YEMEN WATERSHED AND RUNOFF POTENTIAL
ANALYSIS METHODOLOGY

INTRODUCTION

Yemen has been consistently affected by severe floods over the past months. The variation of rain patterns due to climate change and the land use greatly contribute to these natural disasters.

The increase of impervious surfaces, related to the increase of urbanized land, directly affects the occurrence of damaging floods in populated areas.

So far, there have been studies on flood risk and susceptibility in Yemen, but there has not been any attempt to evaluate the surface runoff during peak rain intensity events. This study aims at providing an understanding of the runoff volumes which could affect Yemen, as well as estimate the runoff routing, by leveraging satellite-based data and GIS technology.

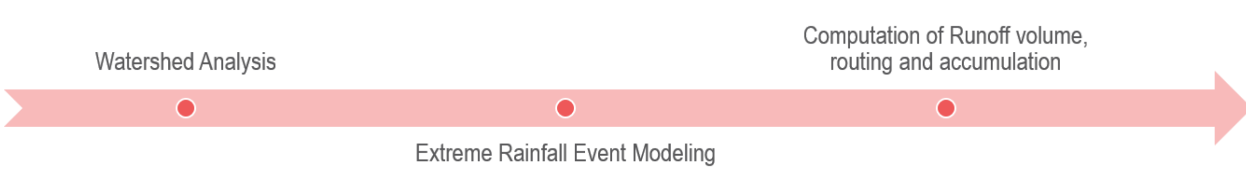
INPUT DATASETS

	Name	Resolution	Period	Reference
Digital Elevation Model	NASADEM Merged DEM Global 1 arc second V001	30 m	~	NASA JPL (2020). NASADEM Merged DEM Global 1 arc second V001 [Data set]. NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/MEaSUREs/NASADEM/NASADEM_HGT.001
Curve Number	GCN250	250 m	~	Jaafar, H.H., Ahmad, F.A. & El Beyrouthy, N. GCN250, new global gridded curve numbers for hydrologic modeling and design. *Sci Data* **6,** 145 (2019). https://doi.org/10.1038/s41597-019-0155-x
Precipitation	CHIRPS Daily: InfraRed Precipitation w/ Station Data v2	0.05 arc degrees	1984 - 2019	Funk, Chris, Pete Peterson, Martin Landsfeld, Diego Pedreros, James Verdin, Shraddhanand Shukla, Gregory Husak, James Rowland, Laura Harrison, Andrew Hoell & Joel Michaelsen. "The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes". Scientific Data 2, 150066. doi:10.1038/sdata.2015.66 2015.

METHODOLOGY

The Runoff analysis was conducted on a country level scale through the following steps:

1. Watershed Analysis
2. Calculation of maximum rain intensity for a given Return period (10, 20 and 50 years respectively, for 24h duration)
3. Runoff Volume
4. Runoff Accumulation



Watershed Analysis

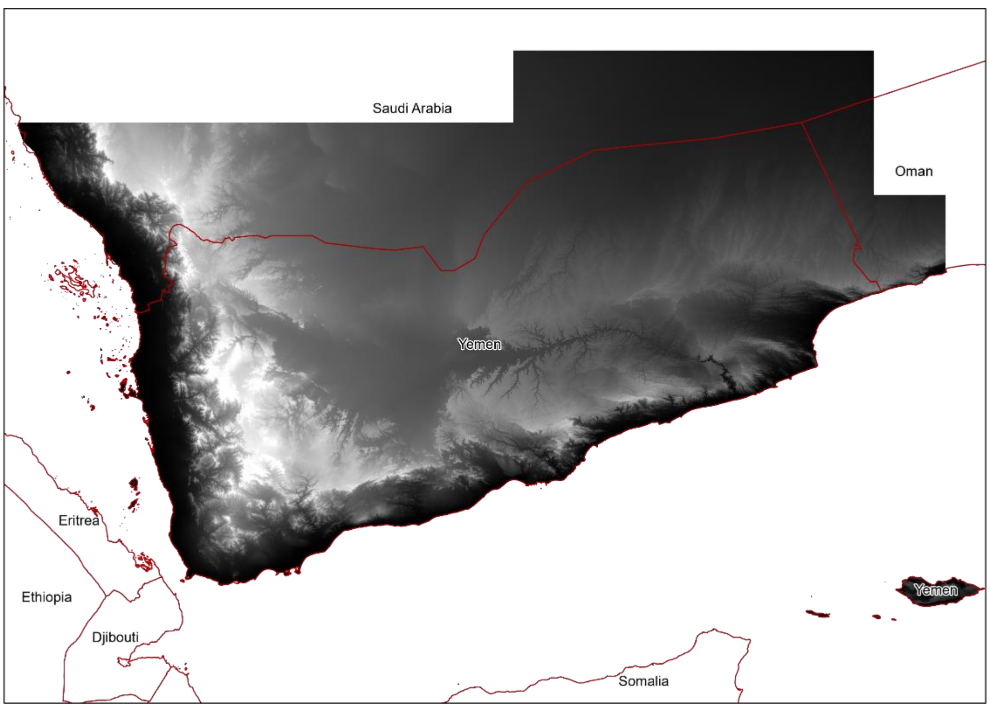
A watershed is an area of land that drains all the streams and rainfall to a common outlet such as the outflow of a reservoir, mouth of a bay, or any point along a stream channel. Watershed analysis refers to the process of using Digital Elevation Model (DEM) and raster data operations to delineate watersheds and to derive features such as streams, stream network, catchment areas, basin etc. Delineation of watersheds can take place at different spatial scales. A large watershed may cover an entire stream system and, within the watershed, there may be smaller watersheds, one for each tributary in the stream system.

In this study the method for watershed delineation is area-based. This divides a study area into a series of watersheds, one for each stream section.

ArcGIS Pro 2.4.0 with Spatial Analyst extension were used to delineate watershed components in this analysis in Yemen by using the Digital Elevation Model (DEM) raster dataset from NASA with 30 meters spatial resolution to extract flow direction as well as flow accumulation, drainage basins, and stream network.

Steps

1. **DEM acquisition**
   Digital Elevation Models (DEM) are a gridded digital representation of terrain, with each pixel value corresponding to a height above a datum. For this analysis, the DEM covers whole Yemen and parts of Saudi Arabia and Oman since watershed boundaries could be extended over country boundaries.

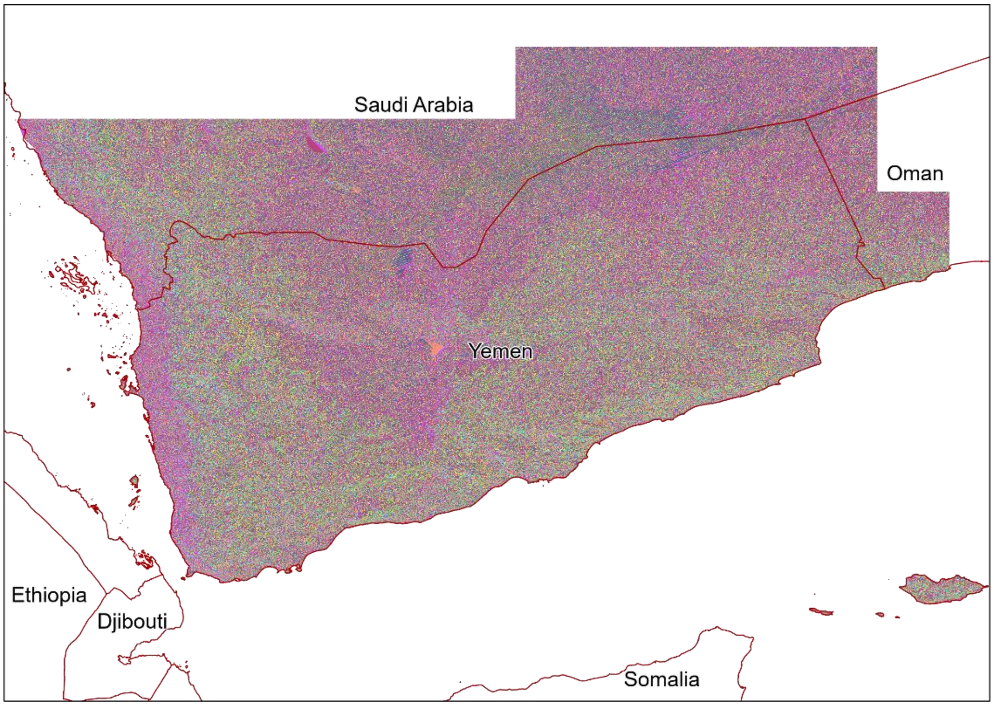


NASADEM

2. **Fill**
   Sinks (and peaks) are often errors due to the resolution of the data or rounding of elevations to the nearest integer value. Sinks should be filled to ensure proper delineation of basins and streams. If the sinks are not filled, a derived drainage network may be discontinuous.

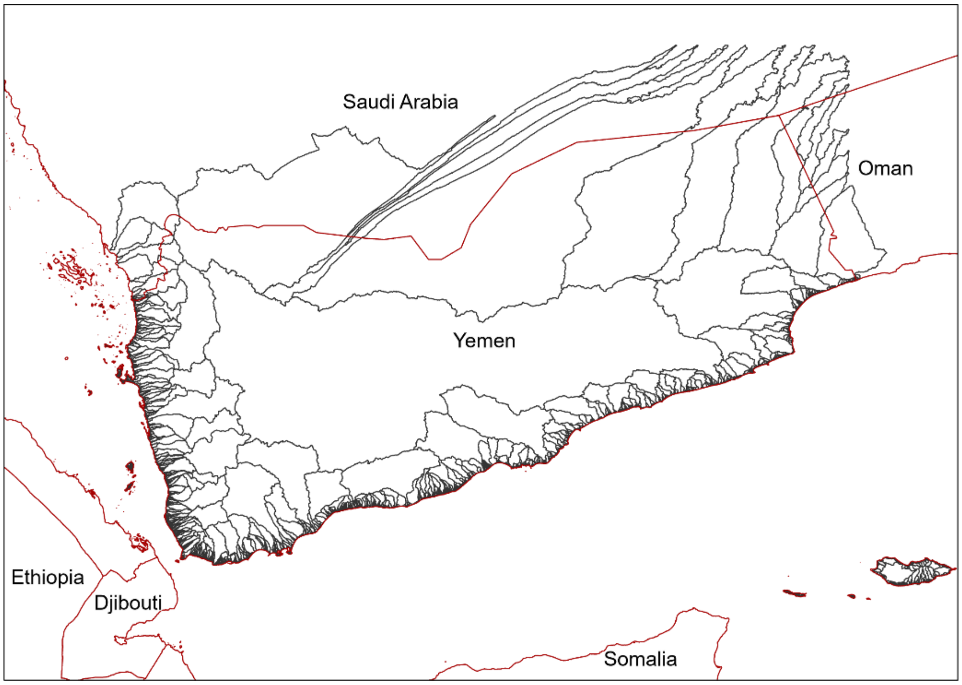
   Fill tool from ArcGIS Pro was used to eliminate those imperfections.

3. **Flow direction**
   Flow direction raster shows the direction of flow out of each cell. The D8 flow method was used to model flow direction from each cell to its steepest downslope neighbor. The Flow Direction tool was used to extract flow direction raster from filled DEM.



Flow Direction

4. **Drainage Basins**
   The drainage basins are delineated by identifying ridge lines between basins. Basin tool was used to extract drainage basin from flow direction raster because it is based on the ridges throughout the raster and divides the entire raster into different basins. It is more useful for large study areas or unfamiliar sites because it automatically selects the pour points.

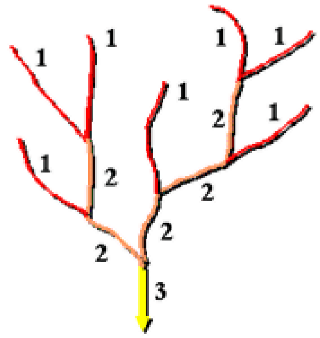


Watershed boundaries

5. **Flow Accumulation**
Flow accumulation raster represents the accumulated flow into each cell. Flow accumulation in its simplest form is the number of upslope cells that flow into each cell. Cells with a high flow accumulation are areas of concentrated flow and may be used to identify stream channels as well as cells with a flow accumulation of 0 are local topographic highs and may be used to identify ridges. Based on D8 flow direction raster, flow accumulation was developed.

6. **Stream Network**
Stream networks can be delineated by applying a threshold value to flow accumulation raster. In this analysis, a threshold value of 5,000 was used to extract the stream network, meaning that all cells with more than 5,000 cells flowing into them are assigned as a stream. The threshold value is higher than normal because the analysis is at national level and cover large area.

Stream Ordering: Stream ordering is a method of assigning a numeric order to links in a stream network. This order is a method for identifying and classifying types of streams based on their numbers of tributaries.
In this analysis, the **Strahler** method was used to define stream order where all links without any tributaries are assigned an order of 1 and are referred to as first order. The stream order increases when streams of the same order intersect. Therefore, the intersection of two first-order links will create a second-order link, the intersection of two second-order links will create a third-order link, and so on. The intersection of two links of different orders, however, will not result in an increase in order.



Strahler Method

Rain Intensity Probability

The return period of an extreme event is defined as the recurrence interval of such event, and it is generally expressed in years. Essentially, in this case, the return period estimates how long it will be between two rainfall events of a given magnitude (Clarke, 2006).

It is important to note that this is based on statistical calculations, therefore two comparable extreme events could occur in a shorter or larger timespan than the calculated return period.

In order to estimate the runoff during an intense storm, the maximum rain intensity with a recurrence interval of 10, 20 and 50 years respectively was calculated (see Figure below). This was done through the Generalized Extreme Value Distribution Function (GEV). A daily rainfall time series of 36 years was used to calculate it (from 1984 to 2019), based on satellite data (Marra et al., 2017). For each year, the maximum precipitation was selected for every area in Yemen. This was then used to calculate the GEV and extract the maximum rainfall for a duration of 24 hours.

The choice of using this function compared to others (e.g. Gumbel or Log – Pearson Type 3) is because it has proved to fit better than others the calculation of maximum intensity events per year (the Annual Maxima Series or AMS) (Marra et al. 2017).

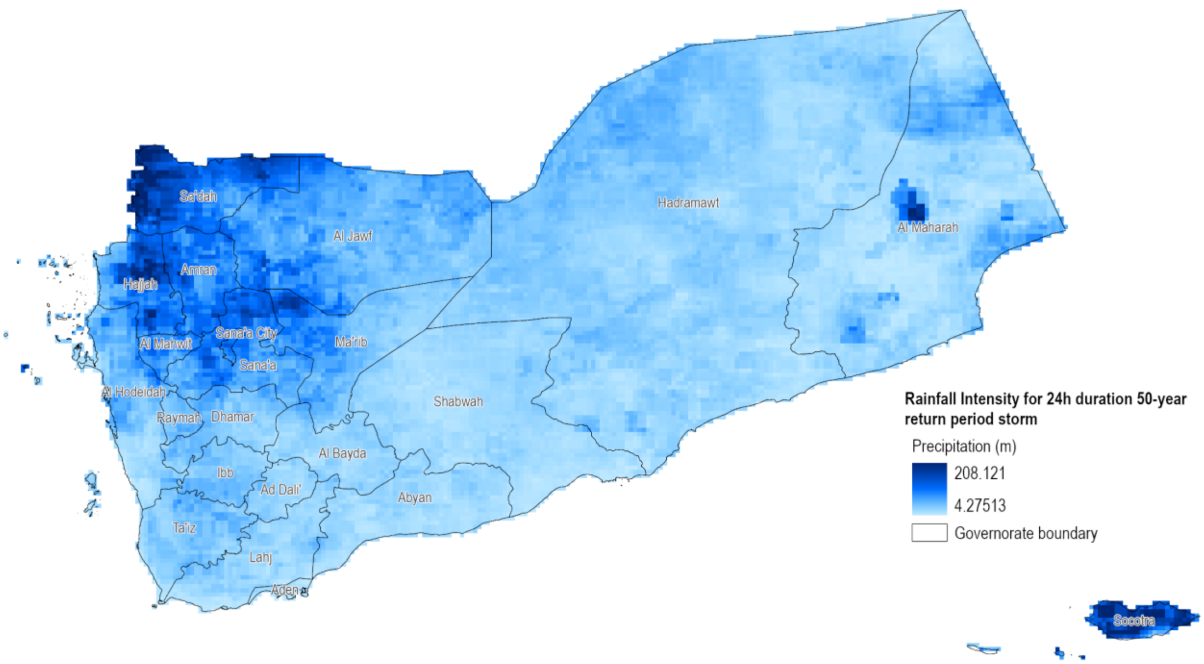
The GEV is a continuous probability distribution function which makes use of 3 parameters: location, scale and shape. Location describes the shift of the distribution on the horizontal axis, the scale provides an indication on how spread out the distribution is, and the shape indicates where the majority of the data lies, and the shape of the tails. The shape parameter derives from the skewness of the distribution (Millington et al., 2011).

Solving the distribution equation by 1 – 1/T, where T is the return period (Generalized Extreme Value distribution and calculation of Return value):

$$P_t = \xi + \left(\frac{a}{k}\right)\{1-(-\log\left(\frac{T-1}{T}\right)^k\}$$

Where P_t is the rainfall intensity, ξ is the location factor, a is the scale factor, k is the shape factor, and T is the return period (in years).

The return period was calculated by using Google Earth Engine for the CHIRPS Daily v2 dataset spanning from 1984 to 2019. (link to code)



Rainfall Intensity for 24h duration 50-year return period storm

Runoff Volume

Runoff is the amount of water that flows overland due to precipitation. It is strongly affected by impervious surfaces, such as urban areas, where there is no infiltration or soil absorption. However, also pervious surfaces can contribute, to a lower degree, to surface runoff. This depends on the soil's level of saturation, type of land cover and presence or absence of vegetation.

Runoff was calculated using the SCS Curve Number method, developed by the USDA NRCS (TR-55, 1986). It is based on a hydrological empirical parameter, the curve number, which uses the hydrological properties of soils and on land cover data as input data. The curve number is an estimation of the runoff potential and it contributes to the calculation of surface runoff depth, Q (in mm) using the following equation:

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

Wher P is precipitation, I_a is Initial abstraction (the amount of water that does not contribute to runoff through infiltration or evapotranspiration) and S is potential maximum retention after runoff begins.

According to the SCS method, it is possible to estimate the initial abstraction as follows:

$$Ia = 0.2S$$

Where

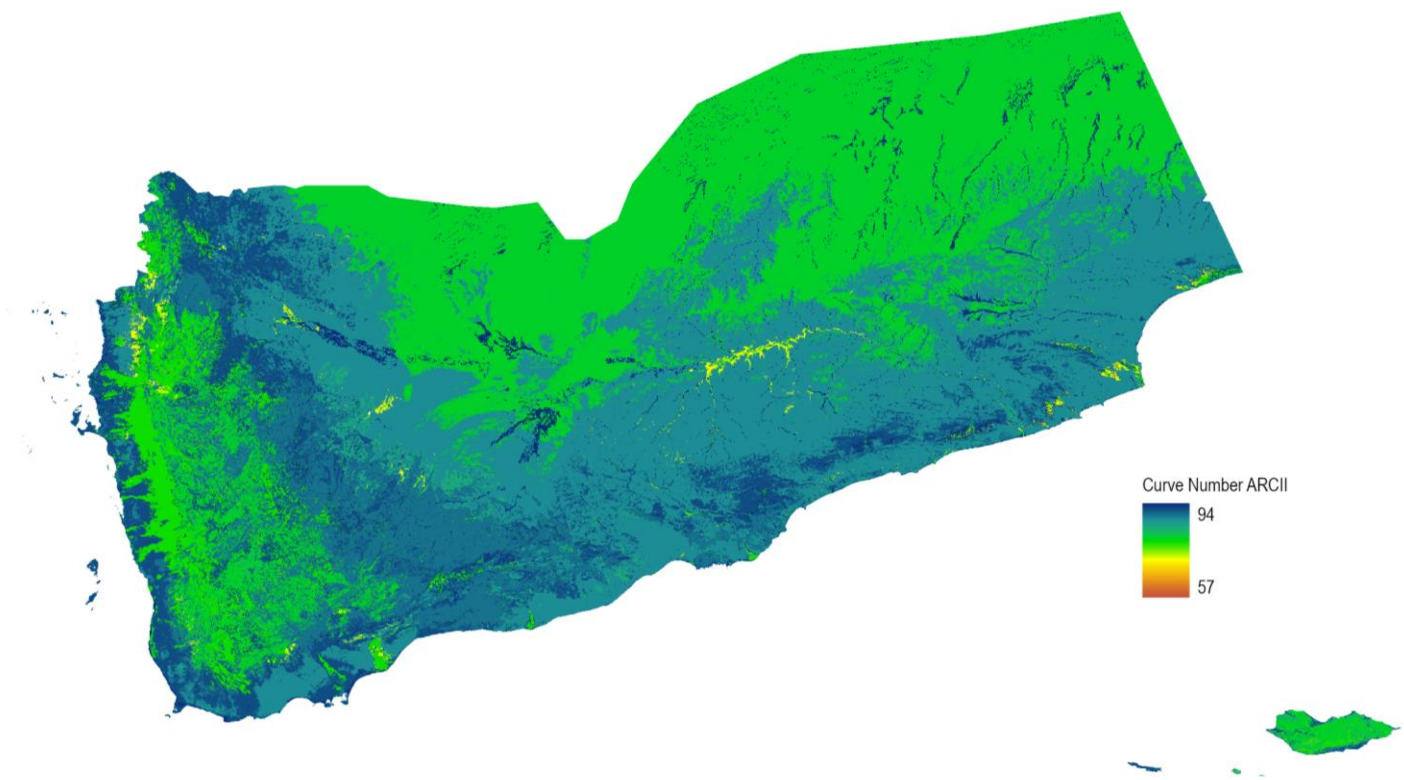
$$S = \frac{25400}{CN} - 254$$

Where CN is the calculated curve number.

In this analysis, the GCN250 global dataset was used (Jaafar et al., 2019), which provides in a gridded format the curve numbers on a global scale at a resolution of 250m. The precipitation data is the Rainfall intensity for a 50, 20 and 10 year return period previously calculated.

The Runoff volume was calculated then by multiplying the runoff depth with the pixel size. These procedures were all conducted using ArcGIS Pro.

It is important to note that in this analysis the Antecedent Moisture Conditions were assumed to be average (AMCII).



Curve Number with average moisture conditions (AMCII). (Jaafar et al. 2019)

Runoff Accumulation

This dataset indicates the flow paths the runoff generated by a storm will travel within the watershed boundary. This was calculated by generating a flow direction model based on the DEM. This latter evaluates the relative height of each cell compared to the surrounding ones and generates the preferred flow paths for surface water. The flow accumulation was generated using as weights the runoff volume previously generated. In this way every cell has the value of the runoff volume, and through the flow accumulation it is possible to understand how the cells are summed, based on the hydrological linkages determined by the flow direction. Therefore, each cell contains the total value of all upstream cells which flow along the same hydrological path. The Runoff Accumulation was generated using the Flow direction and Flow Accumulation tools on ArcGIS Pro. The input dataset was the 30 m resolution enhanced SRTM (NASADEM).

RESULTS

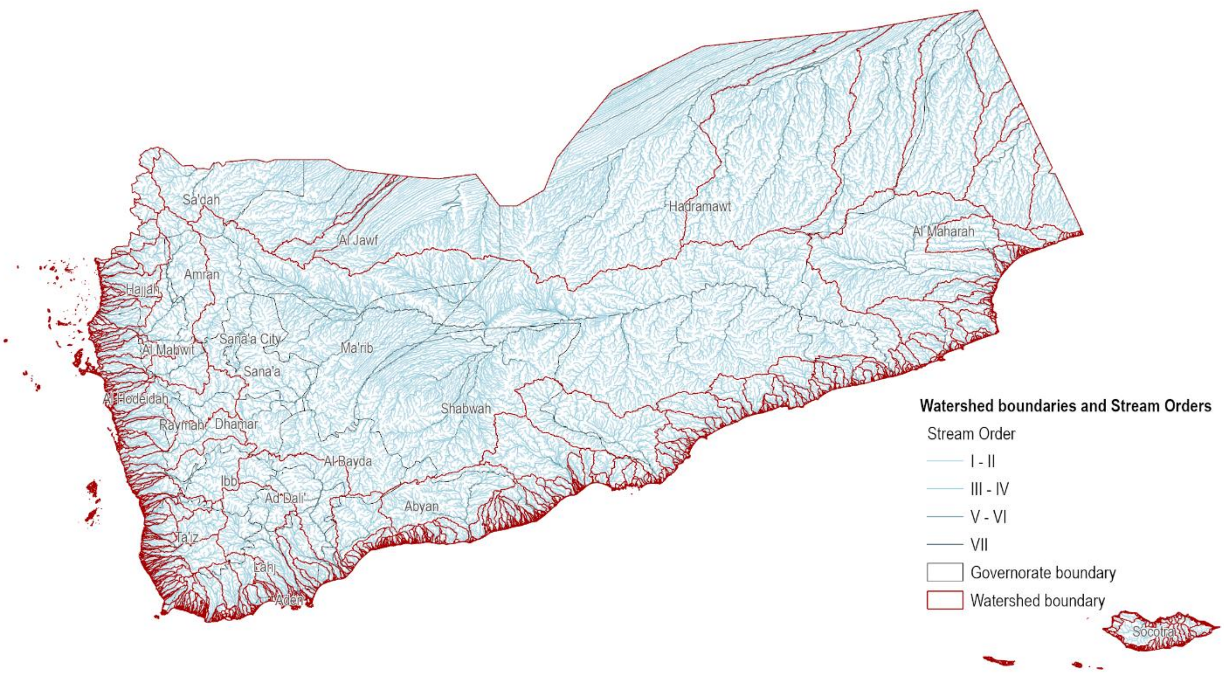
The results are expressed in three different type of products, maps depicting the watershed boundaries and the stream network, those illustrating runoff volume and runoff accumulation maps.

Watershed Boundaries and Stream network

This map illustrates the topographical features that influence the hydrology of Yemen. More specifically the runoff preferential flow paths (stream network) and catchment boundaries.

The stream and river network is classified in 7 orders. Streams with higher order have a larger contributing area within the watershed.

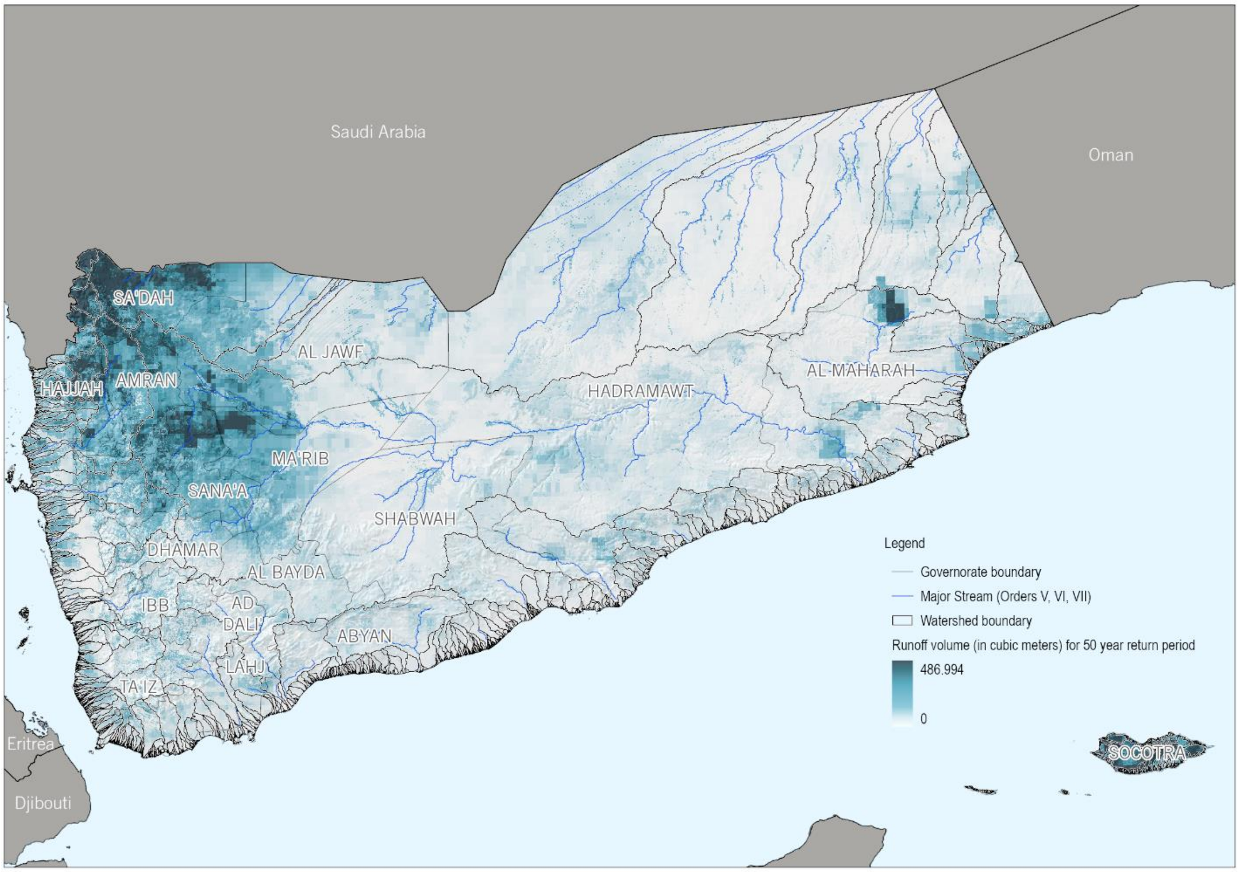
The watershed boundaries outline the drainage basins in the territory. Due to the method used (area-based) which considers the lowest topographical points for every watershed, the dataset does not describe sub-catchments, but only the country level watersheds.



Watershed boundaries and stream network, ordered according to the Strahler method

Runoff Volume

This map describes areas where surface runoff is generated due to precipitation and different grades of impervious surfaces. It indicates the volume of surface water the rainfall event produces in all areas of Yemen. It does not, however, provide any information of runoff routing or flow paths.

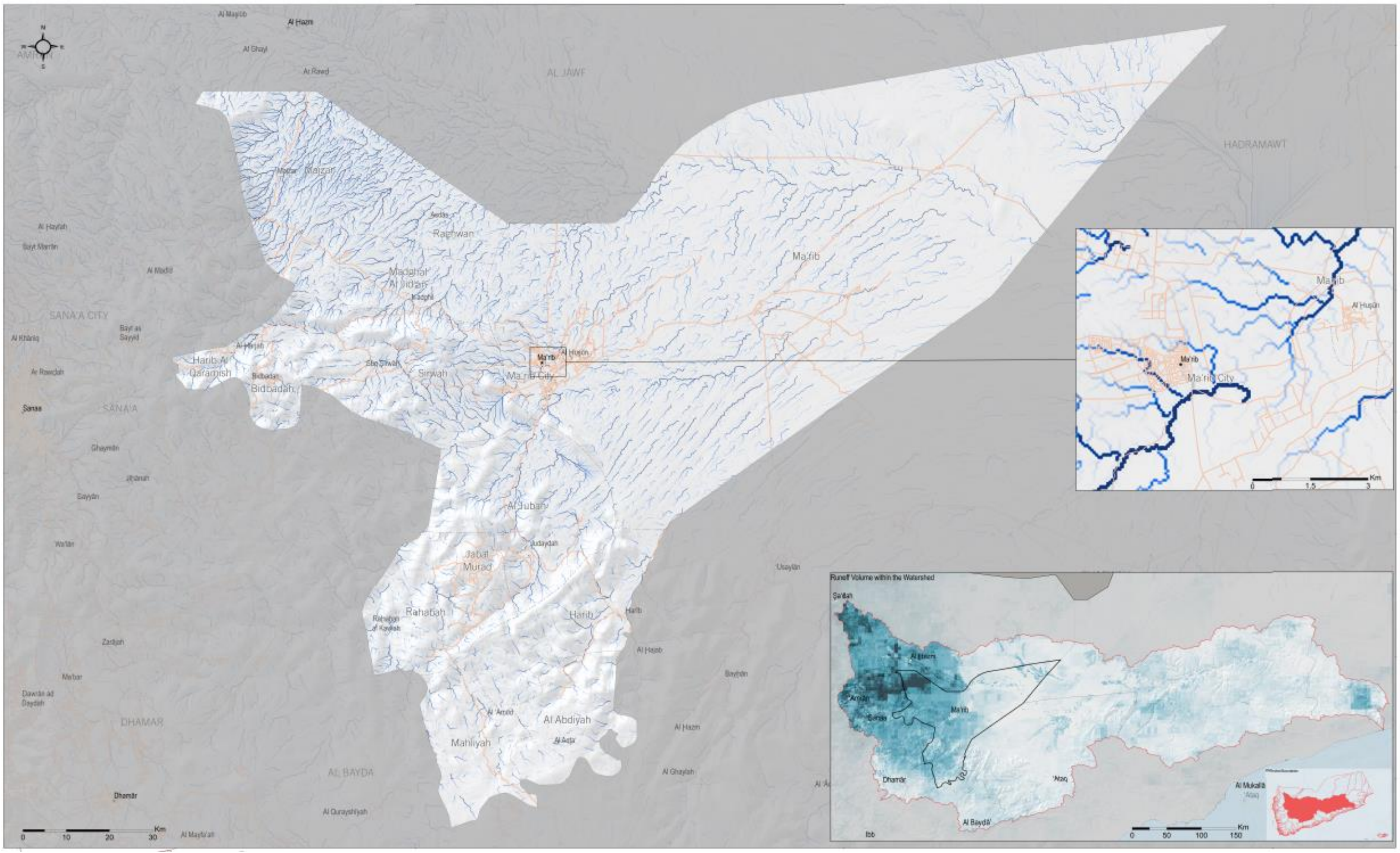


Runoff volume for a 50 year return-period rain storm

Runoff Accumulation

This dataset indicates the flow paths the runoff generated by a storm will travel within the watershed boundary. The values show in which channels the runoff volume accumulates downstream, based on the watershed and topographical data. This kind of product does not indicate the time it takes for the runoff water flow downhill, nor does it take into account how the water will flood out of the stream channels. It however gives an indication of the preferential routes runoff flows through and the distribution of the runoff volume within the watershed.

All runoff final products have the spatial resolution of 60 m.



Runoff Accumulation for a 50 year return-period rain storm

RECOMMENDATIONS

1. The current runoff model assumes that the soil moisture conditions prior to the storm event are average. It is recommended to better to evaluate the (average) antecedent moisture conditions in order to improve the quality of the SCS Curve number method (Brocca et al. 2009, Mishra and Singh, 2006). A possible starting point could be the integration of ERA Land dataset in the model, in order to assess soil moisture.
2. According to the approach used, it is possible to empirically estimate the initial abstraction as 20% of the maximum soil retention potential. However, through remote sensing data it is possible to estimate the average evapotranspiration and infiltration over Yemen.
3. The Curve Numbers used for this analysis are based on MODIS Land Cover dataset (~500 m resolution) and on HYSOG250 dataset for the hydrological soil properties (250 m resolution). The generation of a higher resolution land cover dataset, could be useful for recalibrating the curve numbers and improving the accuracy of the model.
4. This analysis does not take into account the concentration time, and therefore the lag time of overland runoff. Therefore, currently none of the products provides information on the time it takes for the surface water to flow downstream to a given area of interest. This is a function of land cover type and surface roughness, and could be calculated over specific areas of interest.
5. Currently the precipitation dataset used is the CHIRPS daily, however it would be better to use hourly or half hourly rainfall datasets, so as to be able to generate an accurate hyetograph and evaluate the duration of peak rain intensity with a finer temporal resolution. Satellite-based datasets such as JAXA's Global Satellite Mapping of Precipitation (GSMaP) dataset (Kubota et al. 2020) or NASA's Global Precipitation Measurement (GPM) v6 (Hoffman et al. 2018) could be leveraged for this model through Google Earth Engine.

REFERENCES

1. Clarke, R.T. (2006). Rainfall Trend Analysis: Return Period. In Encyclopedia of Hydrological Sciences (eds M.G. Anderson and J.J. McDonnell). doi:10.1002/0470848944.hsa040
2. Marra, F. and Morin, E. and Peleg, N. and Mei, Y. and Anagnostou, E. N. (2017). Intensity--duration--frequency curves from remote sensing rainfall estimates: comparing satellite and weather radar over the eastern Mediterranean. Hydrology and Earth System Sciences, 21, 5, pp. 2389-2404, doi:10.5194/hess-21-2389-2017
3. Millington, N., Das, S., & Simonovic, S.P. (2011). The Comparison of GEV, Log-Pearson Type 3 and Gumbel Distributions in the Upper Thames River Watershed under Global Climate Models.
4. USDA-SCS (1986) Urban Hydrology for Small Watersheds. Technical Release No. 55 (TR-55). USDASCS, Washington DC.
5. Jaafar, H.H., Ahmad, F.A. & El Beyrouthy, N. GCN250, new global gridded curve numbers for hydrologic modeling and design. *Sci Data* **6,** 145 (2019). https://doi.org/10.1038/s41597-019-0155-x
6. Brocca, Luca & Melone, F. & Moramarco, T. & Singh, Vijay. (2009). Assimilation of Observed Soil Moisture Data in Storm Rainfall-Runoff Modeling. Journal of Hydrologic Engineering. 14. 153-165. 10.1061/(ASCE)1084-0699(2009)14:2(153).
7. Mishra, S.K. and Singh, V.P. (2006), A relook at NEH-4 curve number data and antecedent moisture condition criteria. Hydrol. Process., 20: 2755-2768. doi:10.1002/hyp.6066
8. Huffman, G.J., E.F. Stocker, D.T. Bolvin, E.J. Nelkin, Jackson Tan (2019), GPM IMERG Final Precipitation L3 Half Hourly 0.1 degree x 0.1 degree V06, Greenbelt, MD, Goddard Earth Sciences Data and Information Services Center (GES DISC), 10.5067/GPM/IMERG/3B-HH/06
9. Kubota, T., K. Aonashi, T. Ushio, S. Shige, Y. N. Takayabu, M. Kachi, Y. Arai, T. Tashima, T. Masaki, N. Kawamoto, T. Mega, M. K. Yamamoto, A. Hamada, M. Yamaji, G. Liu and R. Oki 2020: Global Satellite Mapping of Precipitation (GSMaP) products in the GPM era, Satellite precipitation measurement, Springer, https://doi.org/10.1007/978-3-030-24568-9_20.

Donors

UKaid from the British people | Department for International Development

USAID FROM THE AMERICAN PEOPLE