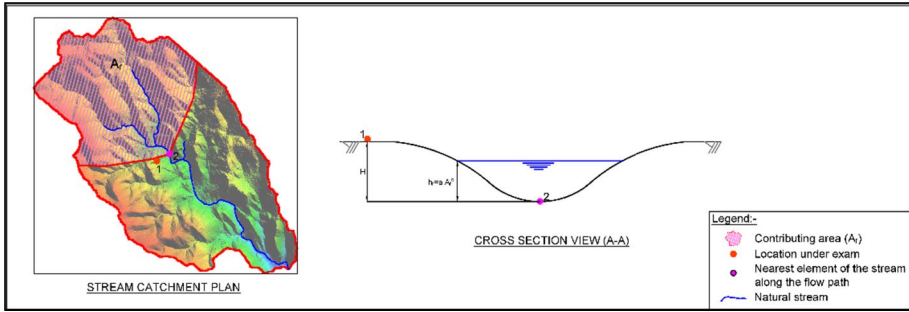


Fig. 2 Main components of the GFI classifier [after Manfreda et al. (2015)]

$$WD = h_r - H \tag{2}$$

If the hydraulic model under study is not covering the entire catchment area and an inflow hydrograph is imposed at the upstream of the hydrodynamic model, the total contributing area—input in Eq. (1)—is in fact subdivided into two sub-areas. Sub-area 1 (shown hatched in Fig. 2) comprises the area upstream of the hydrodynamic model. It is determined via catchment delineation software, using the common, flow direction, flow accumulation, etc. procedure. Moreover, the remaining Sub-area 2 (part not hatched in Fig. 2) accounts for the part within the hydrodynamic model boundary. As we move further toward the downstream, the contributing area (within the model boundary) increases, added to Sub-area 1 (at the upstream of the model as previously mentioned). It should also be highlighted that the catchment area at the most downstream of the model is equal to Sub-area 1 + the model area, and as such, the part that is modeled via the hydrodynamic model (as well as the GFI tool) constitutes a complete sub-catchment area.

The GFA tool then normalizes the GFI values using as much as needed of iterations to calibrate the optimal threshold (t) that produces the flood potential at a required location (i.e., $WD \geq t$, the cell is considered flooded). Moreover, it expands the classification between flood-prone and non-flood-prone locations to cover the entire study area.

The GFA uses a confusion matrix as shown in Table 1A, where each column represents the reference classification and each row represents the predicted classification. Table 1B illustrates four different performance measures which are used to calibrate the optimal threshold (t). The optimal threshold (t) value for a specific inundation map corresponding to a certain period is calculated by testing the threshold of the classifier in the range from -1 to 1 and then comparing the resulting GFI classifier with the floodplain maps used for calibration.

2.5 Performance criteria

Four various performance criteria were investigated to compare between each cell in the reference maps and the corresponding one in the predicted maps of each GFI value using the confusion matrix (Fathi et al. 2021) to obtain the optimum GFA threshold:

- The receiver operating characteristic (ROC) graph is a method to quantify a binary classifier as its limit (Fawcett 2006). The ROC values in Fig. 3 are calculated by the

Table 1 Confusion matrix with the performance measures to assess the calibration

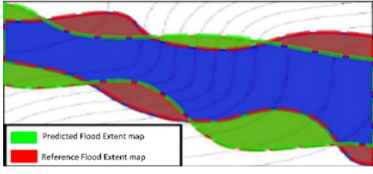
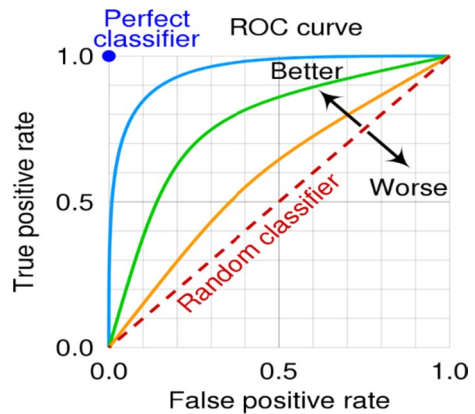
A-The confusion matrix for each element			
	<u>Predicted Flood Extent</u>		
	Flood	No Flood	
<u>Reference Flood Extent</u>	Flooded	True Positive (TP)	False Negative (FN)
	Not Flooded	False Positive (FP)	True Negative (TN)
B-The performance measures to assess the calibration process			
True Positive Rate	$R_{TP} = \frac{TP}{TP + FN}$	Sensitivity	
False Negative Rate	$R_{FN} = 1 - R_{TP}$	Underestimation	
True Negative Rate	$R_{TN} = \frac{TN}{TN + FP}$	Specificity	
False Positive Rate	$R_{FP} = 1 - R_{TN}$	Overestimation	

Fig. 3 Receiver operating characteristic (ROC) graph, drawn by CMG Lee based on <http://commons.wikimedia.org/wiki/File:roc-draft-xkcd-style.svg>



scatter plot of the true positive (R_{TP}) against false positive (R_{FP}) rates as illustrated in Table 1B. Equation (3) presents the calculation of the area under the curve (Bradley 1997).

$$ROC = \frac{1 + R_{TP} - R_{FP}}{2} \tag{3}$$

- Matthew’s correlation coefficient (MCC) is a method to assess binary (two-class) classifications (Matthews et al. 1979). This measure is similar to the Pearson correlation coefficient in its interpretation. The coefficient takes into account true and false positives and negatives. It is generally regarded as a balanced measure which can be used, even if the classes are of very different sizes. The MCC can be calculated directly from the confusion matrix using the formula:

$$MCC = \frac{TP * TN - FP * FN}{\sqrt{(TP + FP) * (TP + FN) * (TN + FP) * (TN + FN)}} \tag{4}$$

- Fowlkes–Mallow’s index is a third performance criteria. A higher value for this index indicates a greater similarity between the clusters and the benchmark classifications, when results of two clustering algorithms are used to evaluate the results. The Fowlkes–Mallow’s index is written as (Halkidi et al. 2001):

$$FM = \sqrt{\frac{TP}{TP + FP} * \frac{TP}{TP + FN}} \tag{5}$$

- *F*-score index is calculated using the precision and recall, where the precision is the number of true positive results divided by the number of all positive results, including those not identified correctly, while the recall is the number of true positive results divided by the number of all samples that should have been identified as positive. It is formulated as follows (Sasaki 2007):

$$F1 = \frac{2 TP}{2 TP + FP + FN} \tag{6}$$

2.6 Assessment methods

After obtaining the optimum threshold (*t*), three independent measures were calculated to further assess the correspondence between the inundation maps generated by the hydrodynamic models and those by the GFA tool using the following equations as shown in Table 2.

As the objective is to predict a flood extent envelope, overpredicting the flood-prone areas might benefit the TP and inflate the R_{TP} and accuracy. This inflation would be of concern if there was no other reported measure that would give an alternative account of the performance. By reporting the R_{FP} , an account of the percentage of cells that are over-predicted is given. At the same time, the critical success index extends the R_{TP} by including the FP, accounting for both underprediction and overprediction. The error bias gives the ratio between the FP and the FN, indicating whether there is a tendency for underprediction or overprediction. The reporting of these three measures should give a reasonable overall account of the performance.

2.7 Regression analysis

Regression analysis is a technique used to develop relations between two or more variables. In the context of assessing flood hazard using the GFI, regression analysis was employed in this research to establish a quantitative relationship between GFI values and rainfall depth, DEM cell size or other variables, with the purpose of understanding the influence of rainfall on flood extend maps produced by the GFA tool.

The regression analysis allows deriving the mathematical relationship between GFI values and rainfall depth. This equation can then be used to predict the flood potential based on an uncalibrated rainfall of 24-h depth. By establishing such a relationship, decision-makers and urban planners can make informed decisions regarding flood management strategies, land-use planning and disaster preparedness. Additionally, it can aid

Table 2 Assessment methods for the obtained threshold

General accuracy	$= \frac{TP+TN}{TP+TN+FP+FN}$	Ranges from 0 (no match) to 1 (identical match)
Critical success index (CSI):	$= \frac{TP}{TP+FN+FP}$	Ranges from 0 (no match) to 1 (identical match) (Bullen and Rockart 1981)
Error bias (EB):	$= \frac{FP}{FN}$	Underestimation, $0 \leq EB < 1$, or overestimation, $EB > 1$, with $EB = 1$ an indication of no bias (Breiman 2001)

in the development of early warning systems that utilize 24-h rainfall depth to anticipate potential flood events and take appropriate preventive measures.

3 Study areas

In recent years, Saudi Arabia has suffered from flash floods which caused major losses in lives and damages. This study encompasses 3 areas as shown in Fig. 4 and Table 3. The first study area for catchment (A) is located in Al-Baha province with high-steep terrain at its upstream and flatten terrain at the downstream, which could be a good setting to evaluate GFA tool. Another study area (B) covers sections (not an entire catchment) of Wadi Hanifa in Riyadh to test the ability of the tool to generate flood maps for part of a stream. The third study area (C) is for a portion of Wadi Bayer (Al-Jouf, Saudi Arabia) in a flat area at the downstream of the wadi. In latest studies, the GFA tool was tested on DEMs with low resolutions 30m for big streams. In the current study, the high-resolution DEMs were resampled for different resolutions from 1.0 to 90.0 m and then utilized in generating flood plain maps using the GFA tool. The tool was also tested for various rainfall depths corresponding to return periods (5, 10, 25, 50 and 100 yr). It is worth mentioning that the area denoted as model area in Table 3 is the one where the HEC-RAS and the GFI tool are applied, while the catchment area upstream is the one at the upstream of the model boundary used in the calculation of A_r in Eq. (1).

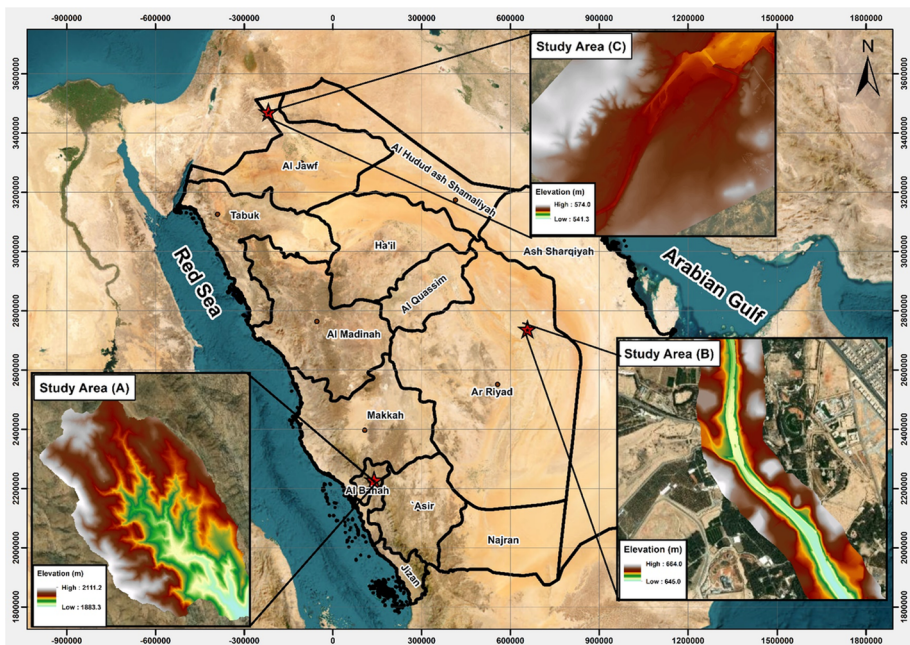


Fig. 4 Study areas and the utilized DEM

Table 3 Summary of the characteristics of the 3 study areas

Study area	Model area (km ²)	Catchment area upstream of model (km ²)	Original DEM cell size (m)	Average stream slope (%)	Max stream length (m)	Inflow peak discharge—100 yr (m ³ /s)	Rainfall depth—100 yr (mm)	Average width of streams (m)
Al-Baha (study area A)	5.80	–	1.00×1.00	38.65	3101	–	188.71	15
Wadi Hamifa (study area B)	4.02	1461	0.50×0.50	9.19	1650	154.82	59.52	52
Wadi Bayer (study area C)	3.14	3175	1.00×1.00	3.22	2146	380.11	57.80	200

4 Results and discussion

The methodology described in Sect. 2 was followed to allow assessing the performance of the GFA tool and to deduce flood extents for uncalibrated rainfall events. First, the benchmark flood maps for the study areas were generated using HEC-RAS 2D models, with various DEM resolutions (cell sizes) and for different terrain types. The 2D models are using two types of boundary conditions: flow hydrographs upstream of the 2D flow area and rain-on-grid using excess rainfall for the entire 2D catchment area. The input hydrographs and excess rainfall depths were used for rainfall events corresponding to various return periods (5, 10, 25, 50 and 100 yr).

The assessment of the performance of the GFA tool was analyzed using the following process:

4.1 Benchmarks flood extent maps

The obtained benchmark maps, in the 1.0-m-resolution DEM, illustrate the maximum flood extents using the input hyetographs or upstream hydrographs. For study areas (B) and (C), an inflow hydrograph is added as upstream boundary condition to account for the flow from the contributing upstream catchment of the modeled area. As shown in Fig. 5 for study area (A), the flood extent for 100-yr return period is narrow at the upstream and gets wider at the downstream, where the small streams are joining to form a large wadi. For area (B), the wadi flood extent for the 100-yr return period is well defined as the cross section of the wadi is V-shaped defined and the water depth is high. For area (C), the flood

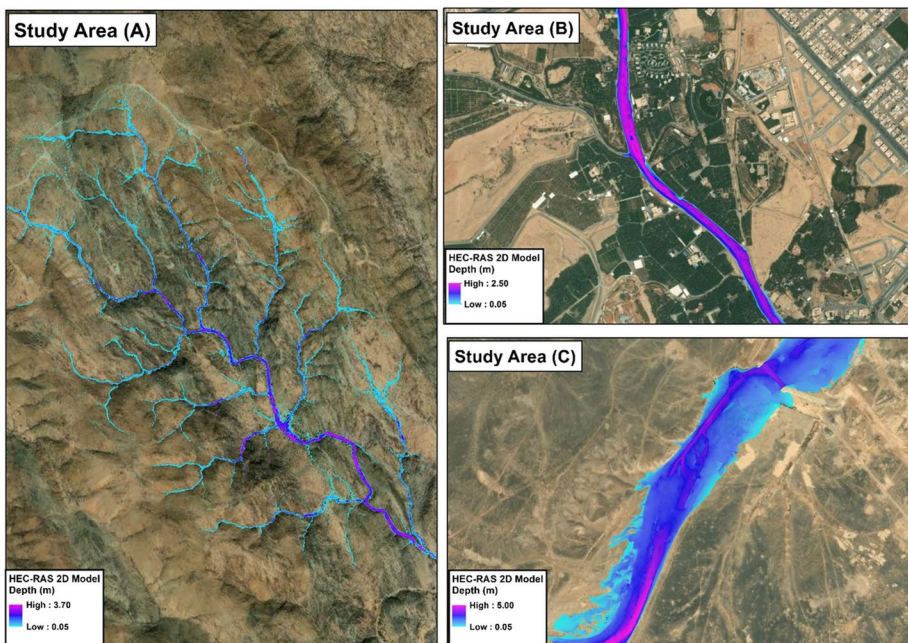


Fig. 5 Maximum flood extent of the 100-yr return period for the three study areas

extent for 100-yr return period is wide as the area is almost flat, and similar results were obtained for the resampled resolutions of the DEM.

4.2 Sensitivity of the results to the used evaluation criteria in the calibration

The GFA tool uses a minimum of 2% of the area of the input flood maps generated using HEC-RAS 2D models as a reference for the calibration process. The tool converts the reference floodplain map into a binary map, where “1” represents flooded areas and “0” represents unflooded areas. To correctly simulate the flood extent, the threshold parameter must be calibrated. The threshold ranges from -1 to 1 where (-1) means maximum width of streams and (1) means minimum width of streams. For study areas (B) and (C), the total contributing area is considered in the calculation of GFI values and output thresholds.

Four evaluation criteria were tested: ROC, MCC, FI and FM to obtain the optimum value of threshold and detect the flooded areas at different return periods. Figure 6 indicates that the GFA tool correctly classifies the raster cells as flood-prone areas and non-flood-prone areas with a probability of success that depends on the terrain types. Figure 6 as an example shows that, for the study area (A), the GFA tool has also a high success rate. It accurately interpolates and estimates missing flood extent. However, in few locations, the GFA tool was not able to accurately simulate it. For defined wadi's topography such as study area (B), the GFA tool could generate flood maps with high accuracy very similar to the reference maps. For the study area (C), the generated flood extent maps using the GFA tool are wider than the reference maps become of the flat topography.

The obtained values of the thresholds for study area (A), shown in Table 4A, range between -0.259 and -0.747 . In general, the absolute value of threshold increases with

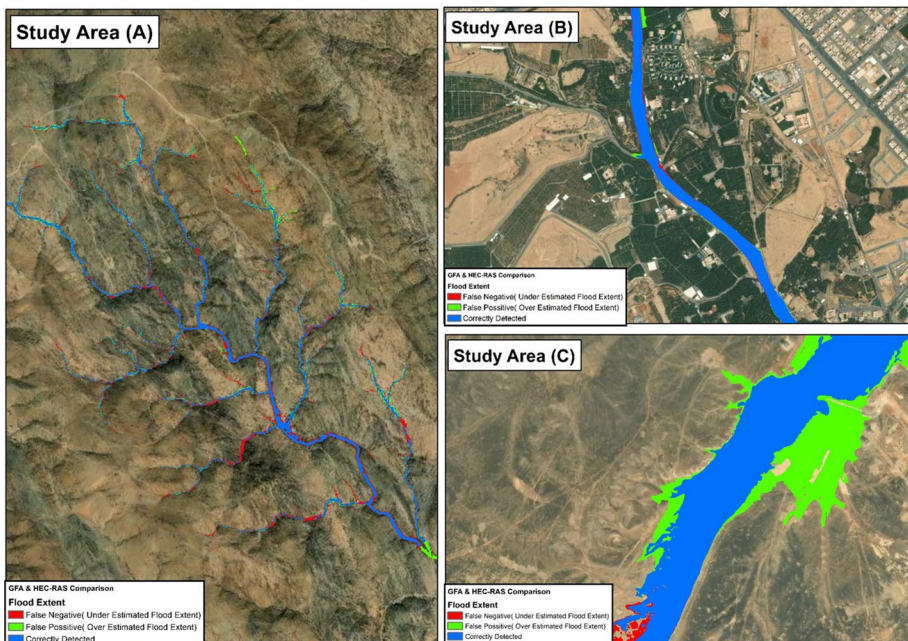


Fig. 6 Flood inundation maps for 100-yr return period for the study areas

Table 4 Obtained threshold values and comparison between calibration methods for study area (A)

A—Threshold value for each calibration method				
Return period	t —ROC	t —MCC	t —F1	t —FM
2	− 0.292	− 0.259	− 0.259	− 0.259
5	− 0.363	− 0.363	− 0.385	− 0.385
10	− 0.384	− 0.384	− 0.427	− 0.427
25	− 0.418	− 0.402	− 0.607	− 0.688
50	− 0.432	− 0.418	− 0.688	− 0.703
100	− 0.450	− 0.426	− 0.703	− 0.747
B—Comparison between calibration methods for event of 100-yr return period				
	ROC	MCC	F1	FM
Threshold	− 0.450	− 0.426	− 0.703	− 0.747
Accuracy	86%	88%	81%	80%
CSI	0.71	0.89	0.59	0.59
EB	0.73	0.84	2.48	4.01

the increase in the rainfall and flood extent. For low-rainfall events, the four evaluation methods resulted in similar threshold values and same extents. The values obtained from the MCC and ROC methods are very similar across return periods. However, the values of thresholds obtained using F1 and FM methods are almost equal for return periods less than 25 yr. Comparing the four methods for the 100-yr return period, as shown in Table 4B, the results show that MCC and ROC methods give better results with accuracy higher than 85%, and critical success index associated with using MCC is about 0.89, which indicates a high similarity between the obtained flood extent maps and the reference maps. Moreover, the error bias is about 0.80 which also indicates that the obtained flood extent maps show a number of underestimated cells higher than the number of overestimated cells as compared to the reference maps.

The obtained threshold values for study area (B) are shown in Table 5A. As the wadi is well defined, the absolute values of thresholds are convergent and range from 0.5 to 0.75. The output thresholds using F1 and FM methods are almost equal. Table 5B shows that the investigated four methods are good for calibration for study area (B) with an accuracy higher than ninety percent. Similar to study area (A), the MCC is the best method among the four methods to obtain the optimum threshold with a high accuracy of 97% and flood extent maps similar to flood maps generated using hydraulic models (HEC-RAS). Furthermore, critical success index associated with using MCC is about 0.97 and the error bias is about 1.04 which also indicates that the obtained flood extent maps are neither underestimating nor overestimating the reference flood extents.

Table 6A presents the obtained values of threshold using the four calibration methods for study area (C). The study area has flat slopes with a wide flood extent. Thus, the absolute values of thresholds are high and range between 0.68 and 0.75. Moreover, the difference between the values of thresholds is small for various return periods. Table 6B shows that the four methods are nearly acceptable to obtain the optimum thresholds with accuracy higher than 75%. However, the MCC gives the highest accuracy of 85% and the critical success index associated with using MCC is about 0.81

Table 5 Output threshold values and comparison each calibration method for study area (B)

A—Threshold value for each calibration method				
Return period	t —ROC	t —MCC	t —F1	t —FM
5	− 0.554	− 0.582	− 0.589	− 0.589
10	− 0.615	− 0.631	− 0.631	− 0.631
25	− 0.649	− 0.667	− 0.671	− 0.671
50	− 0.666	− 0.690	− 0.693	− 0.693
100	− 0.679	− 0.708	− 0.750	− 0.750
B—Comparison between calibration methods for event of 100-yr return period				
	ROC	MCC	F1	FM
Threshold	− 0.68	− 0.71	− 0.75	− 0.75
Accuracy	91%	97%	91%	91%
CSI	0.91	0.97	0.90	0.90
EB	0.84	0.87	0.86	0.86

Table 6 Output threshold values and comparison each calibration method for study area (C)

A—Threshold value for each calibration method				
Return period	t —ROC	t —MCC	t —F1	t —FM
5	− 0.728	− 0.728	− 0.731	− 0.736
10	− 0.736	− 0.736	− 0.732	− 0.732
25	− 0.783	− 0.789	− 0.831	− 0.831
50	− 0.789	− 0.848	− 0.873	− 0.874
100	− 0.791	− 0.858	− 0.906	− 0.906
B—Comparison between calibration methods for event of 100-yr return period				
	ROC	MCC	F1	FM
Threshold	− 0.791	− 0.858	− 0.906	− 0.906
Accuracy	82%	85%	75%	75%
CSI	0.81	0.81	0.74	0.74
EB	0.84	1.24	1.41	1.41

while the error bias is about 1.24 which indicates that the obtained flood extent maps are overestimated compared to the reference maps.

From the above investigation and the obtained evaluation criteria values, both ROC and MCC are best as calibration criteria with an advantage of MCC in all 3 terrain types (study areas) as the MCC produces better accuracy and critical success index with an error bias close to unity.

4.3 Sensitivity analysis of GFI parameters

Various values of parameters a and n were examined to check the sensitivity of changing each parameter on the obtained threshold. For the same study areas, results show that different values of parameter (a) with a fixed value of (n) have no impact on the obtained values of threshold (t). On the contrary, changing the value of parameter (n) has a significant effect on the threshold. The optimum values of n range from 0.1 to 0.5 to get the optimum values of threshold with high accuracy as shown in Table 7. For defined wadis with suitable slopes, such as study area (B), the value of parameter n has no impact on the obtained optimum value of threshold.

4.4 Relationships between rainfall depths and threshold values

Figure 7 presents an overall visualization of this sensitivity analysis results. In general, there is an increase in the absolute value of threshold with the increase in the rainfall depth. The regression analysis shows that square root transformation of the rainfall depth provides the highest R^2 between rainfall depth and threshold (generated for the thresholds related to the 5-yr to 100-yr return periods, obtained using MCC as calibration criterion).

Table 8 shows that there are strong relationships between rainfall depths and thresholds with low standard error, significant p -values for all parameters and high adjusted R^2 . A rainfall depth value between rainfall depths of 50- and 100-yr return period is utilized to test the predicted thresholds using developed relations.

As shown in Table 9, the predicted threshold shows a high skill in obtaining flood extent maps similar to the generated flood extent maps using hydrodynamic models. The accuracy

Table 7 Obtained threshold values for each n value for three study areas

n	MCC			
	Threshold	Accuracy (%)	CSI	EB
<i>A—study area (A)</i>				
0.1	− 0.426	88	0.89	0.84
0.3	− 0.363	79	0.80	0.75
0.5	− 0.363	79	0.80	0.74
0.7	− 0.222	68	0.78	0.68
1	− 0.097	28	0.48	0.47
<i>B—study area (B)</i>				
0.1	− 0.709	97	0.97	0.87
0.3	− 0.709	97	0.97	0.87
0.5	− 0.709	97	0.97	0.87
0.7	− 0.707	97	0.97	0.85
1	− 0.708	97	0.97	0.86
<i>C—study area (C)</i>				
0.1	− 0.858	85	0.81	1.24
0.3	− 0.887	84	0.76	1.33
0.5	− 0.890	84	0.75	1.35
0.7	− 0.892	84	0.75	1.39
1	− 0.890	83	0.73	1.39

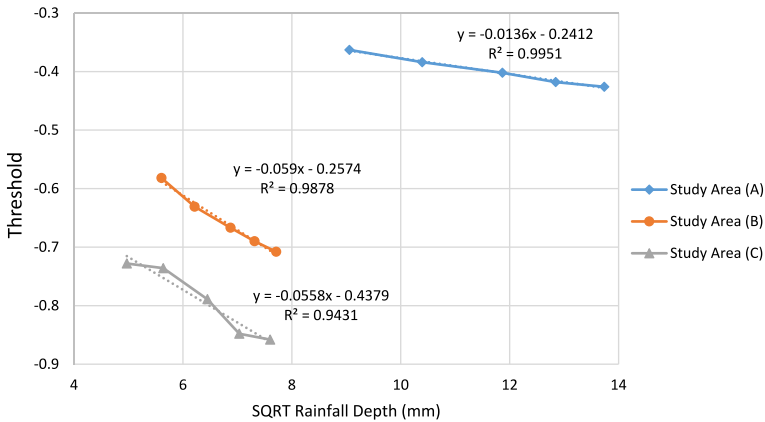


Fig. 7 Relations between rainfall depths and thresholds for the study areas

of flood extent maps obtained from predicted threshold is from 87 to 98%. Furthermore, the critical success index is very high which indicates high rate of true positive values. Moreover, the error bias is close to unity which means that number of underestimated cells is almost equal to the number of overestimated cells. Figure 8 shows flood extent maps generated using HEC-RAS and flood extent maps generated based on the value of threshold calculated from the developed regression equations.

4.4.1 Sensitivity to the DEM resolution

Finally, to investigate the sensitivity of the GFA tool results to the DEM resolution, the digital elevation model of study area (A) (originally of 1.0m resolution) was resampled at different DEM cell sizes (3.0, 5.0, 10.0, 30.0 and 90.0 m). Figure 9 shows that the relationships between DEM cell size and the obtained threshold values have an asymptotic pattern.

As above shown in Fig. 9, the value of the threshold is constant for various return periods in DEM resolution of 90 m. As a general trend, the threshold tends to be lower (more negative) with coarser resolutions. In fact, the relationship between DEM resolutions and the obtained threshold values shows similar patterns for DEM of 1m to 10m resolution. The threshold values converge starting from 30-m-resolution DEM.

5 Conclusions

The GFA is a valuable tool for assessing flood risk and understanding the potential impacts of flooding in a given area. The tool utilizes advanced geomorphic analysis techniques to evaluate the likelihood and severity of flooding. One of the key strengths of the GFI tool is its ability to integrate multiple data sources and generate comprehensive flood risk maps, as demonstrated by the current study. It is also important to acknowledge that the accuracy and reliability of the GFI tool depend on the quality and availability of input digital elevation model data.

We have investigated the capabilities of the GFA tool to generate inundation maps and detect flood-prone areas in various locations in arid regions using digital elevation models

Table 8 Results of regression analysis between rainfall depths and thresholds for the three study areas

Study area	Developed relation	Adjusted R^2	Significance of F -statistics	Standard error	P -Value for slope	P -value for intercept
(A)	$(t) = -0.0136\sqrt{D} - 0.2412$	0.993	0.00014	0.002	0.00014	0.00004
(B)	$(t) = -0.059\sqrt{D} - 0.2574$	0.983	0.0005	0.006	0.00057	0.00211
(C)	$(t) = -0.0558\sqrt{D} - 0.4379$	0.924	0.005	0.016	0.0059	0.00327

where D is the rainfall depth (mm)

Table 9 Predicted threshold values for uncalibrated rainfall depth

Study area	Rainfall depth (mm)	Predicted threshold (t)	Accuracy (%)	CSI	EB
(A)	120	-0.3901	88	0.87	0.852
(B)	50	-0.6745	98	0.82	0.85
(C)	50	-0.8324	87	0.87	1.13

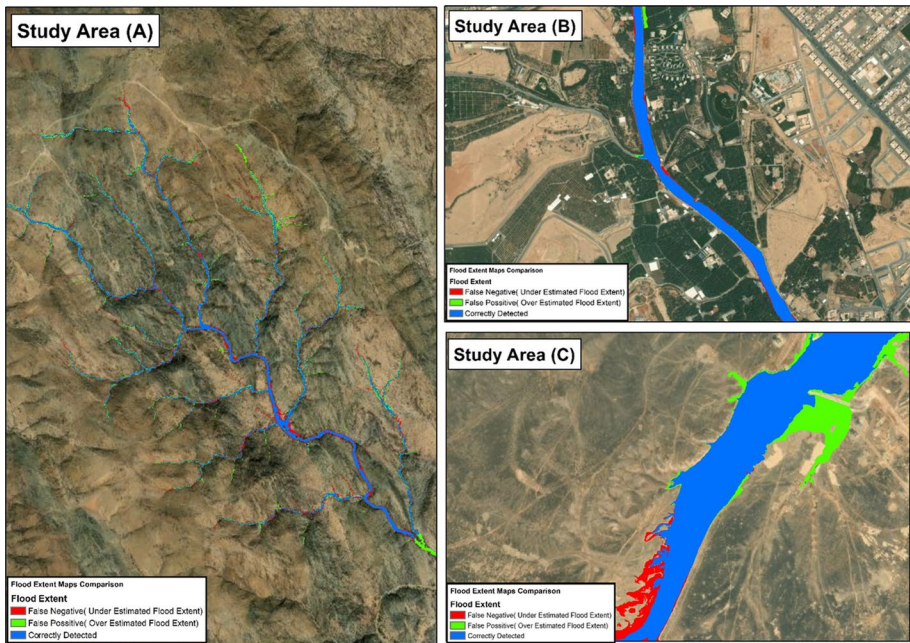


Fig. 8 Flood plain maps for tested rainfall events for the study areas

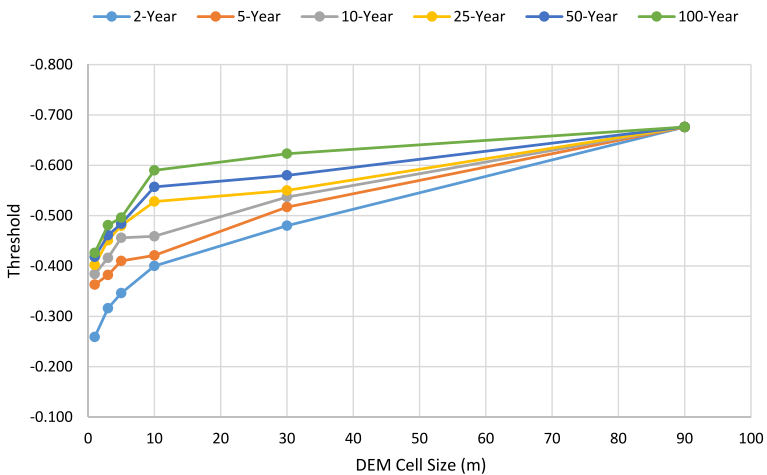


Fig. 9 Threshold values with various DEM resolutions for different return periods

with high resolutions and applied various rainfall depths corresponding to different return periods. The main conclusions drawn from this research could be summarized as follows:

- In the GFA method, the resulting value of threshold is not affected by changing the value of parameter (a) while fixing the value of parameter (n) (i.e., no impact on flood extent). On the other hand, the obtained value of threshold (t) changes noticeably when the value of parameter (n) is changed. The best values of parameter (n) range from 0.1 to 0.5 in order to obtain the optimum threshold and increase the accuracy of the flood extent maps.
- To get more accurate inundation maps in arid regions, the MCC evaluation criterion is compared to ROC, F1 and FM evaluation criteria. The outputs show that MCC method is the best calibration criterion to get optimum threshold values for different types of topographic. Moreover, for large defined wadis MCC and ROC methods could be similar in obtaining the optimum value of threshold.
- The GFA tool is best suited to predict flood extent maps for areas with high slopes and defined streams. Conversely, generating flood extent maps in flat areas with low slopes requires more enhancement and research.
- We found that the threshold value has strong relationship with the square root of the rainfall depth, which extends the applications of the GFA tool to generate inundation maps for uncalibrated rainfall events. Nevertheless, the application of the derived equation is limited to return periods from 10 yr to 100 yr (or even above).
- For areas with small streams widths, especially in mountains areas, to generate accurate flood extent maps, it is recommended to use digital elevation models with high resolution (cell size equal or less than 10 m). For coarser resolutions, the threshold values are similar for various rainfall depths (i.e., the same flood extent is predicted regardless of the input rainfall depths).

Author contributions All authors contributed to the study conception and design. Research idea and conceptualization were proposed by Ayman G. Awadallah. Material preparation, data collection and analysis were performed by Mohamed A. Hamouda. The first draft of the manuscript was written by Mohamed A. Hamouda, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Open access funding provided by The Science, Technology & Innovation Funding Authority (STDF) in cooperation with The Egyptian Knowledge Bank (EKB). The authors declare that no funds, grants or other support were received during the preparation of this manuscript.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Al Saud M (2010) La Cartographie des zones potentielles de stockage de l'eau souterraine dans le bassin Wadi Aurnah, située à l'Ouest de la Péninsule Arabique, à l'aide de la Télédétection et le Système d'Information Géographique. *Hydrogeol J* 18:1481–1495. <https://doi.org/10.1007/s10040-010-0598-9>
- Alipour A, Jafarzadegan K, Moradkhani H (2022) Global sensitivity analysis in hydrodynamic modeling and flood inundation mapping. *Environ Model Softw* 152:105398
- Basha E, Rus D (2007) Design of early warning flood detection systems for developing countries. In: 2007 International Conference on Information and Communication Technologies and Development, pp 1–10. <https://doi.org/10.1109/ICTD.2007.4937387>
- Billa L, Mansor S, Mahmud AR (2004) Spatial information technology in flood early warning systems: an overview of theory, application and latest developments in Malaysia. *Disaster Prev Manag an Int J* 13:356–363. <https://doi.org/10.1108/09653560410568471>
- Billa L, Mansor S, Mahmud AR (2011) Pre-flood inundation mapping for flood early warning. *J Flood Risk Manag* 4:318–327. <https://doi.org/10.1111/j.1753-318X.2011.01115.x>
- Bradley AP (1997) The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recogn* 30:1145–1159. [https://doi.org/10.1016/S0031-3203\(96\)00142-2](https://doi.org/10.1016/S0031-3203(96)00142-2)
- Breiman L (2001) Random forests. *Mach Learn* 45:5–32. <https://doi.org/10.1023/A:1010933404324>
- Bullen CV, Rockart JF (1981) A Primer on Critical Success Factors. Massachusetts Institute of Technology, Sloan School of Management, Massachusetts, USA
- Coulthard TJ, Neal JC, Bates PD et al (2013) Integrating the LISFLOOD-FP 2D hydrodynamic model with the CAESAR model: implications for modelling landscape evolution. *Earth Surf Process Landforms* 38:1897–1906
- Degiorgis M, Gnecco G, Gorni S et al (2012) Classifiers for the detection of flood-prone areas using remote sensed elevation data. *J Hydrol* 470–471:302–315. <https://doi.org/10.1016/j.jhydrol.2012.09.006>
- Fathi MM, Awadallah AG, Awadallah NA (2021) Estimation of regional sub-daily rainfall ratios using SKATER algorithm and logistic regression. *Water Resour Manag* 35:555–571. <https://doi.org/10.1007/s11269-020-02730-1>
- Fawcett T (2006) An introduction to ROC analysis. *Pattern Recogn Lett* 27:861–874. <https://doi.org/10.1016/j.patrec.2005.10.010>
- Giannoni F, Roth G, Rudari R (2005) A procedure for drainage network identification from geomorphology and its application to the prediction of the hydrologic response. *Adv Water Resour* 28:567–581. <https://doi.org/10.1016/j.advwatres.2004.11.013>
- Glade T, Nadim F (2014) Early warning systems for natural hazards and risks. *Nat Hazards* 70:1669–1671. <https://doi.org/10.1007/s11069-013-1000-8>
- Gracheva IV, Orlova NV (1975) Vliianie kontsentratsii fosfora na obrazovanie novobiotsina produktentom Act. spheroides [Effect of the phosphorus concentration on novobiocin formation by the producer Act. spheroides]. *Antibiotiki* 20(10):871–876
- Halkidi M, Batistakis Y, Vazirgiannis M (2001) On clustering validation techniques. *J Intell Inf Syst* 17:107–145. <https://doi.org/10.1023/A:1012801612483>
- Horváth Z, Waser J, Perdígão RAP et al (2015) A two-dimensional numerical scheme of dry/wet fronts for the Saint-Venant system of shallow water equations. *Int J Numer Methods Fluids* 77:159–182. <https://doi.org/10.1002/flid.3983>
- Kumar A (2013) Natural hazards of the Arabian Peninsula: their causes and possible remediation. Earth system processes and disaster management. Springer, Berlin, pp 155–180
- Leopold LB, Maddock T (1953) The hydraulic geometry of stream channels and some physiographic implications, vol 252. US Government Printing Office
- Liu C, Guo L, Ye L et al (2018) A review of advances in China's flash flood early-warning system. *Nat Hazards* 92:619–634. <https://doi.org/10.1007/s11069-018-3173-7>
- Mandyam S, Priya S, Suresh S, Srinivasan K (2022) A correlation analysis and visualization of climate change using post-disaster heterogeneous datasets. *arXiv prepr arXiv:2205.12474*
- Manfreda S, Samela C (2019) A digital elevation model based method for a rapid estimation of flood inundation depth. *J Flood Risk Manag* 12:1–10. <https://doi.org/10.1111/jfr3.12541>
- Manfreda S, Di Leo M, Sole A (2011) Detection of flood-prone areas using digital elevation models. *J Hydrol Eng* 16:781–790. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000367](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000367)
- Manfreda S, Samela C, Gioia A et al (2015) Flood-prone areas assessment using linear binary classifiers based on flood maps obtained from 1D and 2D hydraulic models. *Nat Hazards* 79:735–754. <https://doi.org/10.1007/s11069-015-1869-5>

- Matthews BW, Fenna RE, Bolognesi MC et al (1979) Structure of a bacteriochlorophyll a-protein from the green photosynthetic bacterium *Prosthecochloris aestuarii*. *J Mol Biol* 131:259–285. [https://doi.org/10.1016/0022-2836\(79\)90076-7](https://doi.org/10.1016/0022-2836(79)90076-7)
- McGrath H, Bourgon J-F, Proulx-Bourque J-S et al (2018) A comparison of simplified conceptual models for rapid web-based flood inundation mapping. *Nat Hazards* 93:905–920. <https://doi.org/10.1007/s11069-018-3331-y>
- Nardi F, Vivoni ER, Grimaldi S (2006) Investigating a floodplain scaling relation using a hydrogeomorphic delineation method. *Water Resour Res* 42:1–15. <https://doi.org/10.1029/2005WR004155>
- Nardi F, Grimaldi S, Santini M et al (2008) Hydrogeomorphic properties of simulated drainage patterns using digital elevation models: the flat area issue. *Hydrol Sci J* 53:1176–1193. <https://doi.org/10.1623/hysj.53.6.1176>
- Nouh M (2006) Wadi flow in the Arabian Gulf states. *Hydrol Process* 20:2393–2413. <https://doi.org/10.1002/hyp.6051>
- Pengel BE, Krzhizhanovskaya VV, Melnikova NB, et al (2013) Flood early warning system: sensors and internet. In: IAHS-AISH publication, pp 445–453
- Samela C, Troy TJ, Sole A, Manfreda S (2016) A new geomorphic index for the detection of flood-prone areas at large scale. In: *Convegno nazionale di idraulica e costruzioni idrauliche*, pp 14–16
- Samela C, Troy TJ, Manfreda S (2017) Geomorphic classifiers for flood-prone areas delineation for data-scarce environments. *Adv Water Resour* 102:13–28
- Samela C, Albano R, Sole A, Manfreda S (2018) A GIS tool for cost-effective delineation of flood-prone areas. *Comput Environ Urban Syst* 70:43–52
- Sasaki Y (2007) The truth of the F-measure. *Teach tutor mater* 1(5):1–5
- Subyani AM, Al-Dakheel AM (2009) Multivariate geostatistical methods of mean annual and seasonal rainfall in southwest Saudi Arabia. *Arab J Geosci* 2:19–27. <https://doi.org/10.1007/s12517-008-0015-z>
- Tarboton DG, Bras RL, Rodriguez-Iturbe I (1991) On the extraction of channel networks from digital elevation data. *Hydrol Process* 5:81–100. <https://doi.org/10.1002/hyp.3360050107>
- Tavares da Costa R, Zanardo S, Bagli S et al (2020) Predictive modeling of envelope flood extents using geomorphic and climatic-hydrologic catchment characteristics. *Water Resour Res* 56:1–23. <https://doi.org/10.1029/2019WR026453>
- Villagrán de León JC, Bogardi J, Dannenmann S, Basher R (2006) Early warning systems in the context of disaster risk management. *Entwicklung Ländlicher Raum* 2:23–25
- Youssef AM, Sefry SA, Pradhan B, Alfadail EA (2016) Analysis on causes of flash flood in Jeddah city (Kingdom of Saudi Arabia) of 2009 and 2011 using multi-sensor remote sensing data and GIS. *Geomat, Nat Hazards Risk* 7:1018–1042. <https://doi.org/10.1080/19475705.2015.1012750>
- Zaidi FK (2012) Hydrological characterization of Mahd Ad Dahab Gold Mine, Saudi Arabia. *Int J Phys Sci* 7:2935–2943. <https://doi.org/10.5897/IJCIR12.206>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.