IMAGERY, AUTOMATION, AND APPLICATIONS FINAL ASSIGNMENT

Coursera's GIS Specialization (4/5 course)

PROJECT: Flooding of the Mississippi River in St. Louis City (Missouri, United States)

Fernando Lopez Blanco, Feb. 10, 2020.

Abstract

As an analyst working for the United Nations, this project evaluates the extent of flooding of the Mississippi River across St. Louis City in May 7, 2019. An event that made the Earth Observatory Image of the Day in May 10, 2019 (https://earthobservatory.nasa.gov/images/145029/flooding-continues-along-the-mississippi). We test the hypothesis that the flooded surface of St. Louis City neighborhoods along the riverbank was larger than the flooded surface of just St. Louis City Centre (Downtown), which also sits by the river. To find out, the project involves two major parts: an initial image acquisition from Landsat 8 satellite and classification, followed by an analysis of the derived products to get the flooded areas along the rivebank.

1. Image Acquisition and Classification

To put the hypothesis to the test, we first need to get satellite imagery for both the normal river extent and the flooding extent. Landsat 8 Collection 1 (C1) Analysis Ready Data (ARD) at 30-meter pixel resolution series (OLI-TIRS sensor) are a good choice because they are images ready for direct use in sensing applications, being pre-processed from raw data to high standards. Namely, the Surface Reflectance by-product is perfect for our purposes. After creating an account with USGS, the two products shown in Table 1 were downloaded from the USGS Earth Explorer website (https://earthexplorer.usgs.gov/). Both images come with 7 bands covering different parts of the electromagnetic spectrum in the visible and infrared ranges. For this analysis, bands 2 (Blue), 5 (NIR) and 6 (SWIR1) were used as Red, Green and Blue filters respectively. The resulting composite images are also shown in Table 1.

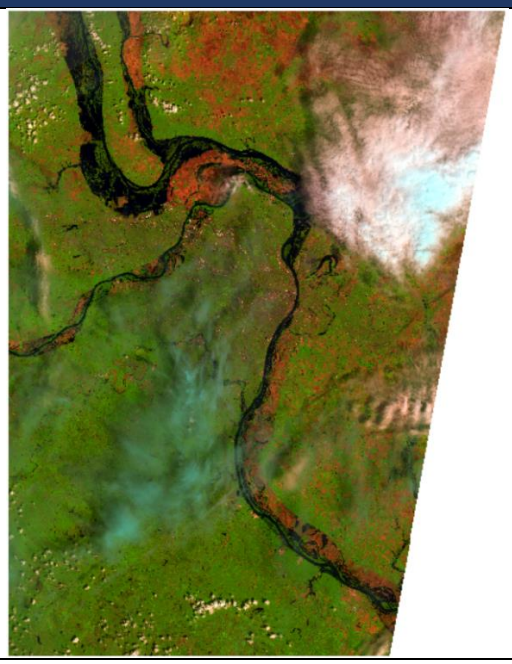
IMAGE OF THE FLOODING	
Tile **ID:** LC08_CU_020010_20190507_20190523_C01_V01 **Acquisition Date:** 07-MAY-19 **Horizontal:** 20 **Vertical:** 10	
Product – Surface Reflectance LC08_CU_020010_20190507_20190523_C01_V01_SR.tar Pixel resolution: 30 meters *Right image*: composite of bands 2 (Blue), 5 (NIR) and 6 (SWIR) of the area round St. Louis City (at centre of the image) to enhance agriculture areas.	

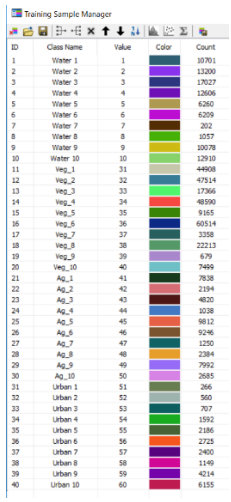
IMAGE FOR COMPARISON	
Tile **ID:** LC08_CU_020010_20181011_20190614_C01_V01 **Acquisition Date:** 11-OCT-18 **Horizontal:** 20 **Vertical:** 10	
Product – Surface Reflectance LC08_CU_020010_20181011_20190614_C01_V01_SR.tar Pixel resolution: 30 meters **Criteria**: *Cloud cover*: less than 20% *Cloud shadow*: less than 20% *Right image*: composite of bands 2 (Blue), 5 (NIR) and 6 (SWIR) of the area round St. Louis City (at centre of the image) to enhance agriculture areas.	

Table 1 – Data on the products downloaded from USGS website for flooding and base comparison images, which are also shown in the right column after compositing them with bands 2, 5 and 6 with ArcGIS Composite Bands Tool.

Even though the image of the flooding shows a good patch of clouds, fortunately they are mostly out of the studied area of St. Louis City. The image for comparison with normal river and water extent was searched as having as least noise as possible (i.e. additional search criteria parameters shown in Table 1) and as close in time as to the flooding event as possible. Mid-autumn, after the summer drought and when some rain is supposed to have already fallen, seemed a good choice for this purpose. Thus, the date of Oct. 11, 2018 was finally selected for comparison.

1.1. Image Classification

To start with, we trained and finally created 40 training polygons on the Oct. 11, 2018 image for the 4 different classes as shown in Picture 1. These classes are Water, Vegetation, Agriculture and Urban areas. We used ArcGIS Image Classification Toolbar for this purpose.



Training Sample Manager

ID	Class Name	Value	Color	Count
1	Water 1	1		10701
2	Water 2	2		13200
3	Water 3	3		17027
4	Water 4	4		12606
5	Water 5	5		6260
6	Water 6	6		6209
7	Water 7	7		202
8	Water 8	8		1057
9	Water 9	9		10078
10	Water 10	10		12910
11	Veg_1	31		44908
12	Veg_2	32		47514
13	Veg_3	33		17366
14	Veg_4	34		48390
15	Veg_5	35		9165
16	Veg_6	36		60514
17	Veg_7	37		3358
18	Veg_8	38		22213
19	Veg_9	39		679
20	Veg_10	40		7499
21	Ag_1	41		7838
22	Ag_2	42		2194
23	Ag_3	43		4820
24	Ag_4	44		1038
25	Ag_5	45		9812
26	Ag_6	46		9246
27	Ag_7	47		1250
28	Ag_8	48		2384
29	Ag_9	49		7992
30	Ag_10	50		2685
31	Urban 1	51		266
32	Urban 2	52		560
33	Urban 3	53		707
34	Urban 4	54		1592
35	Urban 5	55		2186
36	Urban 6	56		2723
37	Urban 7	57		2400
38	Urban 8	58		1149
39	Urban 9	59		4214
40	Urban 10	60		6155

Image 1 - Training Sample Manager screenshot showing all 40 polygons to classify St. Louis base image.

Then, we ran a Maximum Likelihood supervised Classification that yielded the result shown in Image 2 left, after the 40 polygons were merged into 4 classes and these were assigned a more natural color. Though it overall seems like a good match to the satellite image, the urban areas are overestimated after the white clouds and their shadows (right and bottom of the image purple patches). However, the urban area around St. Louis is fairly well estimated and even both radial and ring roads can be seen. The water extent (rivers, streams, dams and ponds) seems also reasonably well estimated and thus, we extracted just that class for using it in our subsequent analysis (Image 2 right) by using the Extract by Attributes Tool.

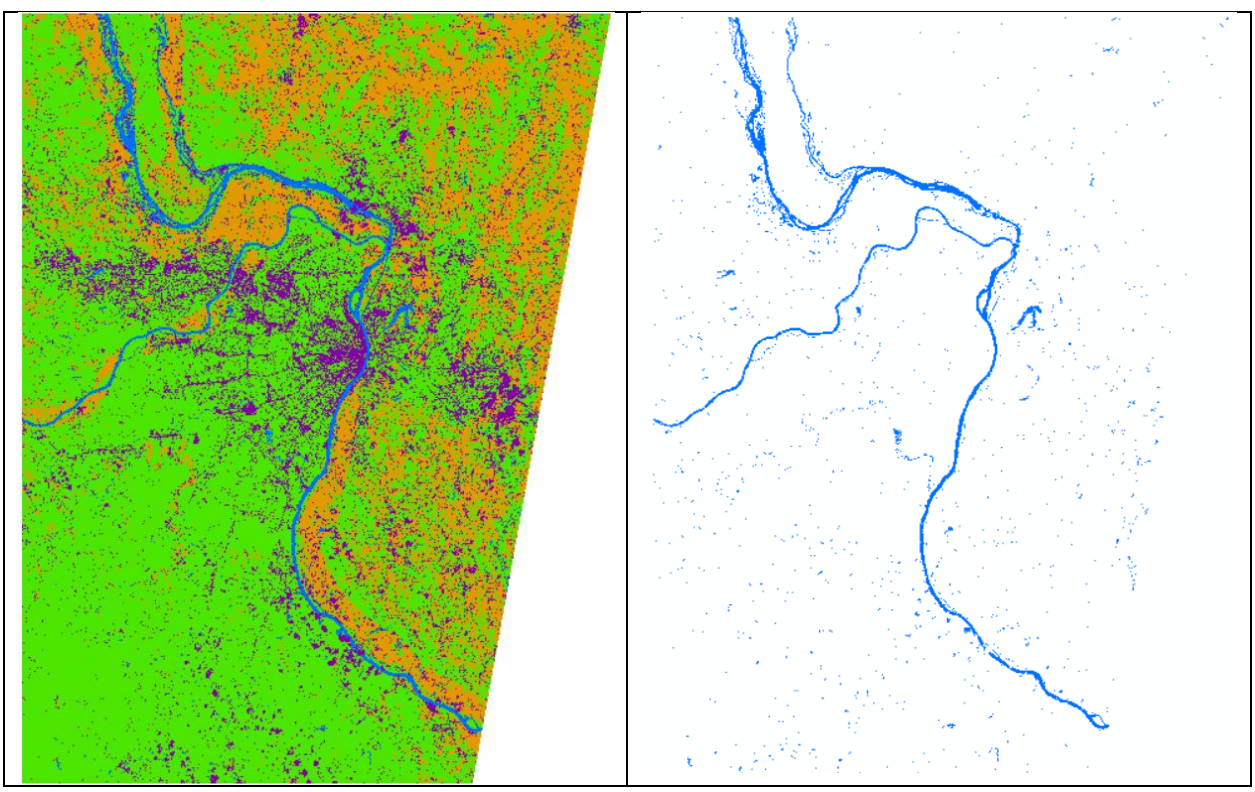


Image 2 – Left: Maximum Likelihood Classification output when ran with the above-mentioned training polygons. Water is blue, Agriculture is shown as orange, Vegetation as green and Urban areas as purple. *Right*: Water extent raster from left image.

Analogously, we created training polygons and ran a Maximum Likelihood Classification on the image of the flooding May 7, 2019. The water extent raster is shown in Image 3. Note the lack of data in the top-right corner where the clouds are on the original image.

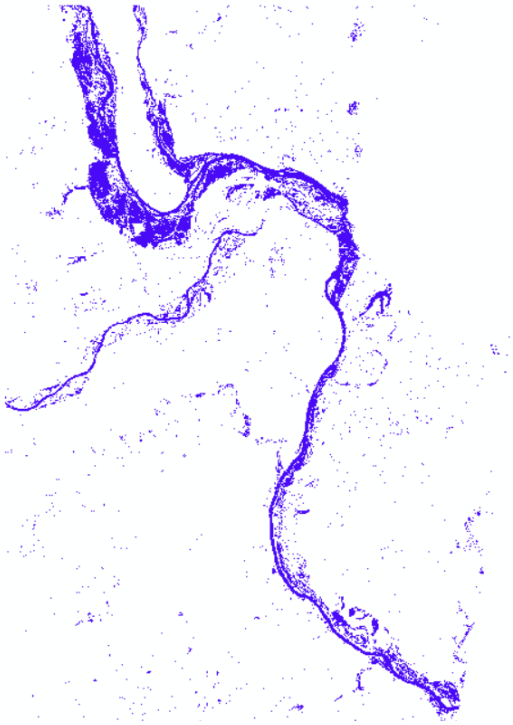
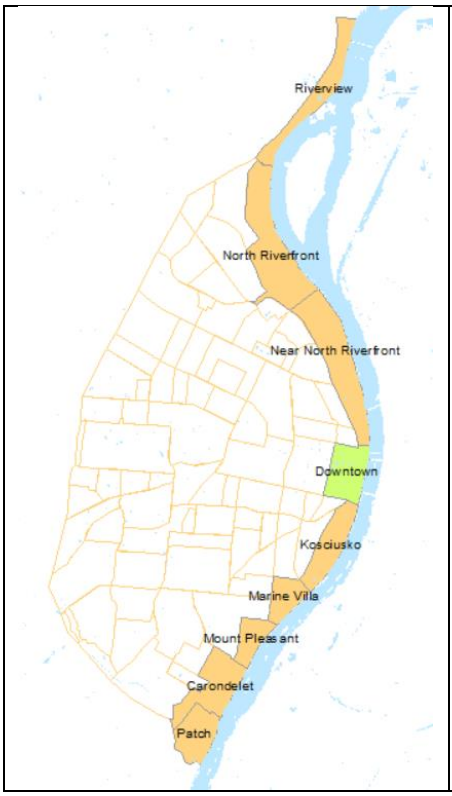


Image 3 –Water extent raster for the flooding image.

1.2. St. Louis City Neighborhoods

We need now the neighborhoods in St. Louis City to extract those along the riverbank, so we can test our hypothesis against. The St. Louis Government page provides this data, along with Wards in a shapefile (https://www.arcgis.com/home/item.html?id=1dff446cfb4e4db59a6c0adff2ab7b67). By manually selecting the Downtown neighborhhood and creating another layer from the selection, we got the area for the city centre. Analogously, we created another layer for the neighborhoods along the Mississippi River and dissolved them to get just a single feature (Table 2 left). From both features, Downtown and Riverside neighborhoods, we can get their total surface in square meters (since Meters is the linear unit used in the rasters) (Table 2 right).



Riverside neighborhoods (orange-shaded areas):

1- Carondelet
2- Patch
17- Mount Pleasant
18- Marine Villa
20 - Kosciusko
64- Near North Riverfront
75- Riverview
79- North Riverfront

Total Area (*TRS*): 44,399,620.255 sq.m

City Centre (green-shaded area):

35- Downtown

Total Area (*TDS*): 4,390,560 sq.m

Table 2 – Map of St. Louis City Neighborhoods outlined. Those along the river are orange-filled, whereas city centre is green-filled. Total surfaces are also shown. Left-hand side number along the neighborhood name refers to their internal County code number (NHD_NUM).

2. Image Analysis

Our hypothesis can be written with the following equation:

$$\frac{Flooded\ Riverside\ Surface\ (FRS)}{Total\ Riverside\ Surface\ (TRS)} > \frac{Flooded\ Downtown\ Surface\ (FDS)}{Total\ Downtown\ Surface\ (TDS)}$$

2.1. Model Builder

We already have both denominators. We need to figure out both numerators. To do that, we created a Model to calculate the flooding extension in a given area by using the Model Builder in ArcGIS (Image 4).

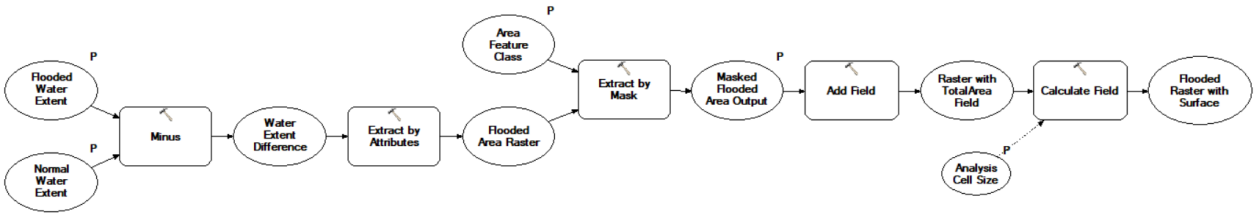


Image 4 –Calculate Flooding Extension Model created within a brand-new Flooding Analysis Toolbox.

The Model has the following input parameters (Image 5):

- Normal Water Extent Raster: this would be the one used as comparison (Oct. 11, 2018).
- Flooded Water Extent Raster: this would be the one with the flooding extent (May. 7, 2019).
- Area Feature Class: this would be a polygon layer (Downtown or Riverside).
- Masked Flooded Area Output Raster: this would be the extent of flooding in the given area.
- Analysis Cell Size: the raster's pixel resolution (30 meters).

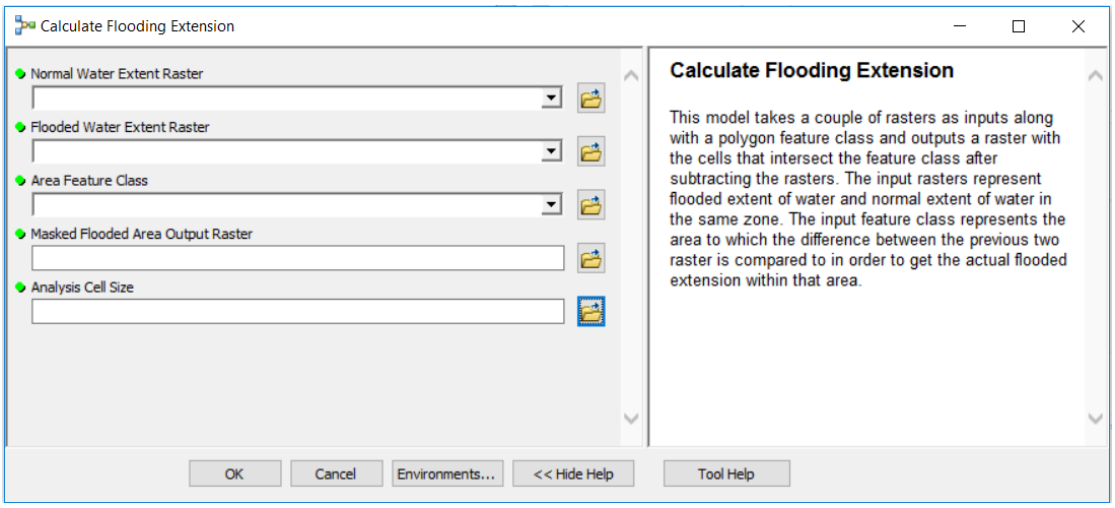


Image 5 –Calculate Flooding Extension tool inputs window.

The Model goes through five different tools to cover the following algorithm:

1) Calculate the difference between input rasters. This was achieved by using the Minus tool. The output raster's cells have 3 possible values: -1, 0 and 1.
2) Get the actual flooded extent from previous raster. This was achieved by using the Extract by Attributes tool with expression 'value = 1' to effectively get the pixels that are in the flooded raster but not in the base comparison raster.
3) Get the pixels that overlap the given polygon area. This was achieved by using the Extract by Mask tool, so when, for instance, the Downtown feature class is passed on as parameter, the output raster contains just the cells of flooding in that area.
4) Finally, to automatically calculate the surfaced covered by water, we added a field to the previous raster's table (Add Field tool) and populated it (Calculate Field tool) with expression '[COUNT] x %Analysis cell size% x %Analysis cell size%' to get a surface in square meters.

2.2. Model Analysis

When both the Downtown and Riverside Neighborhoods feature classes are respectively passed into the model, alongside the water extent rasters we had after image classification, we obtained the results shown in Table 3.

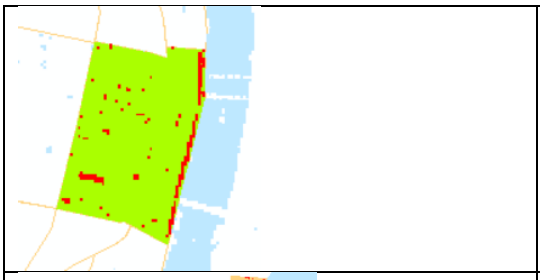
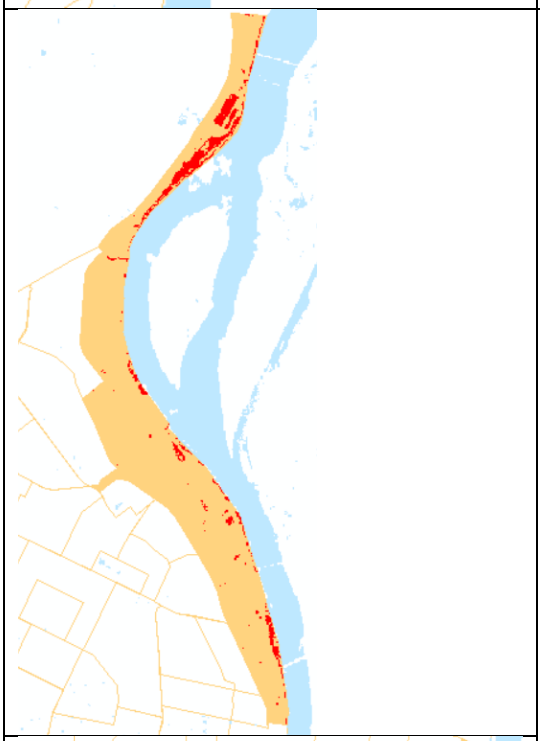
	Flooding in Downtown shown as red pixels. Pixel count: 179 **Total Area (*FDS*)**: 161,100 sq.m
	Flooding in riverside neighborhoods shown as red pixels (north and south of downtown) Pixel count: 1,596 **Total Area (*FRS*)**: 1,436,400 sq.m

Table 3 – Flooded extent rasters for both Downtown and Riverside Neighborhoods (north and south of Downtown) after being processed by our model. Total flooded surfaces are also shown.

Plugging in now all 4 surface figures into our equation, this yields:

$$\frac{1{,}436{,}400}{44{,}399{,}620.255} > \frac{161{,}100}{4{,}390{,}560};\ 0.0323 > 0.0366$$

3. Conclusions

The hypothesis we tested is not true, since according to the previous result, the surface flooded in Downtown during the May. 7 2019 event was slightly larger than the surface flooded in riverside neighborhoods: a 3.66% vs. 3.23% of their total areas, respectively. This is however an initial rough analysis that only took into account two satellite images.

There are various ways to improve this analysis, starting with an improved image classification based on a more numerous set of training polygons and a larger set of images for averaging normal water extent, for example. On this regard, there are quite a few pixels of flooding in Downtown far from the riverbank that it is not clear they were actually flooded areas or due to a poor image classification on those pixels. Without them, the result might have been different.

There is also an intrinsic limitation due to the relatively large pixel size: if available, Landsat 8 imagery at 15-m resolution or better would help refine the results.

A further analysis could also involve the use of other variables like a DEM of St. Louis City, slope and aspect rasters to better assess risk areas in yet-to-come Mississippi floodings.