Alternative methods for mapping processes in the Ganges-Brahmaputra-Meghna river delta

Tess Jarriel

Final Report

CE 394K

12/2/2016

1. Introduction

River delta systems are one of the environmental ecosystems most threatened by climate change and anthropogenic activity. While their low elevation gradients and fertile soil have made them optimal for human inhabitation and diverse ecologic growth, it also makes them susceptible to adverse effects of sea level rise, flooding, subsidence, and manmade structures such as dams, levees, and dikes. With more than 500 million people live on river delta systems [1], it is important to determine how deltas will respond to changing forcings to inform policy decisions.

One particularly large and threatened delta that is the focus of this study, is the Ganges-Brahmaputra-Meghna Delta (GBMD) on the southern coast of Bangladesh/West Bengal India. Not only is this delta home to over 170 million people [2], but it also hosts the largest continuous mangrove forest in the world, contains important marine and fluvial navigation pathways inland, and produces agricultural resources required for the maintenance of the rest of the region. Because of its importance, studies have been conducted on the GBMD to help us better understand how attributes of the delta network can be linked to known processes acting on the delta [3].

The objective of this study is to improve upon analyses presented in Passalacqua's 2013 paper titled, "Geomorphic signatures of deltaic processes and vegetation: The Ganges-Brahmaputra-Jamuna case study" by repeating the analyses using variations of ArcGIS functions, more recent satellite images, and an alternative method of channel extraction.

This paper is organized as follows. Section 2 describes the aforementioned Passalacqua 2013 paper's data sources, methods, and conclusions. Section 3 outlines the alternative methods used in this paper that have been completed for the scope of this term project. Section 4 outlines the methods that have not yet been completed, but are part of my long term research goals. Section 5 draws conclusions from the study thus far by outlining some lessons learned and future work.

2 Previous Study Summary

Delta networks are difficult to characterize because of the uniqueness of each one, but one way to assess the overall delta is by looking at the land islands formed in and around delta's channels. These islands are representative of the network because their characteristics reflect how sediment is transported and reflect the channel characteristics. In her 2013 article titled, "Geomorphic signatures of deltaic processes and vegetation: The Ganges-Brahmaputra-Jamuna case study" Passalacqua characterized the GBMD delta network based on several delta characterization metrics proposed originally by Edmonds et al [4]. This paper used a combination of Landsat Imagery, ArcGIS mapping tools, statistical analyses, and geological survey data to draw connections between delta metrics and the processes acting on the delta. Specifically, it looked at metrics of island area, shape factor, aspect ratio, oxbow density, channel width, and nearest-edge distance.

2.1 Data Acquisition

The first step in the original study was obtaining the aerial imagery required for analysis. This paper used Orthorectified Landsat Thematic Mapper (TM) Mosaics obtained from landcover.org. The two image tiles used to cover the entire study area (Figure 1) were prepared by

mosaicking separate images from the years 1987-1990 Landsat 5. The three bands used in this analysis were Band 2, Band 4, and Band 7.
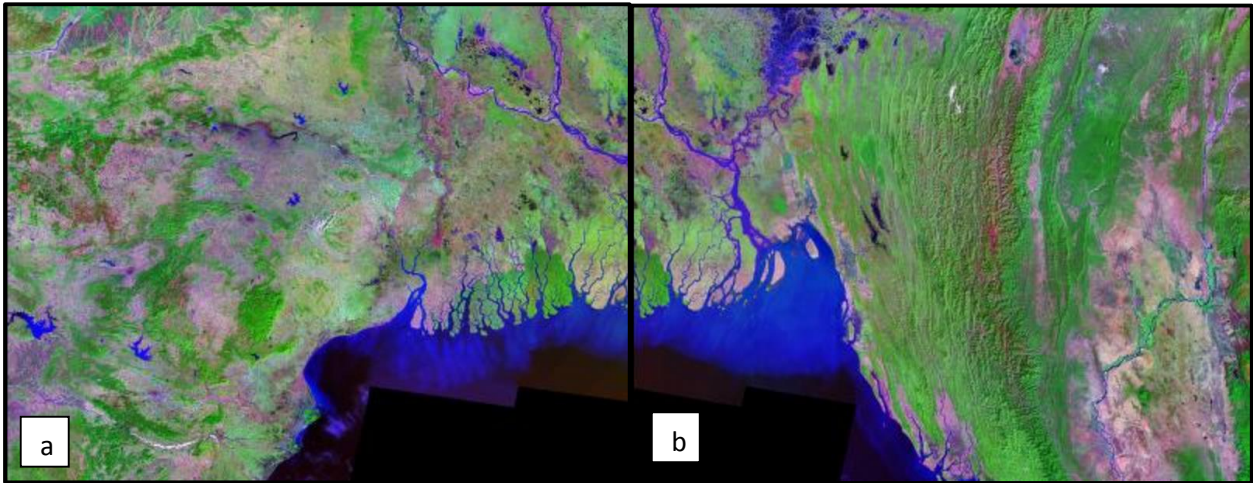

Figure 1: (a) N-45-20 mosaic tile (b) N-46-20 mosaic tile (c) image used in original report

2.2 Channel Extraction

The next step was to use ArcGIS methods to obtain the channel and island networks from the imagery. This was done by using unsupervised and supervised classifications. These respective techniques look at each pixel's spectral signature and determine its ternary (deep water, vegetation cover, or soil) and binary (land or water) classification. The raw output of this method step can be seen in Figure 2a and 2b. As one can see this technique results in a somewhat messy image of the channel. Spurious features were manually edited out and known connections from Google Earth imagery (Figure 2c) were manually entered in to create the final channel image in Figure 2d.
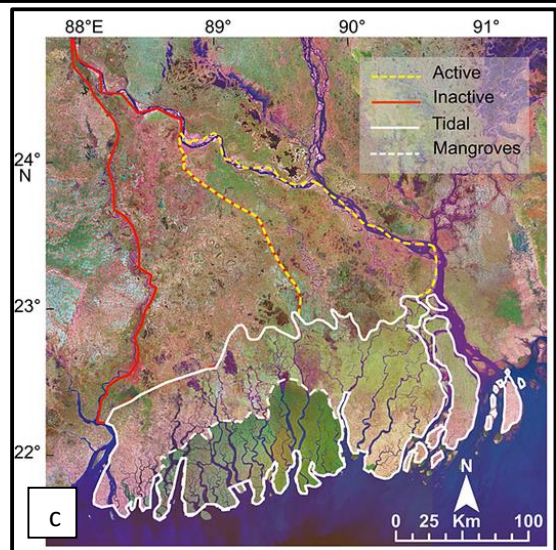
2.3 Obtaining Metric Data

Six delta metrics were obtained from the completed channel network: island area, island shape factor, island aspect ratio, nearest-edge distance, channel width, and oxbow density. Island area (A) is an indicator of the channel connectivity, with smaller islands indicating high connectivity, and larger islands indicating areas of the delta with lower connectivity. Island shape factor (β) is defined as the ratio($P/\sqrt{A}$), where P is the wetted perimeter. This metric is an
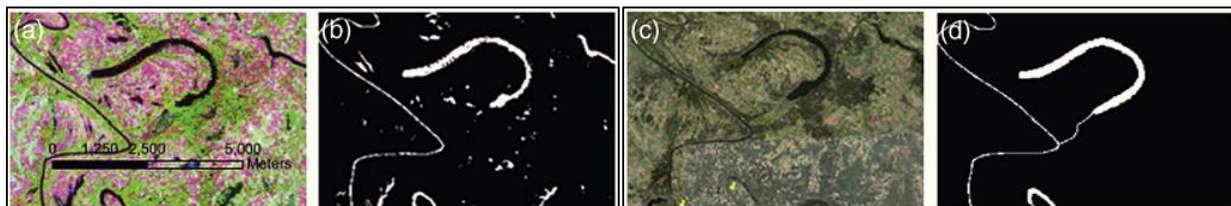

Figure 2: Manual channel network edits: (a) and (b) display the unedited raw classification output. (c) is the google earth imagery used to justify making the manual edits displayed in (d).

indicator of how well drained an island is and the islands roughness. Island aspect ratio (γ) is defined by imagining the island as an ellipse with the same normalized second central moment as the island. Here the aspect ratio is the ratio of the major to the minor axis of this imagined ellipse, and is an indicator of the elongation of an island. The nearest-edge distance (L) is a measure of the shortest straight-line distance from a land pixel to a water pixel. If one looks at the maximum value of this for a particular island, the nearest-edge distance is an indicator of how available water is to the island. Channel width is defined as the minimum distance between channel edges, and the mean value for a channel is an indicator of the strength of the channel's connection to the rest of the network. Oxbow density (O) is a measure of the presence of abandoned channels and oxbow lakes, and can be an indicator of the changes in channel location over time.

2.4 Power-Law Analysis

The data for each of the metrics is statistically analyzed using methods outlined in the Clauset et al paper on "Power-Law Distribution in Empirical Data" [5]. According to Clauset 2009, while some data can be characterized with a simple mean and variance using the normal probability distribution, other data is scale-free and does not have a distribution about a mean value that accurately reflects the data as a whole. This sort of data can be characterized by a power-law probability distribution displayed in Equation 1,

$$p(x) \propto x^{-\alpha} \tag{1}$$

where x is the quantity being analyzed, p(x) is the probability of that value x occurring, and α is the scaling parameter. In actuality, entire datasets rarely exhibit power-law over their entire range, rather exhibit the behavior over certain ranges of x. Regions of a dataset that display power-law behavior can be identified by taking the log of both sides of equation 1 to get:

$$\log p(x) = \alpha \log x + constant \tag{2}$$

By graphing the data on a log-log scale, one can see that the linear region of the plot represents the region where power law applies, and that the slope of this linear region is equal to α. In the hypothetical dataset created in Figure 3, that this data behaves a power law distribution of $p(x) = x^{-2.5}$ for values of x greater than 10.
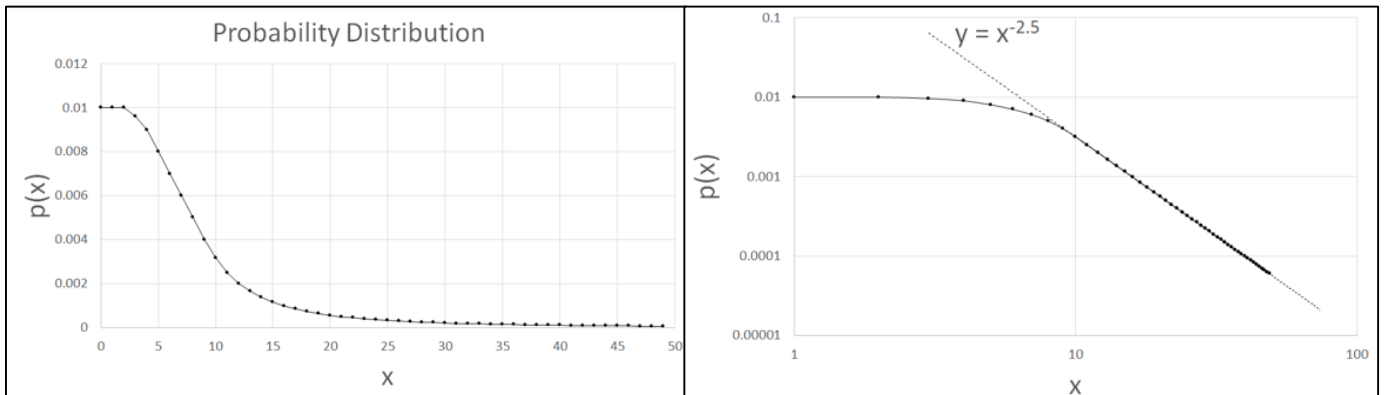

Figure 3: Hypothetical probability distribution function graphed first on a normal scale, then on a log-log scale to highlight the linear region where power-law applies.

Breaks in power-law behavior indicate breaks in the characteristic scales of delta processes, so the value below which power-law no longer applies is a useful number to distinguish zones in a delta. In the report this minimum value was designated as x_{min}. The number of points that fall in the power-law region was designated as n_{tail}. The goodness-of-fit p-value of the dataset was designated as p. For each of the outlined metrics, a value of x_{min}, α, n_{tail}, and p was reported.

2.5 Relating statistics to processes

Once values of x_{min} were chosen for each metric dataset, islands and channels with values lower than that x_{min} (falling outside the power law region) could be isolated and visualized in ArcGIS. By comparing the location of breaks in power law behavior to the known physiographic regions developed by the Geological Survey of Bangladesh [6] (Figure 4), relationships between the metric and the deltaic processes were created. The zones referenced in this figure are active, inactive, tidal, and mangrove.

2.6 Results

It was found that island area distribution behaved power-law for values above an A_{min} of 1.96E7 m^2. The small islands outside this minimum threshold were mostly located in the tidal region, particularly in the mangrove region subset. The largest islands were mostly located in the inactive region of the delta. Island shape factor distribution behaved power-law for values above a β_{min} of 13.7. The less rough islands with smaller values of β were located in the same regions as the islands of smaller area, indicating a relationship between island area and island shape factor. Island aspect ratio distribution behaved power-law for values above a γ_{min} of 2.7. Islands

Figure 4: Recreated zones of the GBMD based on the Geological Survey of Bangladesh

with high aspect ratios and large elongation were found to mainly be along the largest channels and main branches of the delta. The nearest-edge-distance distribution behaved power-law for values above an L_{min} of 1.154E3 m. The areas with L values outside of this range were found mostly in the tidal region and mangrove forests.

Channel width was analyzed in a slightly different fashion than the preceding metrics. In order to assess the sensitivity of the preceding metrics to changing forcing on the system, the weakest channels (smallest width channels below a set 57 m threshold were removed from the system. It was found that when the weakest links were removed, the inactive zone of the delta behaved as one large island. Oxbow density distribution power law parameters were not analyzed, however, simply by mapping their locations, it could be seen that oxbow lakes almost exclusively
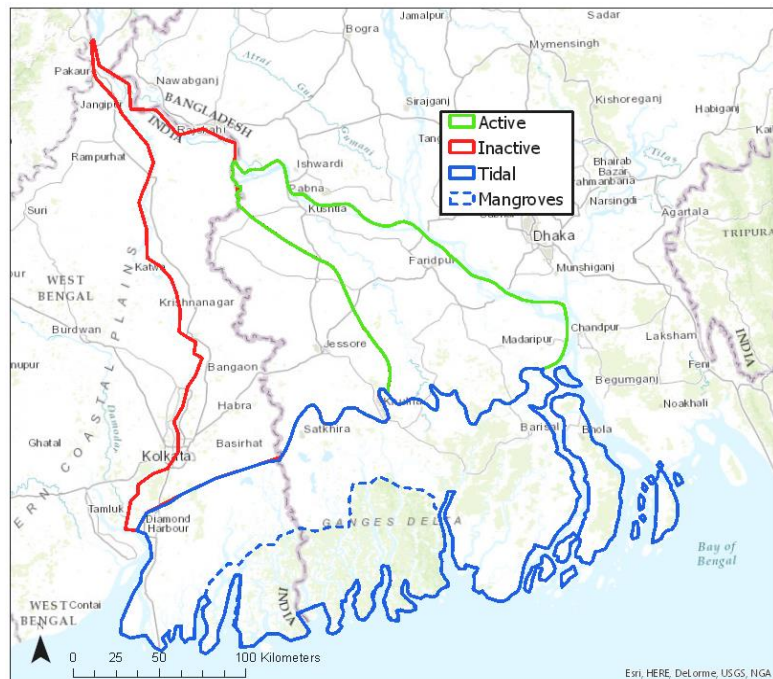
occurred in the inactive zone. This finding suggested that oxbows are more likely to form in regions with weak channel connections.

The paper concludes that the delta can be separated into lower regions of active flow and sediment transport, and upper regions of lower flow and sediment transport, and that the active region can be characterized by high channel density, small islands, good drainage, and short nearest-edge distances [3]. These conclusions were reached based off of Landsat data from 1990, by using ArcGIS to extract the channel network, and by using MATLAB functions to calculate metrics. The rest of this paper will focus on my personal efforts to create alternative methods to accomplish the same goals as Passalacqua's 2013 paper, and compare the results. Specifically, counter to the previous paper, I will do this by using more recent Landsat data, using Rivamap automated channel extraction methods to obtain the channel network, and exploring some different ArcGIS methods for metric calculation.

3. Completed Methods

3.1 Data Acquisition

Because of the plethora of Landsat data variations available online, there were many options of aerial imagery for my analysis. The first option explored was the Orthorectified Enhanced Thematic Mapper (ETM+) Mosaics from Landsat 7. These images used the same bands as the original study; however, they were mosaicked at a more recent time from 1999-2001 images, and they have a higher pixel resolution than the original Landsat 5 images. Unfortunately, as one can see in Figure 5, the available updated Enhanced Thematic Mapper mosaics didn't cover the entirety of the study region outlined in red, and therefore could not be used in isolation.

As a work-around to this issue, I used clipping, conditional classification, composite, and mosaicking techniques in ArcGIS to attempt to join 13 available TIFF images (Figure 6) into one manually mosaicked image that covered the entire study area. Although my efforts were relatively successful, the resulting image had mismatched stretch values and was unnecessarily time consuming.
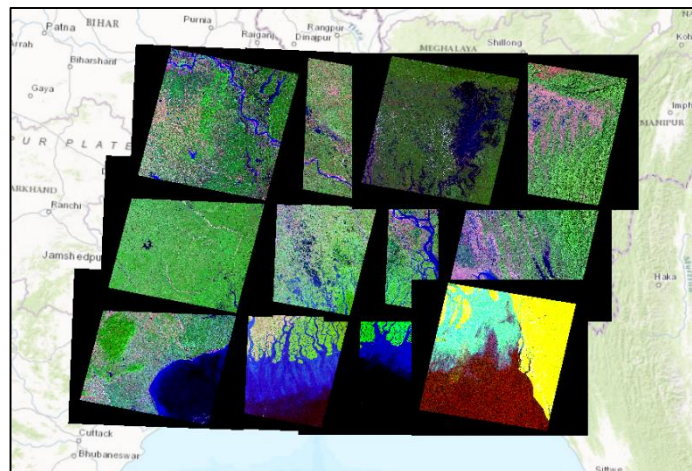

Figure 5: (above) Two ETM+ mosaic tiles from 1999-2001 that potentially could have been used in this study but didn't cover entire study area shown in red.

Figure 6: (right) 13 individual ETM+ images covering whole study area

Although using Landsat 5 data for my study would have made the methods very comparable, for the sake of getting the most up to date imagery, I opted to use data from Landsat 8 instead. Additionally, Landsat 8's raw Band 3 and Band 6 wavelengths are exactly in the range of required values for the Rivamap channel extraction tool, removing the need for extra image processing. I obtained this imagery using Google Earth Engine Code Editor, an online tool that uses Google servers to run scripts you enter yourself. Figure 7 displays the code I used to extract the two bands of Landsat 8 imagery, clip it to the study area, and send the resulting files to my own Google Drive account. This code contains parts directly from the sample scripts provided by the site and others, and parts edited personally by me. At the end of this step, I was left with two TIFF files, one of Landsat 8 Band 3 (green 0.53-0.59 µm) and one of Landsat 8 Band 6 (short wave infrared, 1.57-1.65 µm) that could then be used in the Rivamap tool to create a channel network.

```
Landsat Simple Composite *
1   // Composite Landsat8
2   var L8 = ee.ImageCollection('LANDSAT/LC8_L1T');
3
4 - var composite = ee.Algorithms.Landsat.simpleComposite({
5     collection: L8.filterDate('2015-1-1', '2015-12-30'),
6     asFloat: true});
7
8   // select bands from the composite
9   var compBands = ee.Image(composite)
10    .select(['B3', 'B6']);
11
12  //add bands to display map
13  Map.addLayer(composite, {bands: 'B3,B6', max: [0.3, 0.4]});
14
15  // Create a geometry representing an export region.
16  var geometry = ee.Geometry.Polygon([[[91.5, 21.0], [91.5, 25.2],
17  [87.6, 25.2],[87.6, 21.0]]]);
18
19  // Export the image
20 - Export.image.toDrive({
21    image: compBands,
22    description: 'B3andB6',
23    scale: 30,
24    region: geometry,
25    maxPixels: 2.3e8});
26
```

Figure 7: google earth engine code. Lines 1-6 select Landsat to use and the range of image dates to include. Lines 8-10 select only bands 3 and 6 and composite them into one file. Lines 12-13 display the bands on the site map for visualization purposes. Lines 15-17 define the area over which to clip the compiled images. Lines 19-25 export the completed file to designated google drive account.

3.2 Rivamap Channel Network Extraction

Rivamap is a tool developed by Leo Isikdogan that extracts channel networks from remote sensed imagery. This is done by employing a multiscale singularity index that responds strongly to curvilinear structures and weakly to edges [7]. The input to this tool are two TIFF bands of wavelength 0.53-0.59 µm and 1.57-1.65 µm, and the outputs of the tool are a binary raster TIFF file of the channel network, and a CSV file with columns for width (pixels), latitude (northing), and longitude (easting). This is accomplished through combining a series of functions created in each of Rivamap's included modules.

Rivamap is available in a MATLAB format and a Python format. Because the MATLAB format is stated on the website as being unmaintained, for this study I worked with the python version. Slight adjustments to the code and my system had to be made to make the code (written in python 2.7) compatible with my python 3.5 operating system. After loading Landsat 8 bands 3 and 6 into the script, the following raster image output was achieved (Figure 8).
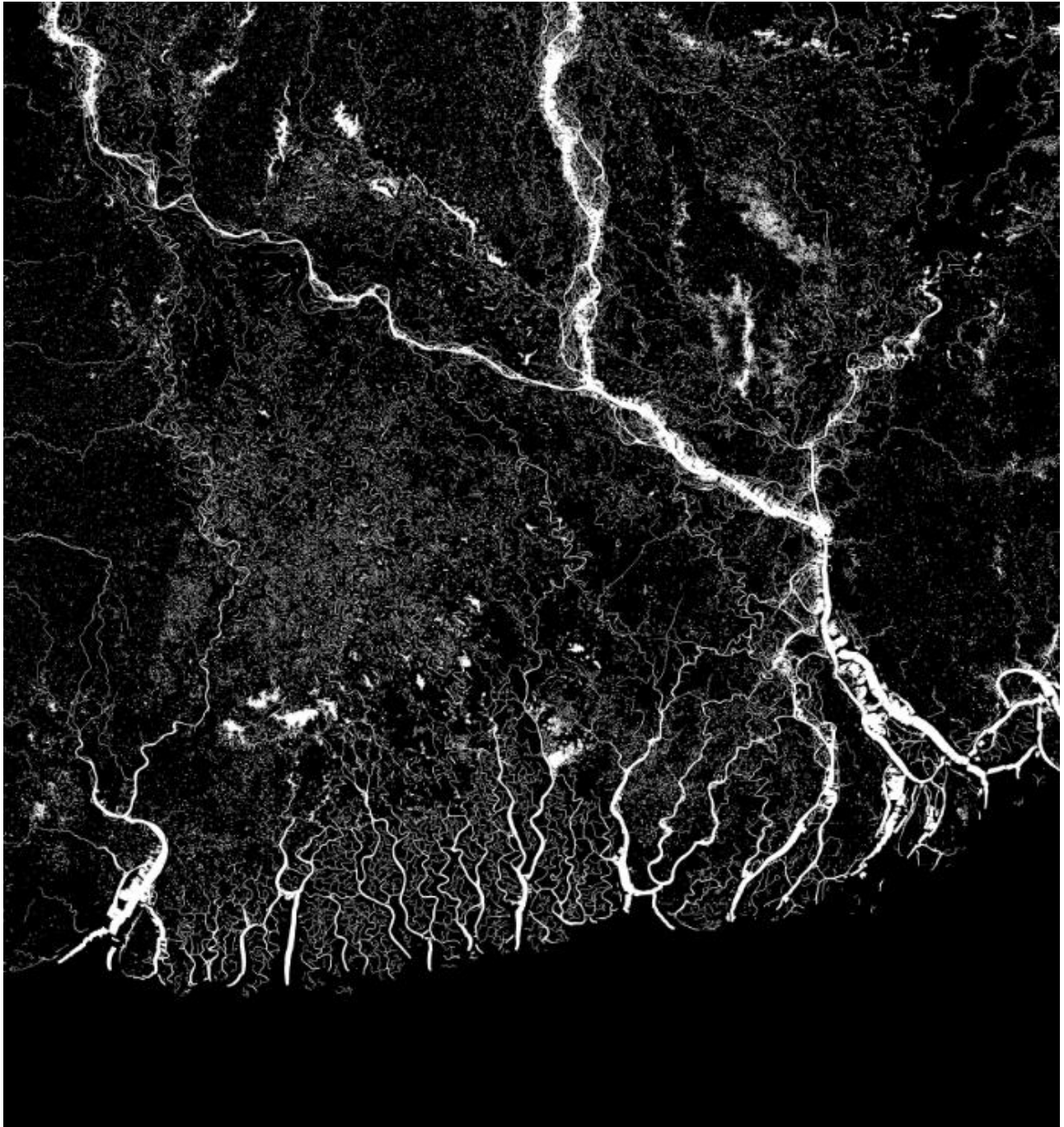

Figure 8: Raster TIFF output file from the Rivamap tool.

3.3 Import into ArcGIS

By loading the CSV of channel centerlines into ArcGIS as a table, you can easily add the centerline points to the map as XY data (Figure 9). When one zooms in on some of the more indistinguishable channel centerlines of the delta (Figure 9c), the power of the Rivamap tool to extract channels from murky images accurately (Figure 9d) is evident.
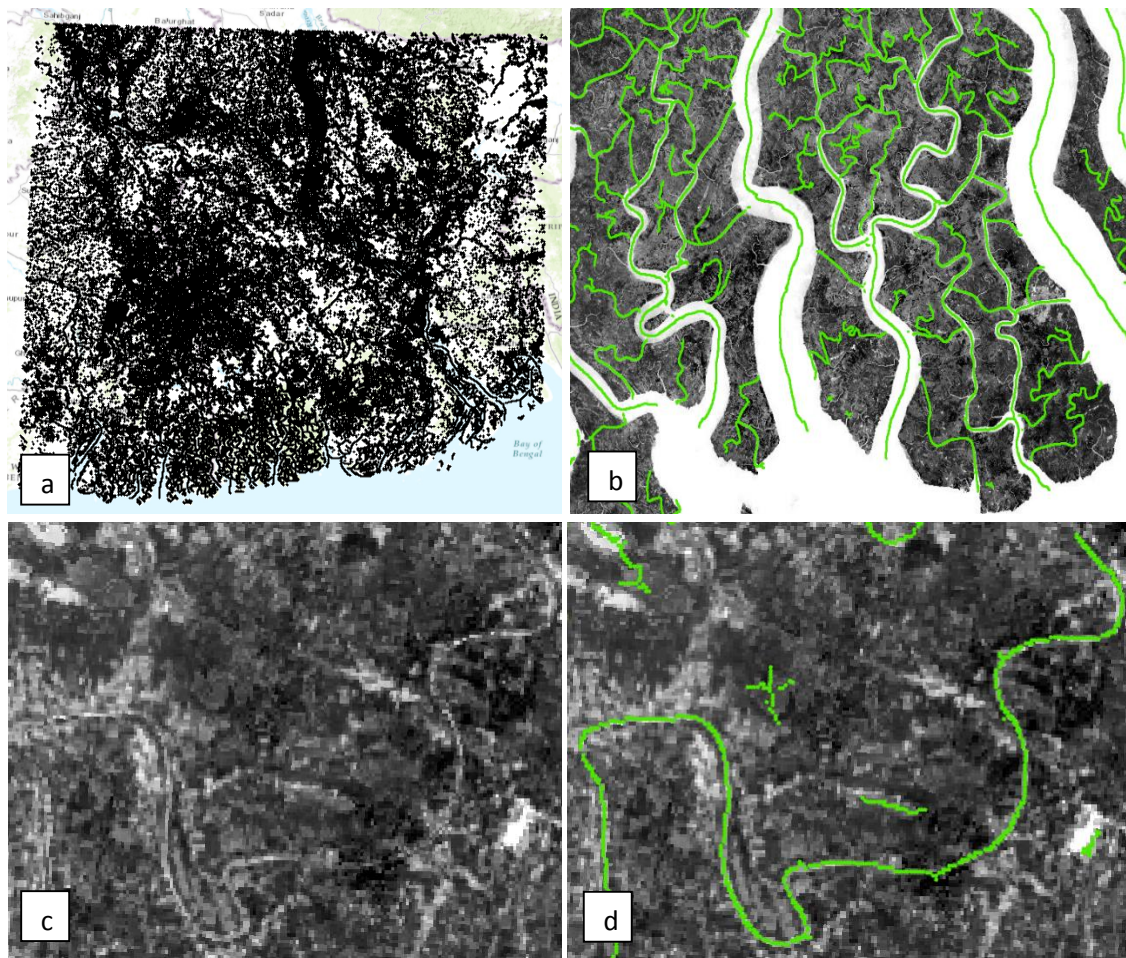

Figure 9: (a) channel centerline points for entire study area, (b) zoomed in section of points file, (c) example of an area where channels are harder to distinguish and (d) a demonstration of Rivamap's capability to extract channel from regions like this.

By adding the width column as an attribute, converting the width from pixels to meters by multiplying by the cell size of 30, and buffering the points using the width field as the indicator for buffer distance, I was able to create the following image of the channel network (Figure 10).
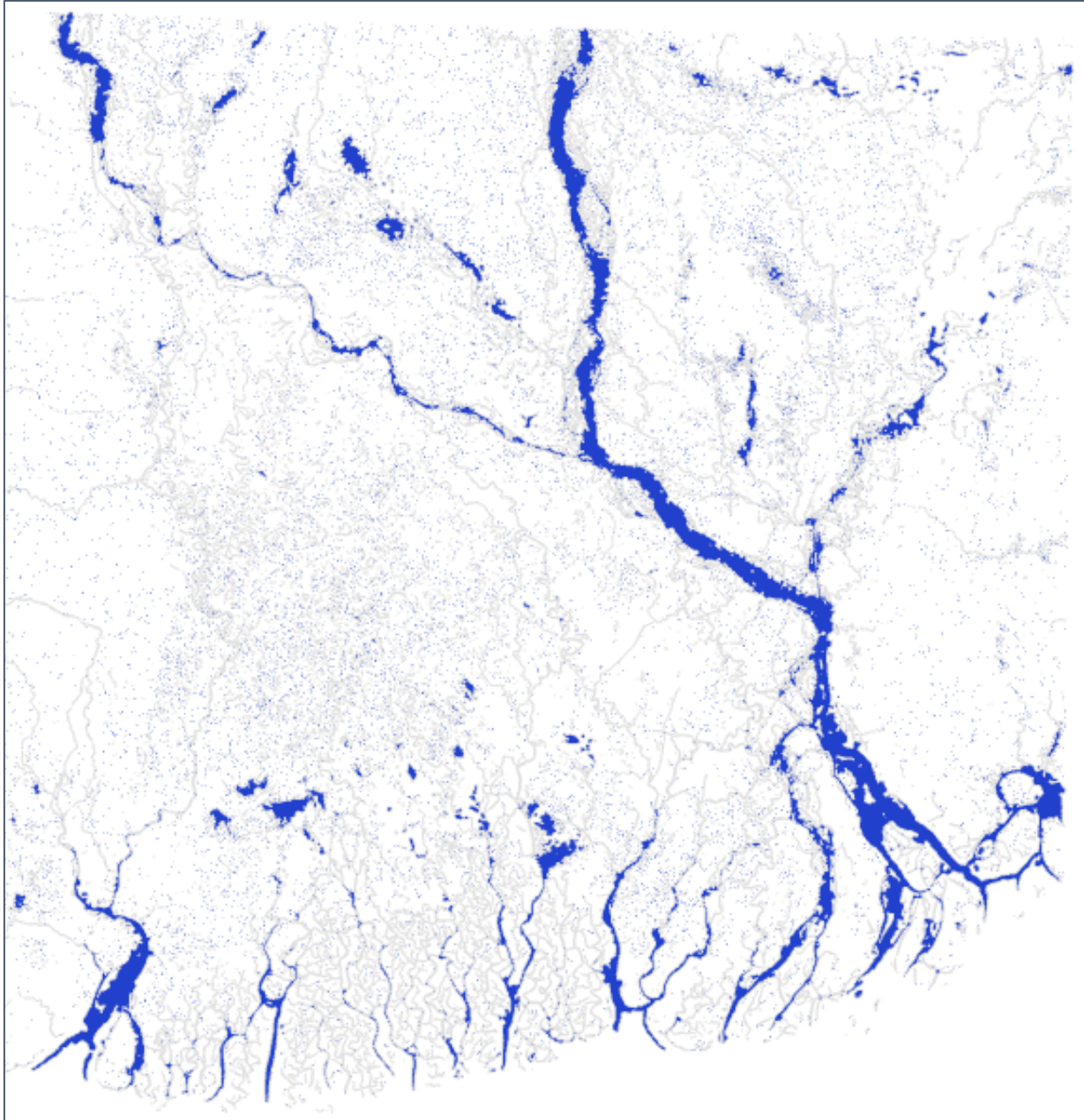
Figure 10: Output from assigning each channel centerline point a buffer equal to its associated width attribute.

At this point in my work I realized a fundamental mistake that precluded me from making any additional analysis on the actual data for the time being. Although it could be seen by overlaying the CSV points on top of the TIFF band images that the centerline points did accurately depict the centers of the channels, when I began to examine the buffer channels I created, I saw an error emerging. As you can see in Figure 11, as one moves down from North to South in the network, the Rivamap channel output becomes less and less accurate to the TIFF image. This error can be attributed to the fact that I neglected to project the two input bands to the Rivamap before running the program.
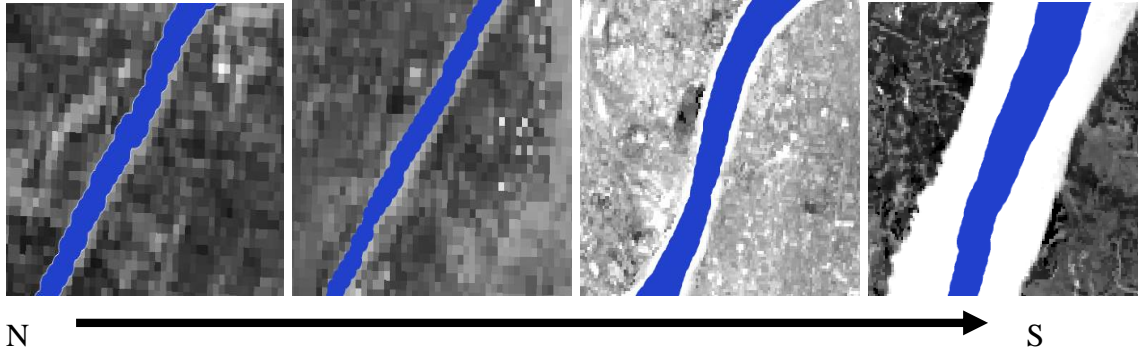
N ➜ S

Figure 11: Channel width increasing error compared to TIFF imagery with decreasing latitude.

The bands were in a geographic coordinate system, and they should have been projected into WGS 1984 UTM Zone 45N before being used in the script. Had I done this, I would have also been able to import the resulting raster file directly into ArcGIS and create a shapefile from it using the raster to polygon tool. This would have simplified the channel creation in ArcGIS, and reduced some of the error associated with the buffer method used in this report.

Although the width data was inaccurate and unusable, I continued to make progress towards my more long term research goals by testing and creating a hypothetical workflow for how I will continue once I get corrected channel width information. These methods are explored in the following section.

4. Planned ArcGIS Metric Analyses

As explained in section 2.2, channel extraction often results in spurious features and incomplete channel linkages. The Rivamap tool was not an exception to these errors. For example, in Figure 12, an incomplete channel from the created channel network is displayed.

Figure 12: Image of the island network with and without manual channel connections. Green lines represent the model interpretation. Red lines represent connections that can be seen in ground truth data.

Were this channel linked as can be seen from the imagery and the red lines manually added for the purposes of this example, the area would be depicted as four separate islands with small shape factors; However, because of the connection error and the failure of the tool to recognize some smaller channels, it will be analyzed as one island with a large shape factor as shown in green. To resolve this issue, similar to the original study, I would have to manually edit all of the errors by comparing the satellite imagery to the mapped channels. This is a very tedious and long task and will not be a part of the scope of my project. Because of this, the metrics of island area, island shape factor, island aspect ratio, and oxbow density are not included in the analysis. However, the remaining two metrics, channel width and nearest-edge distance, do not depend on the complete closure of the channels, and can both be examined if the accurate channel feature class and island feature classes are created.

The reminder of this section is dedicated to outlining ArcGIS methods for creating channel and island features from the raster Rivamap output, ArcGIS methods for obtaining nearest-edge-distance values, and methods to replicate the distribution analyses. As stated, the channel map created in this study was not usable, so the remainder of this paper will use fictitious simplified channel images created by myself to emphasize points and do a sort of "proof of concept" for my proposed methods.

4.1 Channel and Island Feature Class Creation

As mentioned, the raster output to the Rivamap tool could be directly converted to a polygon using the raster to polygon ArcGIS tool. One method to create islands from this newly created channel polygon that I explored was by using the union spatial analysis tool in ArcGIS. This tool allows you to calculate the geometric union of polygons. If I were to create a boundary around the entire delta network (shown in red Figure 13) and union it with the channel network (shown in blue Figure 13), the result is three features: one representing the old channels, one being linear clips of where the channels crossed the boundary, and one being the outline of the white space in Figure 13a. By erasing the first two features, you are left with just the area representing the islands (shown in green figure 13b). This is delivered as one single feature, but by using the multipart to singlepart ArcGIS tool, the polygons can be separated and made ready for statistical analysis.
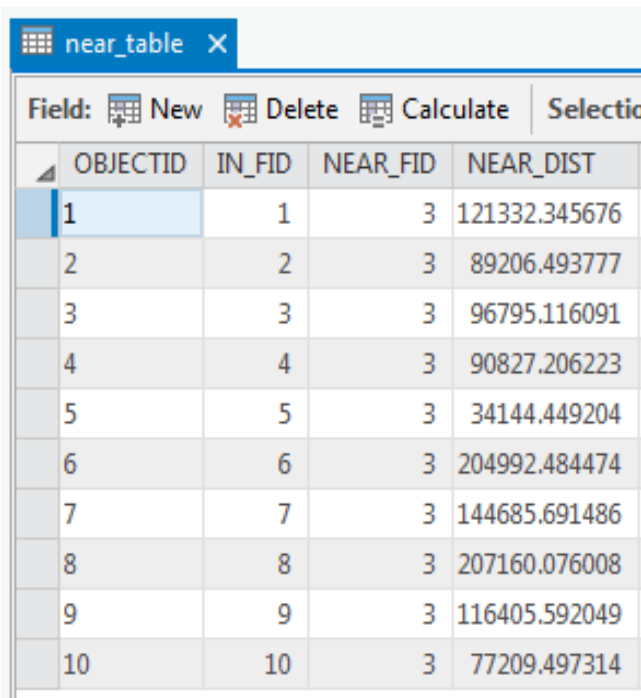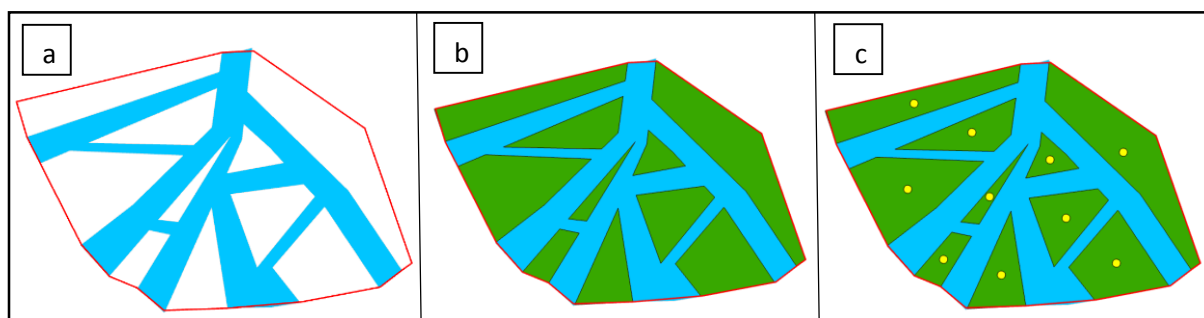

Figure 13: (a) Hypothetical channel network polygon with border of area in red (b) result of the union analysis is creation of island polygons in green (c) result of adding centroid geometry point to each island for nearest-edge-distance assessment.

4.2 Nearest-edge-distance Calculation

While the original paper uses MATLAB functions to obtain data on the nearest edge distance for all pixel points on an island, and then only returns and analyzes the maximum value for the island, an alternative method is to only focus on the maximum distance from the start to cut down on computation time. This can be done by recognizing that the largest value of nearest-edge-distance will occur at the point on an island furthest from the channel, at the islands centroid.

Using the add geometry attribute tool with the newly created island feature class, the centroid points of each of the islands can be added, first to the attribute table of the island polygons, and then to the map as their own set of points (Figure 13c). The nearest-edge-distance for each of the points (corresponding to a particular island), can be calculated using the near table tool. This tool returns a table of calculated distances between features, so by having the centroid points feature class as input, and the as the channels as a near feature, a value of nearest-edge distance for each point is created (Figure 14). It is important when creating the centroid points separate class that one includes the island object ID as a field, so that the results of the near tool can be joined back to the island feature class for later analysis.

Figure 14: Sample output nearest distance for example islands

4.3 Distribution Analyses

While the methods outlined in this subsection are not markedly different than the original study, they will be described nonetheless for completeness sake.

First, to analyze channel width in a similar fashion to the study, I will remove channels less than 57 meters wide and see how the network changes compared to the original study. Here, the buffer method of channel visualization described in section 3.3 becomes useful. While importing the raster TIFF output of python directly into ArcGIS and using the raster to polygon tool, you are left with a single feature where widths at individual channels are lost. If you use the buffer method, the individual widths of each centerline point are preserved. By selecting only points with channel width attribute over the chosen threshold and exporting them to new feature classes, different versions of the channel network can be easily created and interpreted.

The nearest-edge-distances for each island can be exported in a table from ArcGIS and opened in Excel (or any similar calculation software). By creating a histogram and appropriately sized bins, a distribution function of the data can be created. Like in the original study, by graphing the data on a log-log scale, the linear region that follows power-law behavior can be isolated, and the x_{min} can be chosen. This x_{min} value can be interpreted visually just by looking where the linear

region seems to start, but this method adds a certain level of subjective error to the analysis. An alternative approach is to use the Kolmogorov-Smirnov (KS) statistic as a measure of the difference between the probability distributions of the observed data and the fitted model. By simply choosing the x_{min} value that minimizes the summed difference between the two lines, an ideal x_{min} can be obtained [5]. All of these new parameters for the nearest-edge-distance power-law distribution (x_{min}, α, n_{tail}, and p) can be compared to the original study values to see what changes the different methods posed in this paper have caused.

5. Conclusions

This term project objective was to explore alternative methods to potentially make improvements upon the work done in Passalacqua's 2013 report "Geomorphic signatures of deltaic processes and vegetation: The Ganges-Brahmaputra-Jamuna case study." My term project is not intended as a critique of the methods chosen in the original paper in the slightest, rather, it is more so a potential update recommended due to the continuously evolving nature of the GIS and remote sensed imagery fields as a whole. The original paper was submitted in January of 2013, but the improved Landsat 8 data only became available in February of 2013. Similarly, there have been countless updates to the ArcGIS software in the past three years that have improved upon its functions. Finally, the automated channel extraction techniques of Rivamap were not created by Leo Isikdogan until November of 2015. There is no guarantee by any means that the methods I have proposed will improve the original analysis or return different results, but completion of this study in over the next few months will answer that question.

In terms of future work required, I will need to implement all of the methods outlined in section 4 of this paper. Additionally, I think that it would be very useful to compare the channel extraction techniques (unsupervised/supervised classification vs. Rivamap) on the same Landsat data source to see what the differences are between the two. In his paper [7] Isikdogan contrasts his model results to the ground truth and shows that Rivamap results are "comparable with the ground truth," but I think it could be useful to compare it to a different extraction technique as well to see which method is better.

This term project was very beneficial to me personally. I enjoyed re-visiting writing code and exploring the ArcGIS software capabilities. Next time I may make more effort to perform some of the large dataset operations on more powerful computers to cut back on some of the processing time. I'd like to thank my advisor Dr. Passalacqua for her guidance and for giving me the inspiration for this study. I'd also like to thank Leo Isikdogan for promptly answering my endless stream of questions about installing the Rivamap modules. Lastly, thank you to the course instructor Dr. Maidment for his advice. Although I only needed his help sparingly, his small hints to use the buffer tool for channel creation and the union tool for island creation were large in their impact.

References:

[1] Giosan, L., Syvitsky, J., Constantinescu, S., Day, J., 2014. Protect the world's deltas. *Nature* 516 (7529), 31-33.

[2] Auerbach, L. W., Goodbred Jr, S. L., Mondal, D. R., Wilson, C. A., Ahmed, K. R., Roy, K. Steckler, M. S., Small, C., Gilligan, J. M., Ackerly, B. A., 2015. Flood Risk of natural and embanked landscapes on the Ganges-Brahmaputra tidal delta plain. *National Climate Change*, 5, 153-157.

[3] Passalacqua, P., S. Lanzoni, C. Paola, A. Rinaldo. 2013. Geomorphic signatures of deltaic processes and vegetation: The Ganges-Brahmaputra-Jamuna case study, *J. Geophys. Res. Earth Surf.,* 118, 1838–1849.

[4] Edmonds, D. A., C. Paola, D. C. J. D. Hoyal, B. A. Sheets. 2011. Quantitative metrics that describe river deltas and their channel network, J. Geophys. Res., 116, F04022.

[5] Clauset, A., Shalizi, C. R., Newman, M. E. J., 2009. Power-law distributions in empirical data. *SIAM Review,* 51(4), 661-703.

[6] Alam, M. K., Hasan, A. K. M. S., Khan, M. R., Whitney, J. W. 1990. Geological map of Bangladesh, *Tech Rep,* Geological Survey of Dhaka, Bangladesh.

[7] Isikdogan, F., Bovik, A., Passalacqua, P. 2015. Automatic channel network extraction from remotely sensed images by singularity analysis. *IEEE Geoscience and Remote Sensing Letters*, 12, 11, 2218-2221.