Identification of Flood Risk Zones in San Marcos, Texas

Joyce Chien, Melanie Ratchford and Tasnuva S Udita

Department of Geography, Texas State University

Abstract: In the city of San Marcos, the costly threat of recent floods can cause serious damage to property as well as people. It is important to be able to locate and identify these areas at risk to help city planners as well as residents reduce damage and the potential loss of life caused by severe flooding. The "Flood Risk Tool" is a Python tool that attempts to identify the risk of flooding in the chosen study area. This Python tool uses ArcGIS and Arcpy to create a basic flood map identifying these areas using three factors. These three factors include proximity to rivers, rainfall and elevation which were determined through research and previous studies of flood risk mapping. This tool is able to create a map for any study area from a local scale or higher and also allows the user to set specific thresholds for each parameter based on their own research and requirements. Using these user defined parameters, the tool calculates the flood risk level based on a specific number value, showing areas from low to high risk. Finally, the tool is also able to save a copy of the resulting flood map both as an ". mxd" and PNG file.

1. Introduction:

1.1. Context of the Problem: San Marcos, a small city in Texas has long been recognized for its vulnerability to both flash floods and river floods. Some of the major underlying reasons behind this are: the number of low water crossings, heavy rainfall, the elevations and three of the most important water bodies; the San Marcos River, the Blanco River and Purgatory Creek. These disasters cause great havoc to lives and properties every year and this was evident in 2015, when two of the most historic floods occurred within only six months. The water level was 100%-300% above normal and the Blanco River water rose up to 40 feet (City of San Marcos 2016). That is why it is very important to identify the areas where the flood risks are higher and take appropriate measurements to avoid the loss.

1.2. Objectives of the Study: The major objective of the study is to identify the flood prone areas and assess the risks with proper justification. The purpose of this paper is to collect and present the appropriate data and information that might help people in building adequate management systems to mitigate the risks in future. The goals have been broken down into the points below:

- To use the python programming language in building a dynamic/ flexible tool for flood risk identification in an area.
- To develop an algorithm in calculating the risk levels based on the three major data inputs.
- To apply this method on the city of San Marcos and categorize it into zones based on risk level.
- Prepare a map of the potential affected areas in terms of area and population.

- To identify and focus more on the high-risk areas for policy makers to propose adequate management systems.
- To make the tool available for potential users to customize/ make modifications for further use.

1.3. Scientific Merit of the Study: The tool can contribute as a beginning point of developing more effective, complex and user friendly scripts. For now, it can be used for any spatial unit for example: city, county, state or country. However, it is considered more suitable to small scale areas like cities. Because the required data for assessing flood risk according to the tool within that scale are sometimes hard to find or might be unavailable, the tool can help in getting those from larger datasets. There is flexibility to assign weights to the parameters. Also as there are two separate functions under it, they can both be used individually. This tool can prove to be very useful to organizations and researchers working with flood management.

2. Literature Review:

To build this script tool, it is important to consider the possible parameters for identifying what areas might be at risk for flooding. Research on the area of study chosen for this analysis was also obtained. Examining information from various government websites, scientific journal articles as well as previous studies in the subject showed that, "Surface water flooding" has a greater chance of occurring in "low-lying areas" and at the bottoms of slopes and when discussing the role of rivers, according to "Flood Advice & Guidance," "Adjacent low-lying properties and land are then liable to be flooded" (Environment Agency 2013). As mentioned above, although the tool can create basic flood maps of almost any location at a chosen scale, the study area chosen for this analysis was the city of San Marcos which is located in Central Texas. As mentioned above, the city of San Marcos has faced recent floods. A map titled "Flood Occurrences in Texas," that was provided in the document shows Hays County experiencing 13-23 floods from 1960-2008 (Division of Emergency Management 2010). To make the final comparison of data, FEMA maps of the 100-year floodplain were obtained.

As mentioned in the "State of Texas Hazard Mitigation Plan," most of the state is at risk for flooding with flooding also accounting for 90% of damages caused by disasters (Division of Emergency Management 2010). While the document states flooding can happen at any time, seasonal rainfall attributes to floods occurring in winter and early spring while severe storms cause those in the summer (Division of Emergency Management 2010). The "State of Texas Hazard Mitigation Plan," asserts that "When runoff exceeds the carrying capacity of the stream or drainage, flooding occurs. Runoff is a product of two major groups of factors, climate and physiography" (Texas Department of Public Safety 2015). If a flood has occurred in an area, it is likely to reoccur; therefore, it is important to look at flooding history. When looking at the climographs provided, within the area that Hays County is located, the two rainiest points of the year are early summer in May and June and early fall in September and October with both times of the year receiving about four inches of rainfall (Texas Department of Public Safety 2015).

The Texas Floodplain Management Association document, "Quick Guide Floodplain Management in Texas", provides information on the floodplain in Texas (Texas Floodplain Management Association 2015). This document features Flood Insurance Rate Maps made by FEMA that display flood risk. The document describes high risk areas as, "High Risk Areas: All A and V Zones –the area that is located within the 1% annual chance floodplain (100-year floodplain) identified as Special Flood Hazard Areas on Flood Insurance Rate Maps" (Texas Floodplain Management Association 2015). It also describes, "Moderate to Low Risk Areas – Zones B (moderate), C and X (low) – area located outside the one-percent annual chance floodplain" (Texas Floodplain Management Association 2015).

In a previous study on the "Rhodope-Evros region" in Greece attempts were made to evaluate locations of flood hazards (Kazakis, et.al. 2015). Their study used seven parameters including, rainfall intensity, slope, flow accumulation, elevation, distance from drainage networks, land uses and geology (Kazakis, et.al. 2015). They weighed each parameter on "a scale between 2 and 10 (rating score)" and based these weights on prior research (Kazakis 2015). The researcher provides the revised index stating the elevation and the distance from the drainage network have the higher weights, while the lowest are assigned to rainfall intensity and the geology (N. Kazakis, et al. 2015). The article also states that this risk is lower the farther from the riverbed and areas that are less than 200 meters away are classified as being "highly flood hazard" (N. Kazakis, et al. 2015). In addition, the article states, "water flows from higher to lower elevations and therefore slope influences the amount of surface runoff and infiltration" (N. Kazakis, et al. 2015).

From previous study of "Urban flood hazard zoning in Tucumán Province, Argentine, using GIS and multicriteria decision analysis," the researchers also used rankings for their analysis on flooding and mentions that "distance to the discharge channels" has great importance in urban flood mapping (Fernández, D., and M. Lutz. 2010). The article states, "According to the records from the local authorities, the most affected areas during floods are those near these channels, as a consequence of overflow" (Fernández, D., and M. Lutz. 2010). The researchers use a "Multi-buffer operation" stating, "The distance intervals used were: <100 m, between 100 and 500 m, between 500 and 1000 m, and >1000 m" (Fernández, D., and M. Lutz. 2010). Thus, this study assigned distances from river highest weighted and elevation and slope high weighted than other parameters. For this study, their most significant variable was "the distance to the discharge channels," this was due to the fact that the surrounding area is often effected (Fernández, D., and M. Lutz., 2010).

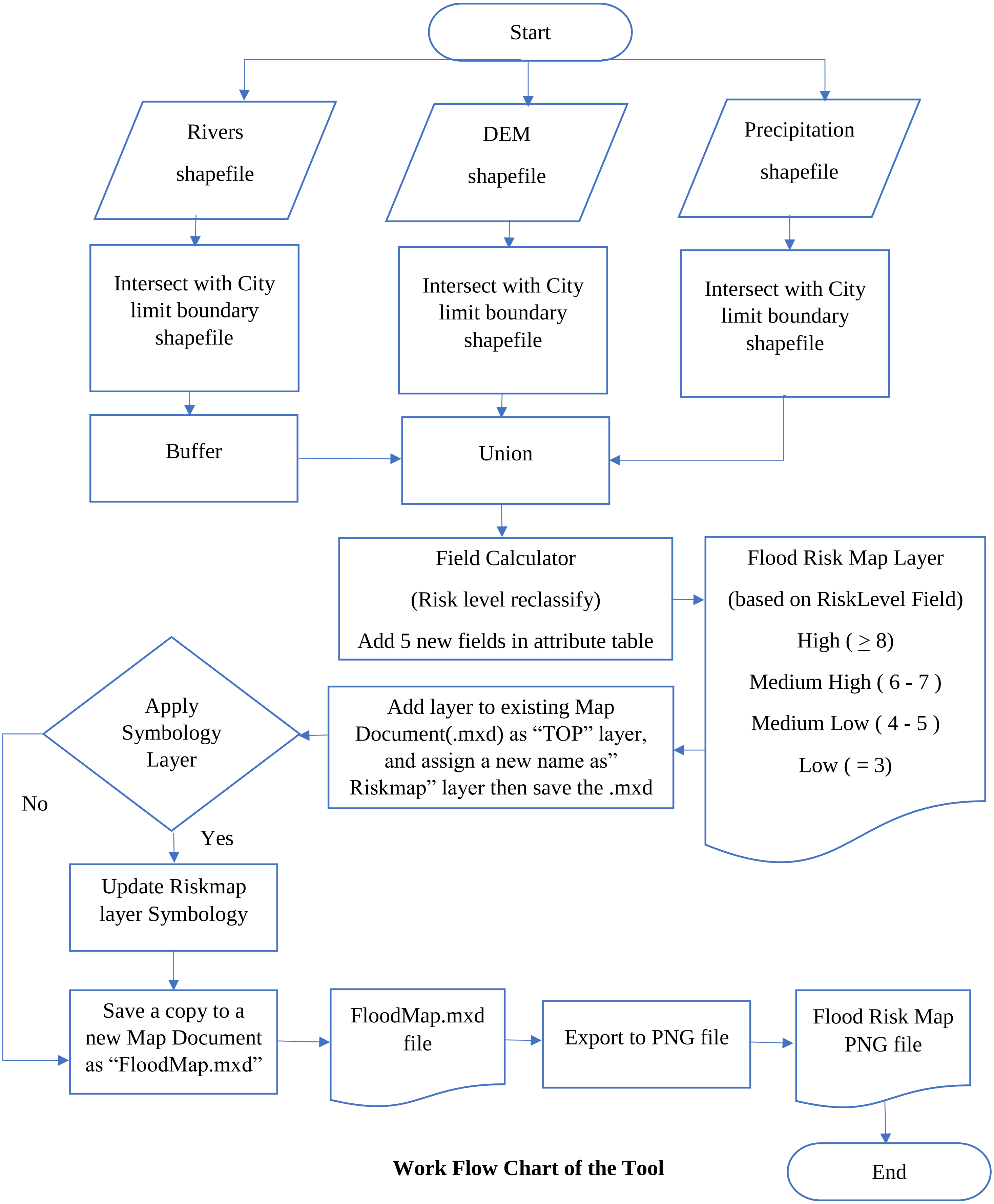
Based on the previous research, this study decides the three parameters which cause flooding zones that are most important are distance from rivers, elevation, and precipitation, and are assigned a weighted level from 1 to 3, where "1" signifies areas with low flood risk and areas with high flood risk are given a value of "3". Since the distance from rivers should have a higher weighted value than other parameters, the study assigns a different value where "1" signifies areas with low flood risk while areas with high flood risk are given a value of "5". In order to make the map easily understandable, a reclassification was performed to define four flood hazard intensity levels/categories as low, medium low, medium high and high. Unlike previous work this study uses the reclassify field by Field calculator instead of the weighted overlay method.

3. Methodology:

3.1. Data Used for the Study: Basically four major datasets have been used in developing the tool algorithm but they were in different file formats. Brief descriptions are given below:

- City boundary of San Marcos - It has been collected from the "City of San Marcos, Texas" official website (http://www.ci.san-marcos.tx.us/index.aspx?page=281). This was a polygon data found as a shapefile (.shp format). This boundary has been used for other datasets which were collected in a different spatial unit (state level).
- Precipitation data - This dataset was acquired for the whole Texas State first, from the NRCS (Natural Resources Conservation Service) of USDA (United States Department of Agriculture) official website (https://www.wcc.nrcs.usda.gov/products.html). The annual average precipitation data found for the most recent time was 1981- 2010. This was also a polygon data found as a shapefile (.shp format). It was pre- processed before using. This dataset was intersected with the San Marcos City boundary to keep all the records within the study area.
- Rivers data - This dataset was acquired for the whole Texas State first, from the TNRIS (Texas Natural Resources Information System) official website (https://tnris.org/data-download/#!/statewide). It was a polyline/ linear data in shapefile (.shp format). To get only the rivers for the study area, it has been intersected with the boundary just like the previous datasets. It generated a shapefile containing the rivers in San Marcos, but then only the major rivers were selected by using the "Select by Attributes" tool in ArcMap 10.4 version.
- Elevation data- The DEM (Digital Elevation Model) datasets have been collected from USGS (United States Geological Survey), Resources of the Geo Community and Digital Data Services online GIS-database (http://data.geocomm.com/dem/demdownload.html). There were six raster files for the Hays County. They were joined altogether using the "Mosaic to New Raster" tool from ArcMap 10.4 to create only one raster file, then clipped to get the record for only San Marcos City.
- Floodplain 100 year - The FEMA (Federal Emergency Management Agency) 100-year floodplain polygon shapefile has been collected within the study area boundary for evaluating the results from using the tool. It was collected from the "City of San Marcos, Texas" official website (http://www.ci.san-marcos.tx.us/index.aspx?page=281).

3.2. The Work Flow Chart of the Tool:



Work Flow Chart of the Tool

3.3. Main Components of the Tool**:** This tool uses two different scripts: The Flood Risk Data Analysis script prepares the data for the second script; The Flood Risk Identification script calculates the flood risk level and exports the final map to a PNG file.

- Component 1-1: Intersecting all the layers to the preferred study area. The user first defines this study area, for example, it can be a shapefile of city limits, state boundaries, counties and more. This also removes any data that might be unnecessary and outside the field of study but also preserves the data tables of the layers. Data for the three selected parameters is also imported to this tool. River data, however, is required to map the proximity to river.
- Component 1-2: Buffer is responsible for showing the proximity to rivers. This parameter allows the user to set a specific distance.
- Component 1-3: Union combines all of the data that has been processed.
- Component 2-1: Calculates the flood risk levels from low to high. These calculations are made based on the user's inputs for the thresholds of each of the parameters. The fields of the thresholds are selected first, then a number value is entered by the user into the thresholds. Using these numbers, the script calculates the levels of flood risk for each area. As mentioned earlier in the paper, the three parameters were chosen based on information from the literature review. Unlike the other two parameters which were assigned a maximum value of three, while proximity to rivers was given a value of five.
- Component 2-2: Export Map. This allows the user to save their results produced by the tool. Users must choose an existing .mxd file to save their flood map and also choose the name and location. The tool generates and saves a PNG formatted image of the final flood risk map.

3.4. Pseudocode of the Conceptual Framework:

Flood Risk Data Analysis Tool	Flood Risk Identification Tool

Import arcpy
Import the modules
Set the path
Set the environment and
Set the overwrite access
Check the analysis tools from extensions
Input the study area = 1st parameter
Input river data = 2nd parameter
Intersect 1 = 3rd parameter
Input second variable = 4th parameter
Intersect 2 = 5th parameter
	Get the user input of Union file from previous flood risk geoprocessing output
	Get the user input: DEM value field, threshold1, threshold2
	Automatically generate field "Risk1"
	Set the expression = """Reclassdem({}, {}, {}, 3, 2, 1)""".format(!DemValueFieldName!, threshold1, threshold2)
	Define the codeblock = """def Reclassdem(field "Risk1", threshold1, threshold2, x, y, z):
	Define the expression1, expression2, expression3 in the codeblock
	    if expression1:
	        return x
	    elif expression2:

Input third variable = 6th parameter Intersect 3 = 7th parameter Value Table1 = [study area, river] Value Table2 = [study area, second variable] Value Table3 = [study area, third variable] Intersect = [Value Table1, Intersect 1] Intersect = [Value Table2, Intersect 2] Intersect = [Value Table3, Intersect 3] Add messages Buffer Distance = 8th parameter Buffer Output = 9th parameter Buffer = [Intersect 1, Buffer Output, Buffer Distance] Add messages Union Output = 10th parameter Value Table = [Intersect 2, Intersect 3, Buffer Output] Union = [Value Table, Union Output] Add messages Check in extensions	return y **else**: return z Add Field in the attribute table Calculate Field of the Risk1 (Similar process repeated 5 times to generate and calculate 5 new field: Risk1, Risk2, Risk3, RiskSUM, RiskLevel) Get the user input: River Buffer value field, threshold1 (same process as blue codes) Get the user input: Precipitation value field, threshold1, threshold2 (same process as blue codes) MakeFeatureLayer to get Flood Risk Map Layer (based on RiskLevel Field) CopyFeatures Get the user input: existing Map Document (.mxd) Add layer to existing Map Document as “TOP” layer and save the .mxd file Get the user input: symbology layer **if** symbology = no input: go to save a copy of mxd directly **else**: Update Riskmap layer based on the Symbology layer Save a copy to a new Map Document as” Flood Map.mxd” Get the user input: PNG path and name Export To PNG del mxd

4. Analysis, Results and Findings:

4.1. Demonstrating the Tool with Real-World Data: Flood Risk Data Processing: As mentioned earlier in this paper that both the tools can be used individually, demonstrating with actual data collected for San Marcos City will help in explaining each parameter more –

Study Area Input: The 1st parameter (marked in red) is for the study area boundary, which depends on the user preferences. This study used the San Marcos City boundary. This input requires a feature class type.

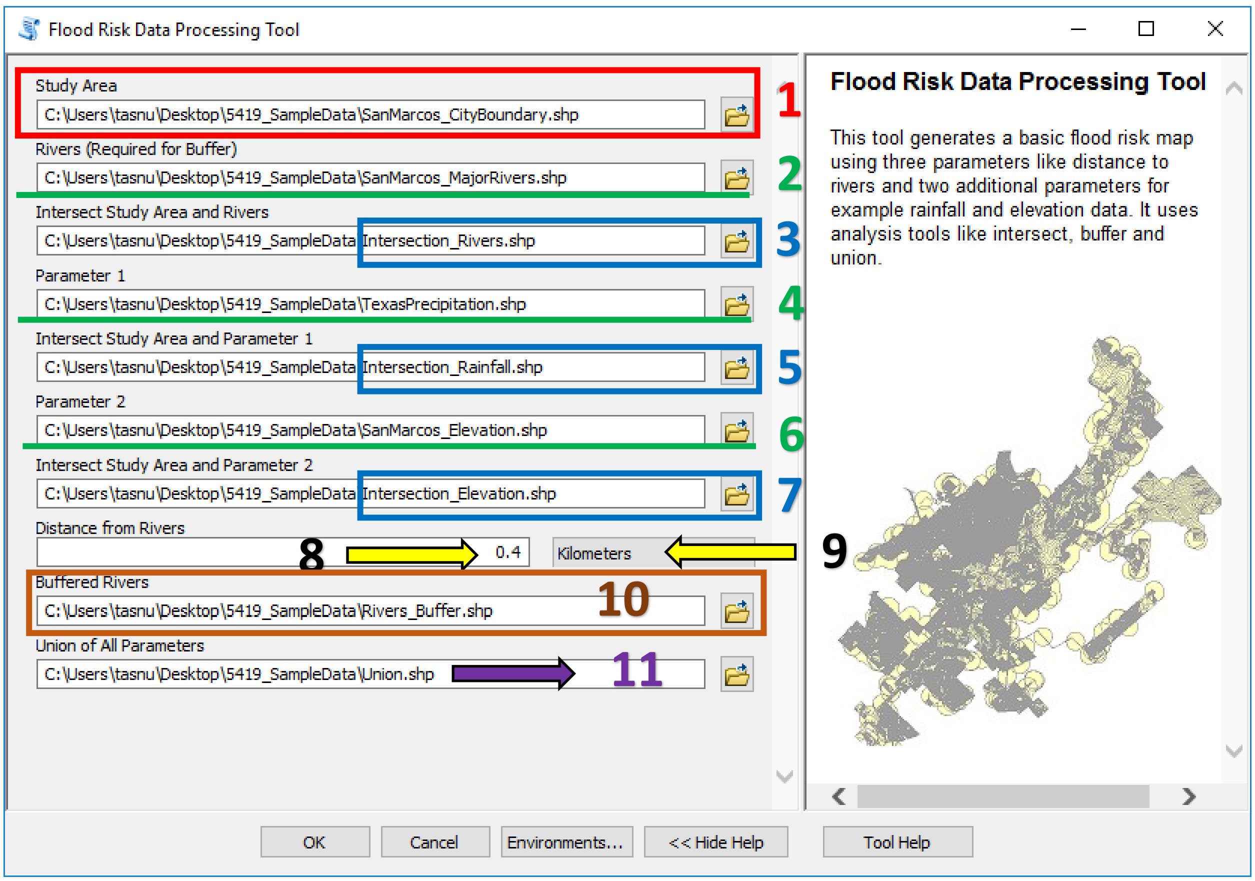
Variable Inputs: The 2nd, 4th and 6th parameters (marked in green) are for the major three variables which are rivers, rainfall and elevation for this study and they are going to be used for calculating final flood risk, intersected within the study area and all of them require feature class data.

Intersection Outputs: The 3rd, 5th and 7th parameters (marked in blue) ask for saving the intersections of the three variables used and they will generate three individual feature classes.

Proximity to River: The 8th and 9th parameters (yellow arrows) ask for a linear unit for creating the river buffer. The suitable buffer distance for this dataset used in this project is 0.4 km because San Marcos is a very small city. This will vary depending on the specific study area used.

Buffered Output: The 10th parameter (marked in brown) asks for saving the buffered river to the desired folder. This will generate a feature class.

Union Output: The 11th parameter (violet arrow) combines all the variables processed above as an output feature class to be saved in a folder preferred. And this is the last step for the first tool.



Input Feature: The 1st parameter (**marked in red**) requires a feature class of the study area with all the major variables data to identify risk levels. In this study that will be the union output.

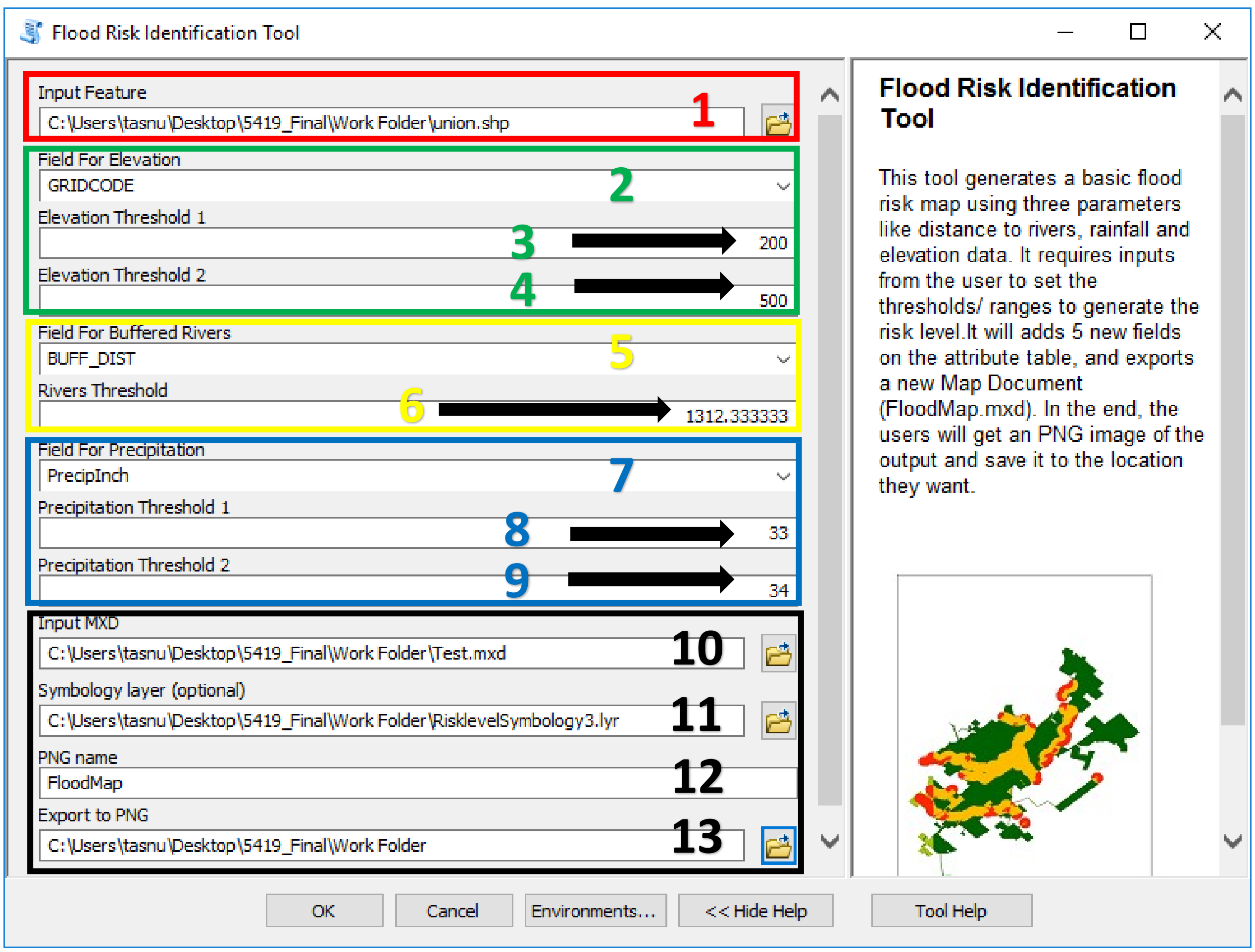
Field and Threshold Value for Elevation: The 2nd, 3rd and 4th parameters (marked in green) are for the fields that have the elevation, choosing the lower and upper limit from there. For this project, these limits have been set as 200 and 500 accordingly based on the “GRIDCODE” field.

Field and Threshold Value for Buffered Rivers: The 5th and 6th parameters (marked in yellow) are for that buff. In here that field is "BUFF_DIST" and distance is 1312.333333 km.

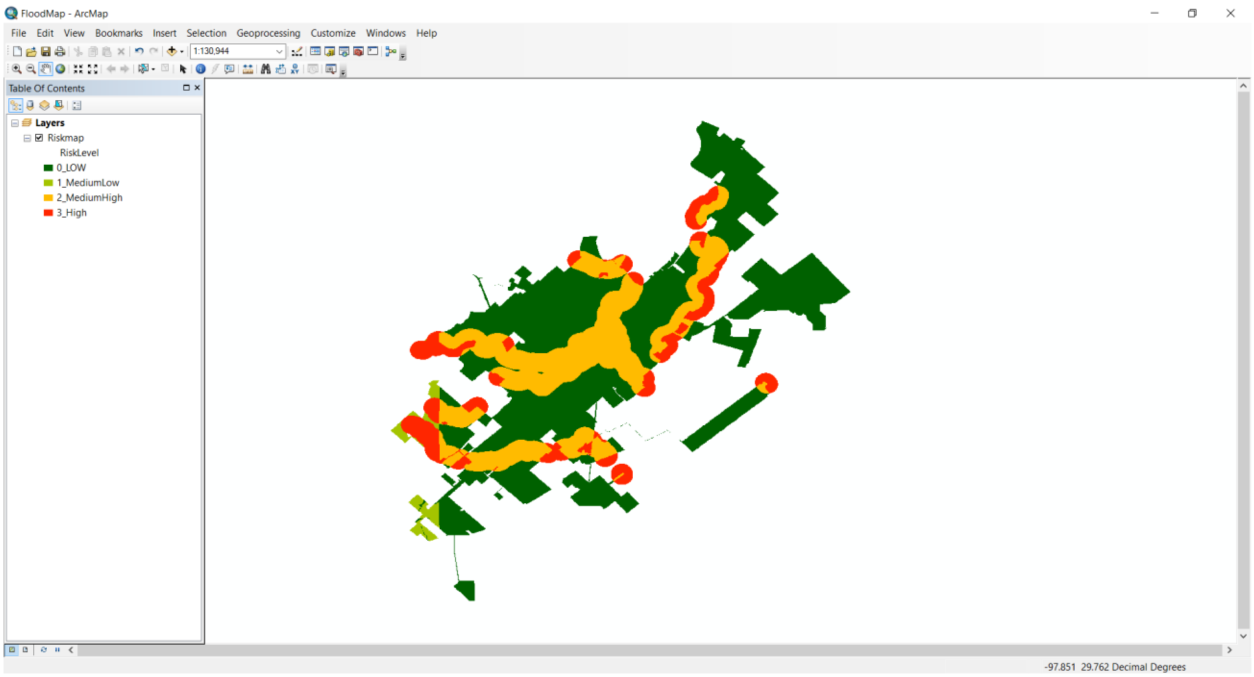
Field and Threshold Value for Precipitation Data: The 7th, 8th and 9th parameters (marked in blue) require the field storing rainfall data, the lower and upper limits which are 33 and 34 for this study based on "PrecipInch" field.

Exporting the Flood Risk Map: The last four parameters (10, 11, 12, 13 marked in black) are to generate the map in ArcMap and export into an image. For that, the user should have an existing empty ". mxd" file to include. For the next step, the user can input an existing symbology layer or leave that default because it is an optional field. The last two steps exports the map into a PNG. The result for this study is shown in the sidebar of the tool description.

Though the fields for thresholds have been defined in here, the user can include any other fields and determine the thresholds accordingly as long as they have those data. But the field for river is a required one for this tool.

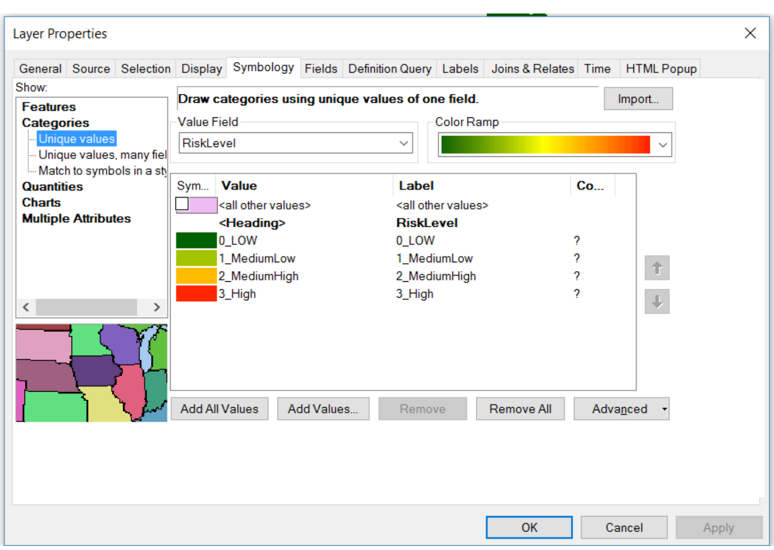


4.2. Results and Maps: Picture 1 shows the output in map document file (. mxd format). This map has been created based on the existing symbology layer. The categories shown are based on the calculated field "Risk Level" where dark green represents Low risk, light green as Medium Low, orange as Medium High and red as High risk level.



Picture 1. The output Map Document file (FloodMap.mxd)

In the next page, picture 2 shows the symbology of the final flood risk map. The user has the flexibility to change the symbology in the Map Document as preferred.



Picture 2. The symbology of the Output Feature (Riskmap)

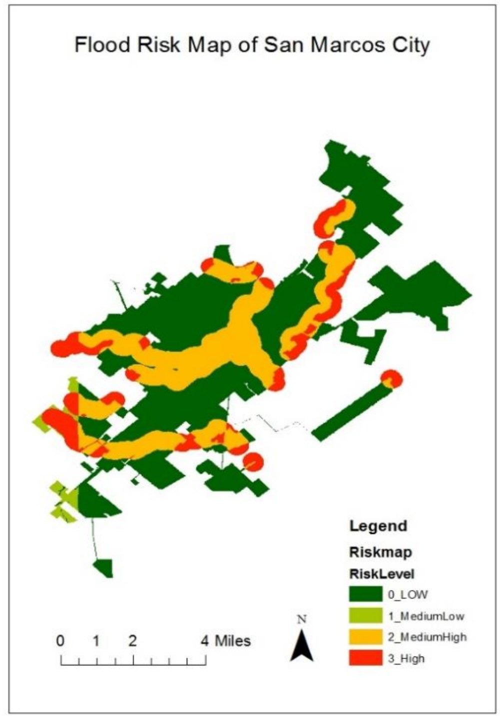
Risk1	Risk2	Risk3	RiskSUM	RiskLevel
3	5	3	11	3_High
3	5	3	11	3_High
3	5	3	11	3_High
3	5	3	11	3_High
3	5	3	11	3_High
3	5	3	11	3_High
3	5	3	11	3_High
3	5	3	11	3_High
3	5	3	11	3_High
3	5	3	11	3_High
3	5	3	11	3_High
3	5	3	11	3_High
3	5	3	11	3_High
3	5	3	11	3_High
3	5	3	11	3_High
3	5	3	11	3_High
3	5	3	11	3_High
3	5	3	11	3_High
3	5	3	11	3_High
3	5	3	11	3_High
3	5	3	11	3_High

Picture 3. Attribute table of the new fields

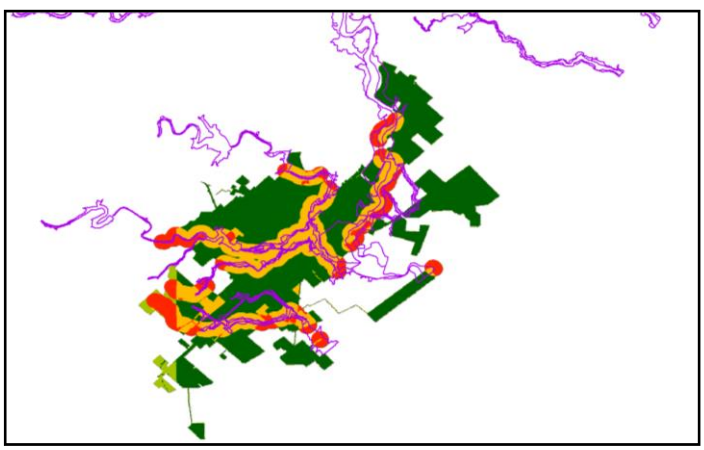
Picture 3 shows the 5 new fields added by the tool in the attribute table: Risk1 (Elevation), Risk2 (Buffer Distance), Risk3 (Precipitation), RiskSUM and RiskLevel (Based on the ranges described in figure). RiskSUM field is calculated with this formula: Risk1 + Risk2 + Risk3, and RiskLevel field is created based on the RiskSUM as Text data type where: Low = 3, Medium Low = 4 -5, Medium High = 6 -7, and High $\geq$ 8.

So basically in the end, the user will get a new output of feature layer named “Riskmap.shp”, 5 new fields added in the attribute table, an output in the new Map Document named “FloodMap.mxd”, and a .PNG file of that map. Other map elements can be added for a different representation of the map which depends solely on the user (Picture 4).

After comparing the final output with the 100-year flood plain data from FEMA, we also validate our result and get the positive answer of our output. The higher risk level will happen near the river, thus buffer from river should be assigned the higher value (Picture 5).



Picture 4. The Final Flood Risk Map with Key Map Elements



Picture 5: The final Flood Risk Map compared with FEMA 100-yr flood plain

4.3. Findings and Implications: The key findings and implications of this study are given below:

- Based on the three values the map shows the areas with low risk appear far from the river, where the elevation is high. The precipitation does not differ much for this dataset.
- Compared to the literature review, it appeared that the area of San Marcos which is most at risk for flooding is the Eastern, and Southern portions of the city close to the rivers.
- Particularly, when comparing the basic flood maps created by this tool and the 100-year floodplain maps created by FEMA of the study area, similar patterns of risk can be seen.

According to the "City of San Marcos Action Plan for Disaster Recovery," during the 2015 floods, "Most of the damage occurred within the 100-year floodplain or right up against the banks of the Blanco and San Marcos Rivers" (The City of San Marcos 2016).

- Because it only requires three parameters to create the final flood map, this tool can prove to be useful when there is little data available to the user.

5. **Conclusion:** The points described below can be considered as limitations or contributions based on the user data availability:

- The tool can take only three variables/ parameters for input. The user can input other parameters instead of these as long as they have the data already in the attribute table, also relate to the purpose and correspond with the execution of the tool.
- The tool works with only vector data but not raster data. Though the first part of the tool can work with both line and polygon data, the second part considers only polygon data (already processed either by the first part of the tool or in some other ways).
- The rivers can represent only one buffered distance given by the user and not multiples.

To overcome the restrictions of the tool, future modifications are recommended that may include more than three parameters for risk identification, change the range of the risk levels and names to increase the flexibility for users. Flood maps can help residents as well as the city plan for potential flooding and allow them to save money caused by damage to homes and buildings.

References:

City of San Marcos Action Plan for Disaster Recovery, (2016). Disaster Relief Appropriations Act. Public Law 114-113.

Division of Emergency Management, 2010. State of Texas Hazard Mitigation Plan. Texas Department of Public Safety. Available at https://www.dps.texas.gov/dem/documents/txHazMitPlan. pdf.

Environment Agency, 2013. FLOOD ADVICE & GUIDANCE. Know Your Flood Risk. Available at http://www.knowyourfloodrisk.co.uk/flood-advice-guidance.

Fernández, D., and M. Lutz. 2010. Urban flood hazard zoning in Tucumán Province, Argentina, using GIS and multicriteria decision analysis. Engineering Geology 111 (1-4):90–98.

Kazakis, N., I. Kougias, and T. Patsialis. 2015. Assessment of flood hazard areas at a regional scale using an index based approach and Analytical Hierarchy Process: Application in Rhodope-Evros region, Greece. Science of the Total Environment 538: 555-563.

Texas Floodplain Management Association. 2015. Quick Guide Floodplain Management in Texas. Available at http://www.twdb.texas.gov/flood/resources/doc/Texas_Quick_Guide. pdf.