Hydrologically Processing a Digital Elevation Model of an Urban Sub-watershed for Hydrological Analysis

Amanda L. Riggs

October 14, 2024

East Carolina University

Title: *Hydrologically Processing a Digital Elevation Model of an Urban Sub-watershed for Hydrological Analysis*

Author: Amanda Riggs
Institution: East Carolina University, Department of Geography, Planning, and Environment
Date of Completion: 10-2024

This work was completed in partial fulfillment of the requirements for the degree of Master of Geography at East Carolina University.

For permissions or inquiries, contact: amanda.l.riggs@outlook.com

1 Objective

The objective of this hydrological processing project is to develop an accurate and hydrologically consistent digital elevation model (DEM) for use in the Natural Capital Project's InVEST® Nutrient Delivery Ratio (NDR) model, as part of a thesis on urban stormwater resource management. Utilizing ArcGIS Pro 3.4.0 and advanced techniques like stream burning, hydrologically (hydro)-flattening, conditioning and enforcing, the project refines the DEM to effectively capture urban water flow dynamics.

Hydro-flattening refers to “flattening” water surfaces to mimic their behavior in digital terrain models (DTMs) (NGP Standards and Specifications, 2024). Standard topographic DEMs represent ground surfaces while managing hydrologic features where roadways over culverts are included, but bridges are excluded. These DEMs are created from mass points and breaklines derived from aerial imagery, defining waterbody edges and enabling smooth flow beneath bridges. Processing the DEM ensures continuous water flow by removing spurious sinks. "Hydro-conditioned" refers to consistent flow across basins, while "hydro-enforced" pertains to mapped drainage features, crucial for understanding catchment relationships. Hydro-enforcement adjusts waterbodies, so lakes are level and streams flow downhill, accurately representing terrain under structures like bridges and improving hydrologic modeling.

The processed DEM enhances the NDR model's ability to simulate nutrient delivery export, shedding light on the connections between urban development, ecological conservation, and water quality. By integrating ArcGIS Pro tools, remote sensing datasets, and validated methodologies, this project aims to support sustainable urban planning and advance hydrological and nutrient modeling practices.

2 Study Area and Data

2.1 Study Area

The study area is in the Greens Mill Run (GMR) sub-watershed and is centrally located in Greenville, NC (Figure 1). It encompasses approximately 8,572 acres and is predominantly urbanized, with approximately 5,195 acres of urban land use, along with 1,031 acres of cropland, 510 acres of pastureland, and 1,832 acres of forested land (EPA, 2024; Hazen and Sawyer, 2016). The GMR sub-watershed has experienced the highest nutrient export and runoff levels among the five central sub-watersheds in Pitt County from 2001 to 2021, according to the Natural Capital Project's InVEST NDR model results ran from an older 2000 DEM acquired from HydroSHEDS version one, developed by World Wildlife Fund (WWF), derived from Shuttle Radar Topography Mission (SRTM) at 3 arc-second resolution (Natural Capital Project, 2024; Lehner, 2022). Such nutrient loading has severely impacted the primary stream of GMR, resulting in significant erosion and rapid changes in stream flow (Boyd, 2015; Hazen and Sawyer, 2016).

Geographically, GMR encompasses much of uptown Greenville and East Carolina University. The sub-watershed includes 70% within city limits, 26% in outer city limits, and 4% in Pitt County (Boyd, 2015; Greenville, North Carolina, n.d.; Hazen and Sawyer, 2016). Boundaries include East 5th Street to the north and east, Greenville Boulevard and Red Banks Road to the south, and US-264 bypass to the west.

GMR drains from west to east into the Tar River, with the eastern half of the sub-watershed highly developed with, over 50% developed, and 23% impervious (Boyd, 2015;

Hazen and Sawyer, 2016). The sub-watershed features various hydrologic soil groups, primarily type A/D and A (Hazen and Sawyer, 2016), and includes four significant tributaries: Reedy Branch, Fornes Run, Patricks Run, an unnamed tributary, and GMR North Fork, with elevations ranging from 6.46 to 21.49 meters (NAVD) and a channel slope of approximately 0.0003 meters (Hazen and Sawyer, 2016). The average annual temperature is 61°F, reaching 84°F in summer and dropping to 44°F in winter, accompanied by an average annual rainfall of 0.02 millimeters and a growing season of 220 days (Boyd, 2015; Greenville, North Carolina, n.d.). Figure 1 depicts a map of the Greens Mill Run (GMR) sub-watershed with its tributaries and bounding roadways, as well as its county and state location.

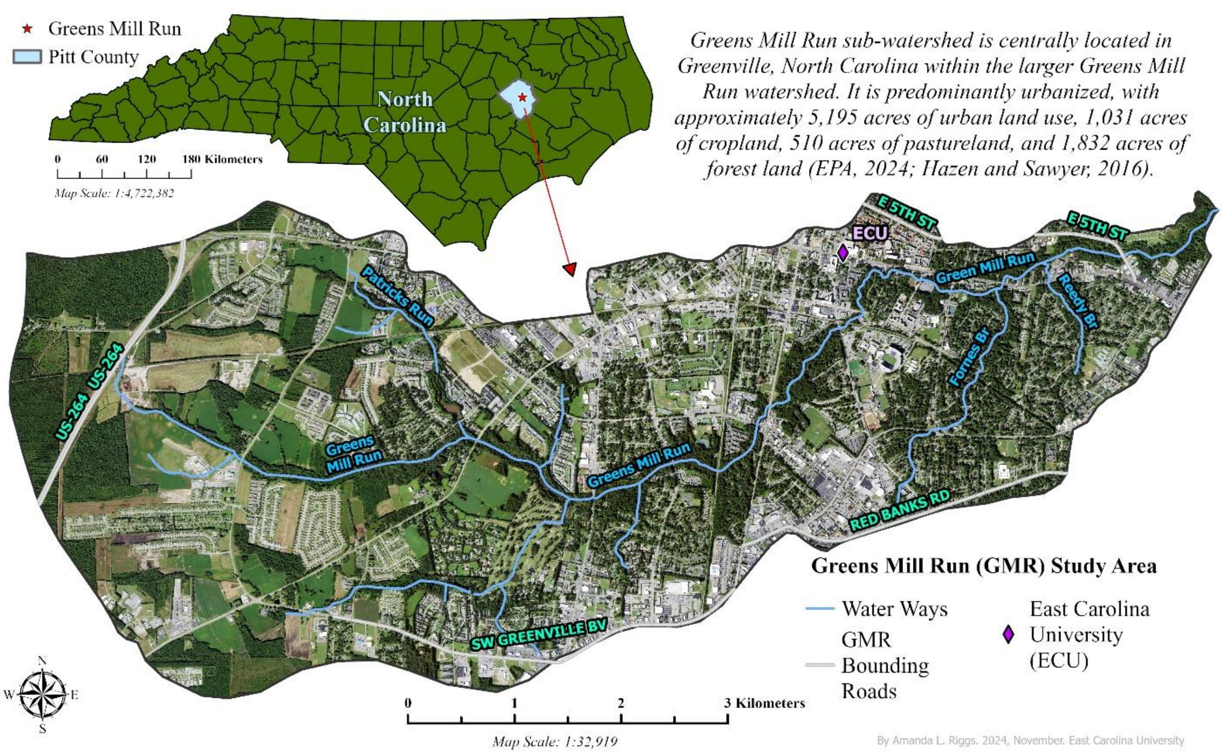


Figure 1. A map of the Greens Mill Run (GMR) sub-watershed with its tributaries and bounding roadways, as well as its county and state location.

2.2 Data

2.2.1 Digital Elevation Model

A Digital Elevation Model (DEM) serves as a critical input for InVEST's Nutrient Delivery Ratio (NDR) model, enabling the accurate simulation of water movement and nutrient delivery. The raw DEM, derived from a Light Detection and Ranging (LiDAR) point cloud, requires hydrological processing, including flattening and conditioning to ensure its suitability for sensitivity analysis and accurate modeling results. The DEM used for this study is the 2017 United States Geological Survey (USGS) 1/3 arc-second DEM (USGS10m), obtained from OpenTopography (https://portal.opentopography.org/datasets), a platform supported by the National Science Foundation (NSF) (United States Geological Survey, 2021). This dataset, part of the USGS 3D Elevation Program (3DEP), provides high-resolution elevation data with horizontal coordinates in North American Datum of 1983 (NAD83) (European Petroleum Survey Group (EPSG): 4269) and vertical coordinates aligned to North American Vertical Dataum of 1988 (NAVD88) (EPSG: 5703). The DEM is distributed in Geographic Tagged Image File Format (GeoTIFF), ensuring compatibility with geospatial analysis and visualization software.

The DEM encompasses the geographic extent defined by coordinates spanning Xmin: -77.46357, Xmax: -77.299461, Ymin: 35.510169, and Ymax: 35.663729, delivering detailed elevation data for the Greenville area. The dataset does not include additional hydrologic terrain analyses or canopy height modeling, nor is pit filling, flow direction, catchment area, and wetness index not enabled during processing. Contour lines were also not generated, making the DEM a valuable resource for hydrological modeling, terrain analysis, and landscape planning, offering precise elevation data aligned with USGS and NSF-supported standards.

2.2.2 Imagery

The 2016 imagery from the National Agriculture Imagery Program (NAIP) serves as a layer for validation in this project. Acquired through National Oceanic and Atmospheric Administration's (NOAA) Digital Coast: Data Access Viewer (https://coast.noaa.gov/dataviewer/#/imagery/search/), the NAIP is administered by the U.S. Department of Agriculture (USDA) Farm Service Agency (FSA) and provides high-resolution ortho-imagery captured during the agricultural growing season (Organizational Change Management (OCM) Partners, 2024). This imagery supports goals related to natural resource stewardship, environmental enhancement, and food security, serving as a foundational tool for USDA farm and conservation programs. NAIP imagery products include Digital Ortho Quarter Quads (DOQQ tiles), Compressed County Mosaics (CCM), and Seamline shapefiles, formatted as 3.75' x 3.75' quarter quadrangles using the Universal Transverse Mercator (UTM) coordinate system (NAD83), with up to 10% allowable cloud cover per tile.

Horizontal accuracy has improved significantly since the program's inception, with all acquisitions after 2009 adhering to a ± 6-meter accuracy to true ground. This standard ensures the imagery's reliability for hydrological analyses and other applications. The 2016 NAIP imagery dataset was collected using Leica Airborne Digital Sensor (ADS)-100, mounted on twin-engine turboprop aircraft flying at 16,500 feet. The imagery achieved a nominal ground sampling distance (GSD) of 0.40 meters with a 23% sidelap, capturing Red, Green, Blue, Near Infrared, and Panchromatic bands. Processing included airborne Global Positioning System (GPS)/Internal measurement Unit (IMU) data refinement using INYS, PosPac, and TerraPos software, followed by aero triangulation and orthorectification with USGS 10-meter DEMs. Radiometric adjustments reduced 12-bit data to 8-bit for the final mosaics, ensuring consistency in sun angle and azimuth. Accuracy was verified by comparing photo-identifiable survey locations with ground control points using ArcGIS.

NAIP imagery is made available within 60 days of the end of the flying season, supporting diverse applications such as USDA farm program monitoring, Common Land Unit boundary establishment, land use planning, natural resource assessments, and disaster response. The program's three-year refresh cycle and commitment to technological innovation ensure the delivery of high-quality, current imagery for agricultural and conservation efforts.

2.2.3 Polyline Stream Vector

The 2022 Topologically Integrated Geographic Encoding and Referencing (TIGER)/Line polyline stream vector dataset for Pitt County, North Carolina, Area Hydrology was acquired from the North Carolina Spatial Data Download platform (https://sdd.nc.gov/DataDownload.aspx#) for stream alignment validation. This dataset provides detailed streamlines, including smaller rivers and streams essential for analyzing the sub-watershed within the study area. Although some minor streams remain unrepresented, this dataset was particularly valuable compared to others that primarily included major hydrographic features. Its inclusion of streams and rivers less than 30 miles in length made it well-suited for the requirements of this analysis.

The TIGER/Line shapefiles are derived from the United States (U.S.) Census Bureau's Master Address File (MAF)/TIGER database. They offer both geographic and cartographic data for a variety of hydrographic features, including ponds, lakes, oceans, swamps, glaciers, and extensive river and canal networks (U.S. Census Bureau, 2022). The dataset utilizes the NAD83 coordinate system (EPSG:4269). While the data proved critical for the analysis, specific accuracy metrics were not included in the accompanying metadata, which may limit precision assessments. The dataset was published in 2022, with the reference system last revised on January 19, 2007.

2.2.4 Watershed Boundary Vector

The Watershed Boundary Dataset (WBD), available through the USDA Geospatial Data Gateway (https://datagateway.nrcs.usda.gov/GDGHome_DirectDownLoad.aspx), is a comprehensive hydrologic unit dataset developed at a 1:24,000 scale. It builds upon the original hydrologic units created in the 1970s by the U.S. Geological Survey under the sponsorship of the Water Resources Council. The WBD encompasses hydrologic units ranging from sub-watersheds smaller than 10,000 acres to entire river systems, organized and attributed using a standardized hierarchical nomenclature (National Hydrography, 2023).

Development of the WBD began in the early 1990s, with state-level delineation and attribution based on diverse methods and source data. Although the WBD metadata does not specify accuracy measurement results, it states that each state's dataset underwent extensive quality assurance to meet the Federal Standard for Delineation of Hydrologic Unit Boundaries. The WBD provides a seamless national framework across six hierarchical levels (HUC-2 to HUC-12) and includes attributes such as hydrologic unit codes, names, downstream units, artificial modifications, and unit types. Associated shapefiles offer polygon and line representations, along with metadata fields for hydrologic unit levels and data sources.

2.2.5 Polyline Road Vector

The polyline road vector is a polyline feature class representing road centerlines, downloaded via an organizational portal in ArcGIS Pro. This layer is configured as a feature service hosted at ArcGIS REST Services. The dataset utilizes the World Geodetic System (WGS) 1984 geographic coordinate system (EPSG: 4326) for its native geometry and the Web Mercator projection (EPSG: 3857) for visualization and analysis. Its spatial reference includes a prime meridian at Greenwich, an angular unit of degrees, and a spheroid defined by the WGS 1984 datum, with a semimajor axis of 6,378,137 meters and inverse flattening of 298.257223563 (ArcGIS REST Services, 2024).

The dataset comprises line features with no Z or M values and does not support attachments, versioning, or archiving. It includes attributes such as street names, classifications, address ranges, maintenance details, pavement width, and length metrics. The extent spans coordinates from X Min: -8622038.8969, Y Min: 4233575.7981 to X Max: -8603208.6091, Y Max: 4255884.6017. Data resolution and tolerances are defined at 0.0001 meters and 0.001 meters, respectively, supporting detailed spatial analysis within this extent.

The dataset, configured with a Mercator Auxiliary Sphere projection, is optimized for visualization and analysis at scales ranging from 1:288,896 to the full extent and supports various query formats, including JavaScript Object Notation (JSON) and geoJSON. With its comprehensive attribute structure and reliable geometry, this dataset provides a reference for transportation analysis, planning, and defining bounding roads around the study area.

2.2.6 County & State Boundary Vectors

Included in Figure 1, is an insert of the study area’s state and county, derived from 2024 TIGER/Line Shapefiles. The TIGER/Line shapefiles provide updated geographic boundaries for legal and statistical entities across the U.S., consistent with 2024 survey data like the American Community Survey, obtained through the United States Census Bureau official website (https://www.census.gov/cgi-bin/geo/shapefiles/index.php). These shapefiles include feature shapefiles that depict essential elements such as roads and rivers, as well as relationship files that offer supplementary attribute information. Each shapefile comes with a projection file (.prj) that specifies the coordinate system, utilizing the Global Coordinate System (GCS) based NAD83, with area measurements in square meters (Census Bureau, 2022).

To maintain accurate geographic data, the Census Bureau regularly updates the MAF/TIGER system, although the completeness of the files can vary based on source documents. While coordinates are provided to six decimal places, this precision doesn’t guarantee high positional accuracy, nor are accuracy measurement results available in the metadata. The TIGER/Line Shapefiles include Federal Information Processing Series (FIPS) codes and the U.S. Geological Survey’s Geographic Names Information System (GNIS) codes for easy identification. The methodology does not include this dataset, as it is simply a reference for the study area in Figure 1.

3 Methods

This study focuses on a refined workflow to produce a DEM that balances hydrological accuracy with visual effects by combining stream extraction, hydro-flattening, stream burning, and gradual stream-sinking techniques; the workflow ensures smooth transitions that align with natural terrain topography. Figure 2 illustrates a contour map of the original DEM before processing, highlighting the artifacts that were present in the original DEM due to irregularities in the elevation data. After flattening the stream, these artifacts are significantly reduced, resulting in a more accurate representation of the stream's flow within the stream channels of the study area.

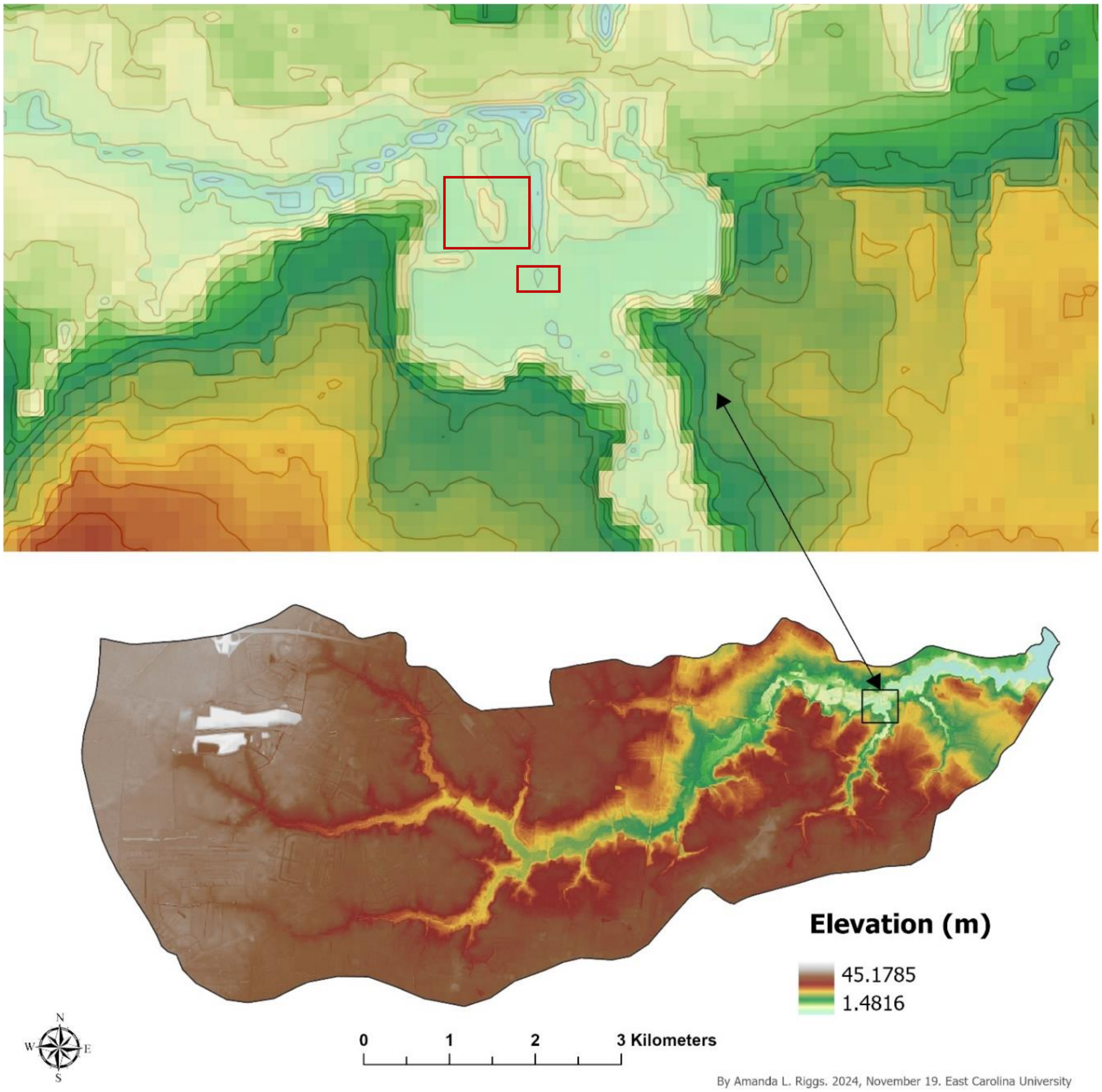


Figure 2. Contour map of the original DEM showing the artifacts present in the original DEM due to irregularities in the elevation data.

The methodology combines practices from referenced sources, drawing on techniques tailored to the study area. Developed with the assistance of OpenAI (2024), the workflow adheres to principles of stream burning and hydro-flattening outlined by Chen et al. (2024), the NGP Standards and Specifications (2024), and Crawford's (2022) guidelines for LiDAR-based

DEM processing, utilizing tools from ArcGIS Pro 3.2.0 and Esri, as shown in Figure 3.

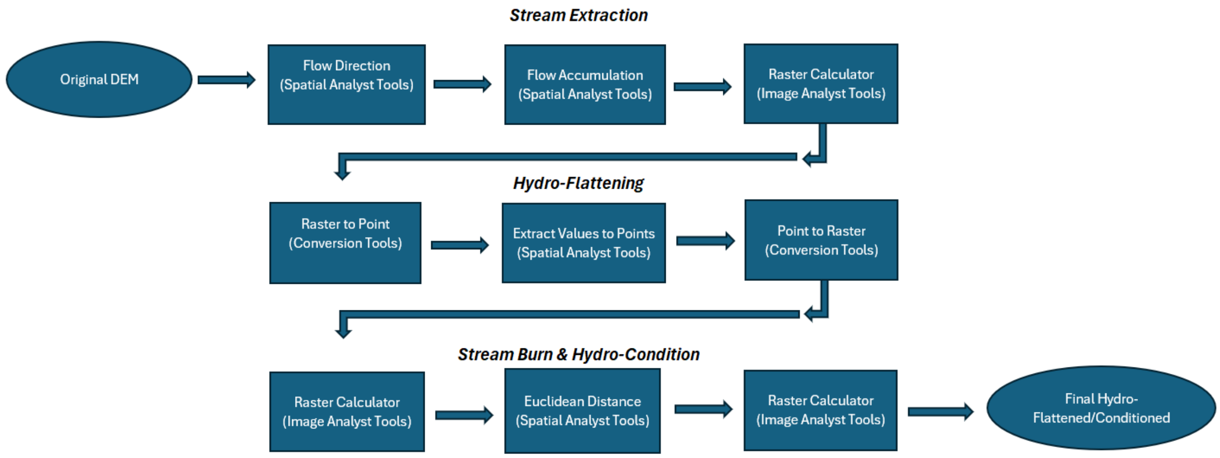


Figure 3. Workflow outlining the three-step process: stream extraction, hydro-flattening, and hydro-conditioning to prepare a refined DEM.

3.1 Hydro-Conditioning & Stream Extraction

Hydro-conditioning began with pre-processing the DEM by filling sinks and pits in the original DEM to ensure that the flow of water is continuous across the terrain surface. Next, the stream extraction began with the creation of a flow accumulation raster from the filled DEM. Using the Flow Direction (Spatial Analyst Tools) tool, water flow paths were modeled based on the steepest downslope path from each cell, using default settings consistent with the D8 algorithm employed in similar studies (Chen et al., 2024). The Flow Accumulation (Spatial Analyst Tools) tool was then used to calculate the volume of water flowing through each cell, also using default methods, including the Deterministic eight-node (D8) algorithm which uses the eight adjacent cells to determine the flow direction from each pixel to its steepest downslope neighbor to produce a raster that highlighted potential stream locations (ArcGIS Enterprise, 2024).

To isolate significant streams, thresholds of 24,000, 34,000, and 44,000 flow accumulation units were tested using the Raster Calculator (Image Analyst Tools) tool to apply the conditional statement:

$$Con(Flow_{Accumulation} > 24000, 1, 0)$$

A map comparing the flow accumulation raster to the three tested binary stream raster outputs is shown in Figure 4. The resulting binary stream raster was validated by overlaying it with aerial imagery and reference data, confirming alignment with major stream channels. Following recommendations from Chen et al. (2024), for threshold-based delineation in large-scale datasets, the threshold of 24,000 was selected as it represented a balance between capturing medium-flow streams and excluding smaller, runoff-driven tributaries that contribute minimal flow (Chen et al., 2024). This output served as the foundation for subsequent hydro-flattening and conditioning steps.

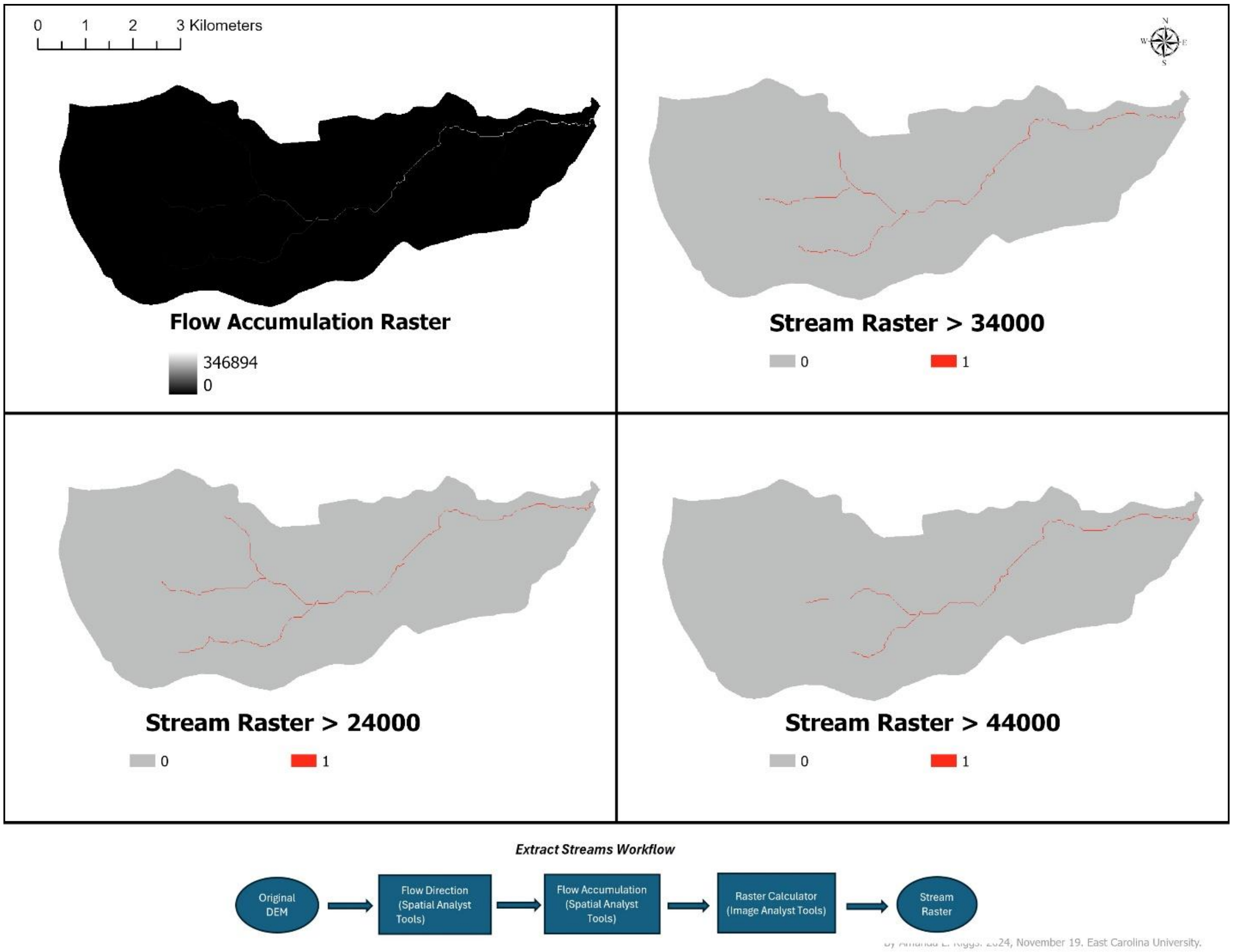


Figure 4. Flow accumulation raster and resulting stream rasters for the three tested thresholds: 24,000, 34,000, and 44,000 flow accumulation units.

3.2 Stream Burning & Hydro-Flattening

Stream burning and hydro-flattening was performed to embed the extracted stream network into the DEM and remove the artifacts from the streams in the DEM. A Stream Elevation Raster was created by converting the binary stream raster to points using the Raster to Point (Conversion Tools) too. Next, the workflow extracted elevation values from the conditioned DEM using the Extract Values to Points (Spatial Analyst Tools) tool. The default tool parameters were used, including the option to interpolate values at the point locations and append all the stream raster elevation values to the point features. These stream points, now containing elevation attributes, were converted back to a raster using the Point to Raster (Conversion Tools) tool. The *RASTERVALU* field which holds the elevation value was used as the input. The cell size matches the original DEM (9.32m) to ensure alignment with the original DEM

The Raster Calculator (Image Analyst Tools) tool was then employed to burn the stream elevations into the DEM. The elevation values of stream cells were replaced with those from the stream elevation raster while preserving the original elevation values for non-stream areas using the conditional statement:

$$Con(\text{Stream\_Raster} == 1, \text{Stream\_Elevation\_Raster}, \text{Original\_DEM})$$

This method supports hydro-enhancement practices outlined in USGS specifications for DEM preparation (USGS, 2024). The resulting hydro-flattened DEM incorporated stream networks within the natural terrain that appear flat and continuous, with less artifacts than the original DEM, confirming visual consistency while maintaining topographical features, as seen in Figure 5.

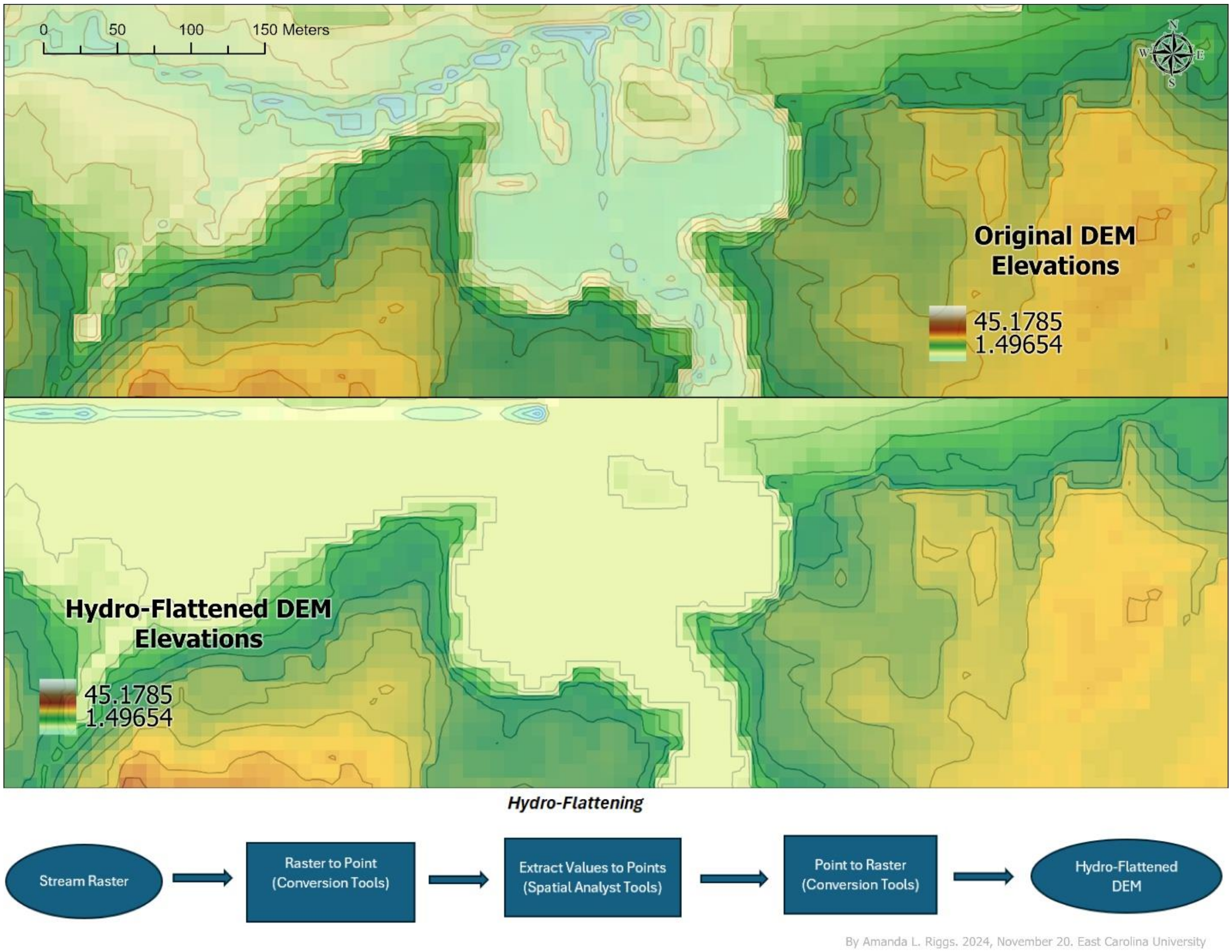


Figure 5. A comparison map between the original DEM and the Hydro-flattened DEM.

3.3 Hydro-Enforcing

To hydro-enforce the DEM, the Euclidean Distance (Spatial Analyst Tools) tool was employed, along with its default settings, Planar distance method, and matching cell size to the original raster (9.32m), calculated the distance of each cell from the stream centerline, generating a distance raster. A gradual sinking effect was applied using the Raster Calculator (Image Analyst Tools), where stream cells were lowered in elevation based on their proximity to the centerline using the conditional statement:

$$Con(\text{“}Stream\_Raster\text{”} == 1, \text{“}DEM\text{”} - (6 \,/\, (1 + \text{“}Distance\_Raster\text{”})), \text{“}DEM\text{”}).$$

This tapering approach ensured a smooth transition between the sunken channels and surrounding terrain, following the principles of natural channel representation, as seen in Figure 6 (Chen et al., 2024). The choice of six as the scaling factor is based on the water table depth of the study area, which ranges from 4.572 to 6.096 meters (Hardy, N.d).

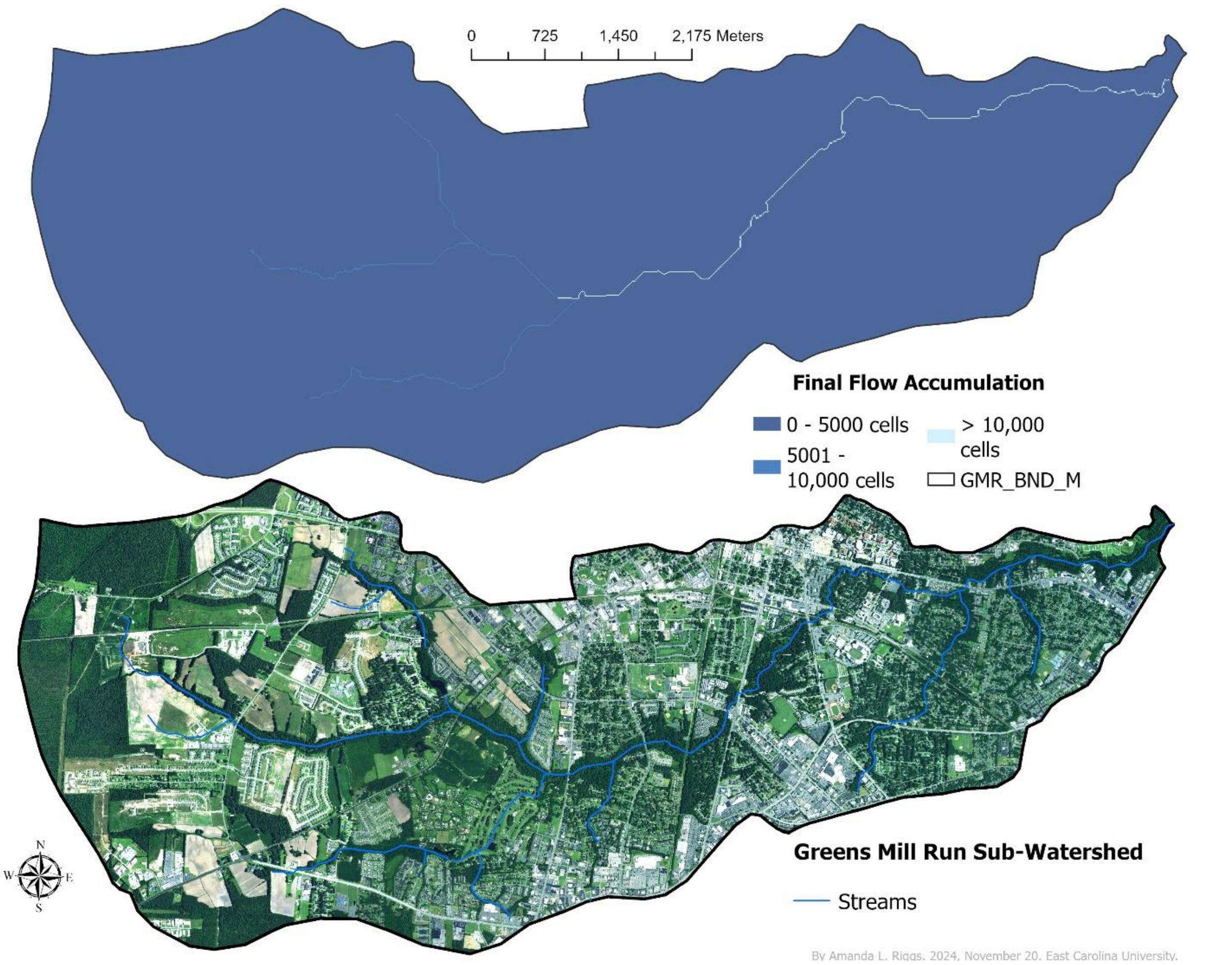


Figure 6. A map of the final flow accumulation raster overlaid on 2016 imagery, with stream vector lines for validation of stream alignment and flow paths.

The final conditioned DEM was validated by re-running the Flow Direction (Spatial Analyst Tools) and Flow Accumulation (Spatial Analyst Tools) tools, with both using the D8 algorithm and the default settings, to confirm hydrological connectivity. Stream profiles were visually inspected to ensure they exhibited realistic depressions consistent with natural topography as well as visual inspection of stream elevation differences around bridges and downstream to ensure elevations consistently flow under bridges and downstream as per USGS guidelines (USGS, 2024). Additional validation metrics and techniques are presented in the subsequent section.

4 Results

The workflow began with the hydro-flattening of streams in the original DEM, setting the stage for subsequent enhancements. Following USGS guidelines for both aesthetic and educational purposes (NGP Standards and Specifications, 2024), the process then progressed to condition and enforce the DEM hydrologically. This initial hydro-flattening effectively removed artifacts while smoothing and flattening the streams, creating a more seamless representation.

Next, the DEM was hydro-conditioned by adjusting the elevations within the streams. This step involved aligning the alterations with the natural flow direction of the streams, ensuring that water flows continuously across the terrain by eliminating any sinks. To finalize the process, the DEM underwent hydro-enforcement, where predicted streams were 'burned' into the model. This step involved lowering the elevations of these streams, which enabled the water to flow downhill realistically and underneath structures, adhering to the specifications outlined by the USGS hydro-flattening reference (NGP Standards and Specifications, 2024).

The result of these comprehensive steps is a hydrologically flattened, conditioned, and enforced DEM. This final product illustrates effective hydrological processing, with widened, flatter, and smoother streams alongside a reduction of artifacts, as shown in Figure 7. To further enhance clarity, both DEMs illustrated in the accompanying map employ consistent color schemes, transparency settings, and layer order, to highlight the differences. A series of validation measures were also conducted to assess the robustness and accuracy of both the workflow and the final processed DEM.

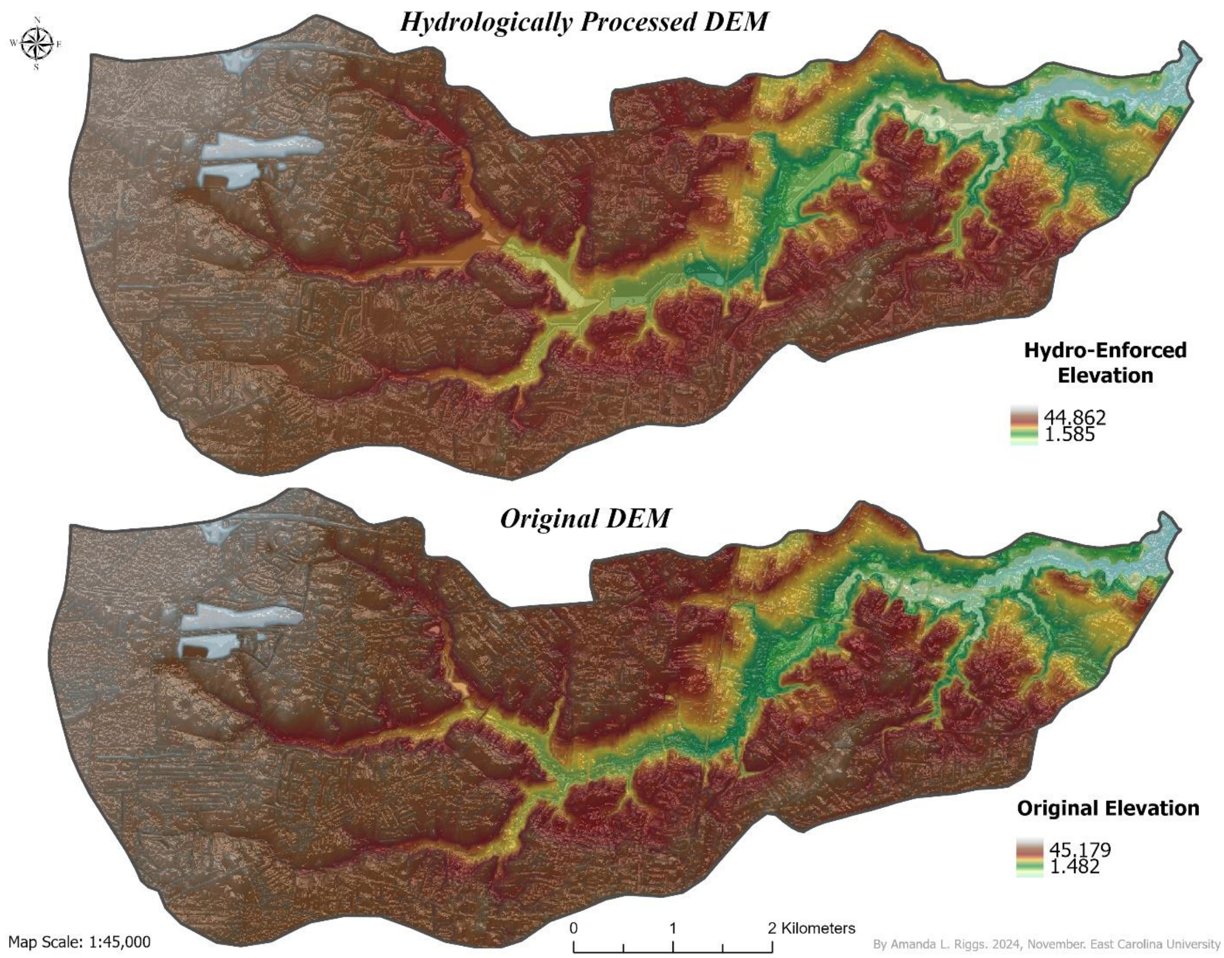


Figure 7. Map comparing the hydro-conditioned DEM to the original Dem. The color schemes, transparency settings, and layer order have been adjusted to clearly highlight the differences.

4.1 Validation

Validation started with the confirmation of the hydro-conditioned Digital Elevation Model (DEM) alignment and connectivity through the application of the Flow Direction and Flow Accumulation tools from the Spatial Analyst suite, as outlined in the methods section, to ensure that the DEM effectively routed water flow and established continuous stream networks. A contour layer was generated for visual validation using the Contour tool, which facilitated a comparative analysis with the original DEM contours to illustrate the effects of terrain smoothing while also verifying the successful removal of triangulation artifacts that had emerged during

DEM processing (Terziotti, 2023). Subsequent sections will discuss additional details on the validation measures adopted from various sources.

4.1.1 Overlap Raio (OR)

The next validation method, derived from Chen et al. (2024), employs the Overlap Ratio (OR) and Intersection over Union (IoU) techniques to assess the alignment between the predicted stream binary layer from the Stream Extraction section of the workflow and the referenced stream centerline vector layer. This evaluation is necessary for understanding how well these two datasets correspond with one another. To assess OR and IoU, the validation workflow incorporated the Intersection (Analyst Tools) tool to compute the geometric intersection of the streams, while the Union (Analyst Tools) tool was employed to determine the geometric union of the predicted and referenced streams.

The validation measure began by converting the predicted stream network from a raster format into a polyline feature layer using the Raster to Polyline (Conversion Tool). This conversion enabled a direct comparison with the reference stream network. Initially, when assessing the overlap between the two datasets using the Intersect (Analyst Tool), the tool generated empty outputs. To correct the output, the Buffer (Analysis Tool) was used to create a one-meter buffer for both the predicted and reference stream networks using the Buffer (Analysis Tool) to address the issue by expanding the streams' spatial footprint to increase the chance of capturing overlaps between the two datasets. The default parameters of the Buffer tool were used in processing, including a full side type to create buffers on both sides of the polylines, a rounded end type, and the option not to dissolve, ensuring that each feature maintained its buffer independently of any overlaps. Given that the data was in a projected coordinate system (UTM Zone 18N), the default planar method was retained, which allowed the Euclidean buffers to represent straight-line distances on a flat surface accurately. By preprocessing both the predicted and reference stream polylines with identical parameters, the tool captured the overlapping geometries.

Once the stream networks were buffered, the Intersect (Analyst Tool) ran successfully, resulting in a layer that contained the overlapping segments of the predicted and reference streams. The next step was to utilize the Union (Analyst Tools) to analyze these overlaps further. Next, the workflow incorporated the Summary Statistics (Analysis Tools) tool to generate a summary of the length field of the intersected features to calculate the total length of overlapping streams, focusing on the Length field and selecting the statistic type "Sum." This same methodology was employed to determine the total length of the reference streams. The resulting statistics summaries were key in validating the stream burning process, using the intersection outputs (length) derived from the two-accuracy metrics. Figure 8 shows comparisons between the predicted and reference streams, as well as the intersections of each.

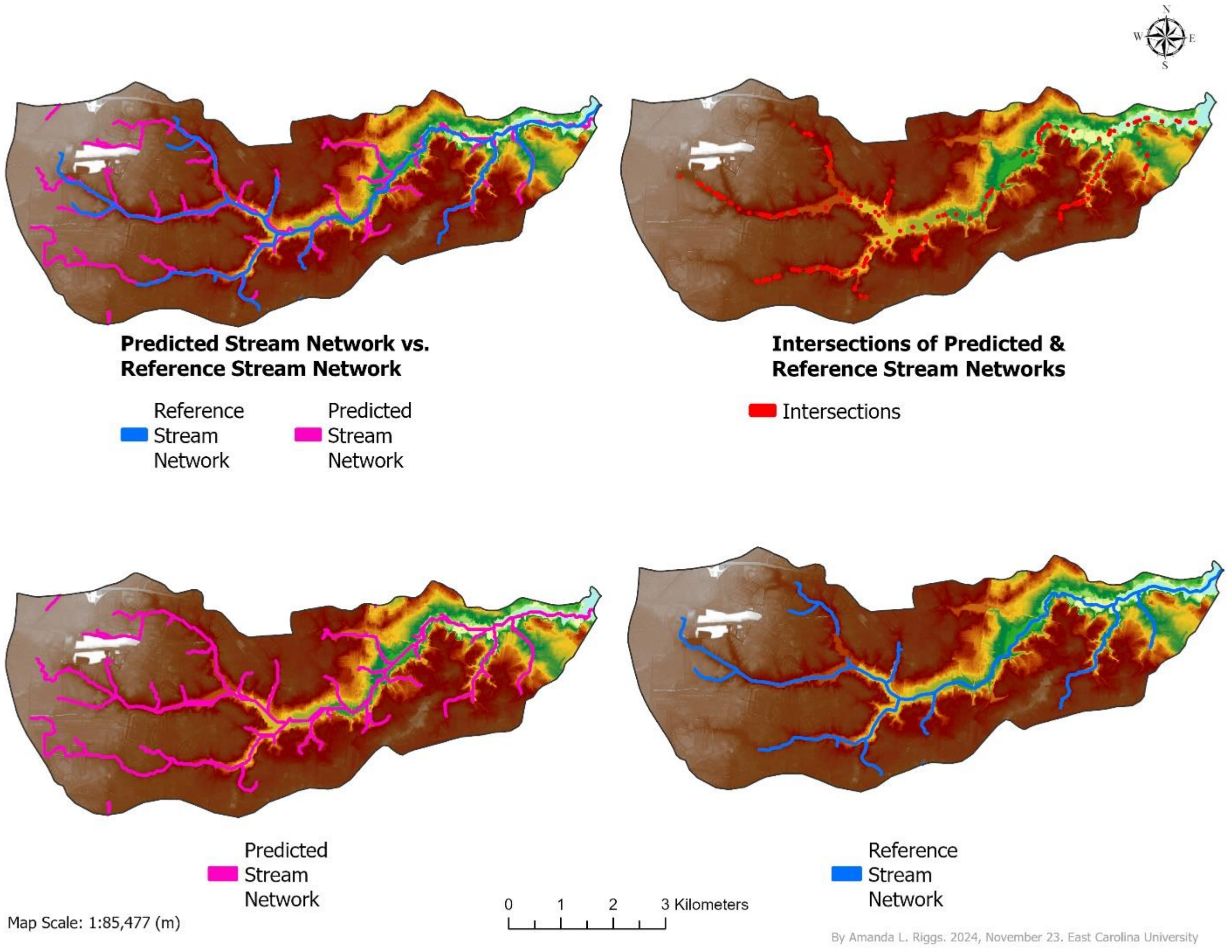


Figure 8. Comparisons between the predicted and reference streams, along with the intersections of the combined streams.

The Overlap Ratio (OR) provides a quantitative measure of the proportion of the extracted stream network that overlaps with the channel label, producing insights into the effectiveness of the stream extraction process.

The equation for OR is:

$$OR = \frac{Length\ of\ Overlapping\ Streams}{Length\ of\ Reference\ Streams}$$

Equation 1. Formula for Overlap Ratio (OR).

For the predicted dataset, the calculation yielded:

$$OR = \frac{4645.03}{53573.68} = 0.09$$

Equation 2. Calculation of OR for the predicted dataset based on the threshold of 24,000 pixels.

An OR greater than one signifies that the predicted stream network is longer than the reference network in areas where they overlap. This situation may arise from over-prediction, where the predicted network includes streams not found in the reference dataset, or it may reflect greater detail in the predicted network, such as the inclusion of smaller tributaries absent from the reference. An OR less than one indicates that the overlapping portion of the predicted network is smaller relative to the overall length of the reference network.

4.1.2 Intersection over Union (IoU)

The next validation measure is the Intersection over Union (IoU) metric, which describes the spatial overlap between the predicted and reference stream networks.

The equation for IoU is:

$$IoU = \frac{Area\ of\ Overlap\ (Intersection)}{Area\ of\ Union}$$

Equation 3. Formula for Intersection over Union (IoU).

For the predicted dataset, the calculation is:

$$IoU = \frac{2559.80}{154720.46} = 0.02$$

Equation 4. Calculation of IoU for the predicted dataset based on the threshold of 24,000 pixels.

The numerator represents the overlapping area between the predicted and reference networks, while the denominator is the total area covered by either dataset. A high IoU indicates strong spatial alignment. A low IoU suggests differences between the networks, potentially occurring from over-prediction and the general detail of the reference data.

For the predicted dataset, the IoU was 0.02, reflecting minimal spatial overlap. This result aligns with the known limitations of the reference dataset, which lacks smaller tributaries that the predicted network caught using the flow accumulation raster derived from the flow direction and the original DEM. While a low IoU typically suggests poor alignment, it can also highlight additional detail captured by the predicted network, including unmapped tributaries.

4.1.3 Kappa Statistics

Cohen's Kappa statistics, as described in the methods outlined by Chen et al. (2024), serve as a valuable validation measure for assessing agreement between two datasets. This statistic not only quantifies the level of agreement but also considers the possibility that this agreement could occur purely by chance. It is particularly valuable in evaluating how well a predicted dataset aligns with a reference dataset (Nicolau et al., 2024).

To assign Kappa metrics for evaluation, the Raster Calculator (Spatial Analyst Tools) was used with a conditional formula to create the kappa statistics raster, assigning values to each cell based on its classification as a True Positive (TP), True Negative (TN), False Positive (FP), or False Negative (FN) as shown in Figure 9. This is done by comparing the predicted stream raster and the reference stream raster.

```
Con((Predicted == 1) & (Reference == 1), 1,  // TP
    Con((Predicted == 0) & (Reference == 0), 2,  // TN
        Con((Predicted == 1) & (Reference == 0), 3,  // FP
            Con((Predicted == 0) & (Reference == 1), 4,  // FN
                0))))  // Default value for unclassified cells
```

Figure 9. Conditional formula to assign Kappa metrics to the predicted and reference streams.

The Kappa statistics are evaluated by constructing a confusion matrix to evaluate stream extraction methods. This matrix facilitates a comparison between the predicted stream network and a reference stream network. Several metrics lie within the confusion matrix: True Positives (TP) represent the predicted streams that correctly correspond to the reference streams; False Positives (FP) indicate predicted streams that do not match any streams in the reference dataset; False Negatives (FN) denote the reference streams that not identified in the predicted streams; and True Negatives (TN) reflect the areas accurately recognized as non-streams in both the predicted and reference datasets.

The formula for Kappa is defined as follows:

$$Kappa\ Coeficient\ = \frac{observed\ agreement - expected\ agreement}{1 - expected\ agreement}$$

$$K = \frac{P_o - P_e}{1 - P_e}$$

Equation 5. Formula to calculate Cohen's Kappa coefficient.

Where:

P_o is the observed agreement between the two rasters.

$$P_o = \frac{TP + TN}{Total\ Cells} = \frac{101 + 1484852}{2991552} = 0.496382145$$

Equation 6. Formula to calculate Po, the observed agreement between two rasters.

P_e is the expected agreement by chance, calculated based on the marginal probabilities of each class.

$$Pe = \frac{(9331 \cdot 1694)}{2{,}991{,}5522^2} + \frac{(1{,}486{,}445 \cdot 1{,}494{,}082)}{2{,}991{,}5522^2} = 0.2502$$

Equation 7. Formula to calculate Pe, the expected agreement by chance, based on the marginal probabilities of each class.

Kappa :

$$K = \frac{P_o - P_e}{1 - P_e} = \frac{0.4964 - 0.2502}{1 - 0.2502} = \frac{0.2462}{0.7498} \approx 0.3283$$

Equation 8. Calculation of Cohen's Kappa coefficient based on the predicted stream network threshold of 24,000.

A kappa value of approximately 0.33 indicates a fair level of agreement between the predicted and reference stream datasets, highlighting the challenges in aligning predicted streams with a generalized reference dataset.

The imbalance within the dataset, featuring 1,494,082 non-stream pixels contrasted with just 1,694 stream pixels, further contributes to the modest kappa value. While the observed agreement (P_o = 0.496) suggests that nearly 50% of the predictions are in line with the reference, the expected agreement by chance (P_e = 0.250) is notably high due to the overwhelming presence of non-stream pixels, limiting the Kappa results to a level of "fair" agreement.

The results indicate that the predicted dataset does capture some stream features accurately; however, it also includes a substantial number of false positives (9,230) and misses some streams entirely (1,593 false negatives), which affects overall alignment. Further examination through metrics such as precision (1.08%), recall (5.96%), and F1-score (1.83%) underscore these issues. The low precision suggests a significant number of false positives, while the low recall reflects missed stream data.

To further investigate Kappa's threshold sensitivity, a second stream network was generated using flow accumulation with a higher threshold of 30,000, designed to exclude smaller tributaries. However, the sensitivity results revealed that this adjustment had no impact on the Kappa statistics even while removing the smaller tributaries, with both Kappa values remaining at 0.329. Table 1 shows comparison tables for each test.

Kappa Statistic Benchmarks (Nicolau et al., 2024)	
Kappa Range	**Interpretation**
0.01–0.20	Slight agreement
0.21–0.40	Fair agreement
0.41–0.60	Moderate agreement
0.61–0.80	Substantial agreement
0.81–1.00	Almost perfect agreement

Zonal Statistics: Predicted Streams		
OBJECTID	**Value**	**COUNT**
True Positives	1	101
True Negatives	2	4615
False Positives	3	531
False Negatives	4	371213

Zonal Statistics: Kappa Sensitivity		
OBJECTID	**Value**	**COUNT**
True Positives	1	235
True Negatives	2	394954
False Positives	3	1229
False Negatives	4	3339

Kappa Matrix: Predicted Streams			
	Reference Stream (1)	**Reference Non-Stream (0)**	**Total**
Predicted Stream (1)	101 (TP)	9230 (FP)	9331
Predicted Non-Stream (0)	1593 (FN)	1484852 (TN)	1486445
Total	1694	1494082	2991552

Kappa Matrix: Kappa Sensitivity			
	Reference Stream (1)	**Reference Non-Stream (0)**	**Total**
Predicted Stream (1)	253 (TP)	1229 (FP)	1482
Predicted Non-Stream (0)	3339 (FN)	394954 (TN)	398293
Total	3592	396183	799550

Kappa Metrics: Predicted Streams		
Metric	**Formula**	**Value**
Precision	$\frac{TP}{TP + FP}$	0.1597
Recal (Sensitivity)	$\frac{TP}{TP + FN}$	0.0003
Accuracy	$\frac{TP + TN}{Total\ Cells}$	0.0016
F1-Score	$2 * \frac{Precision * Recall}{Precision + Recall}$	0.0005
Kappa	$\frac{P_o - P_e}{1 - P_e}$	0.3285

Kappa Metrics: Kappa Sensitivity		
Metric	**Formula**	**Value**
Precision	$\frac{TP}{TP + FP}$	0.1707
Recal (Sensitivity)	$\frac{TP}{TP + FN}$	0.0704
Accuracy	$\frac{TP + TN}{Total\ Cells}$	0.4943
F1-Score	$2 * \frac{Precision * Recall}{Precision + Recall}$	0.0997
Kappa	$\frac{P_o - P_e}{1 - P_e}$	0.3285

Table 1. Comparison of Kappa statistics between the predicted stream of 24,000 pixels vs the predicted stream of 30,000 pixels, including Kappa benchmarks, zonal statistics, Kappa matrix, and Kappa results.

4.2 Discussion

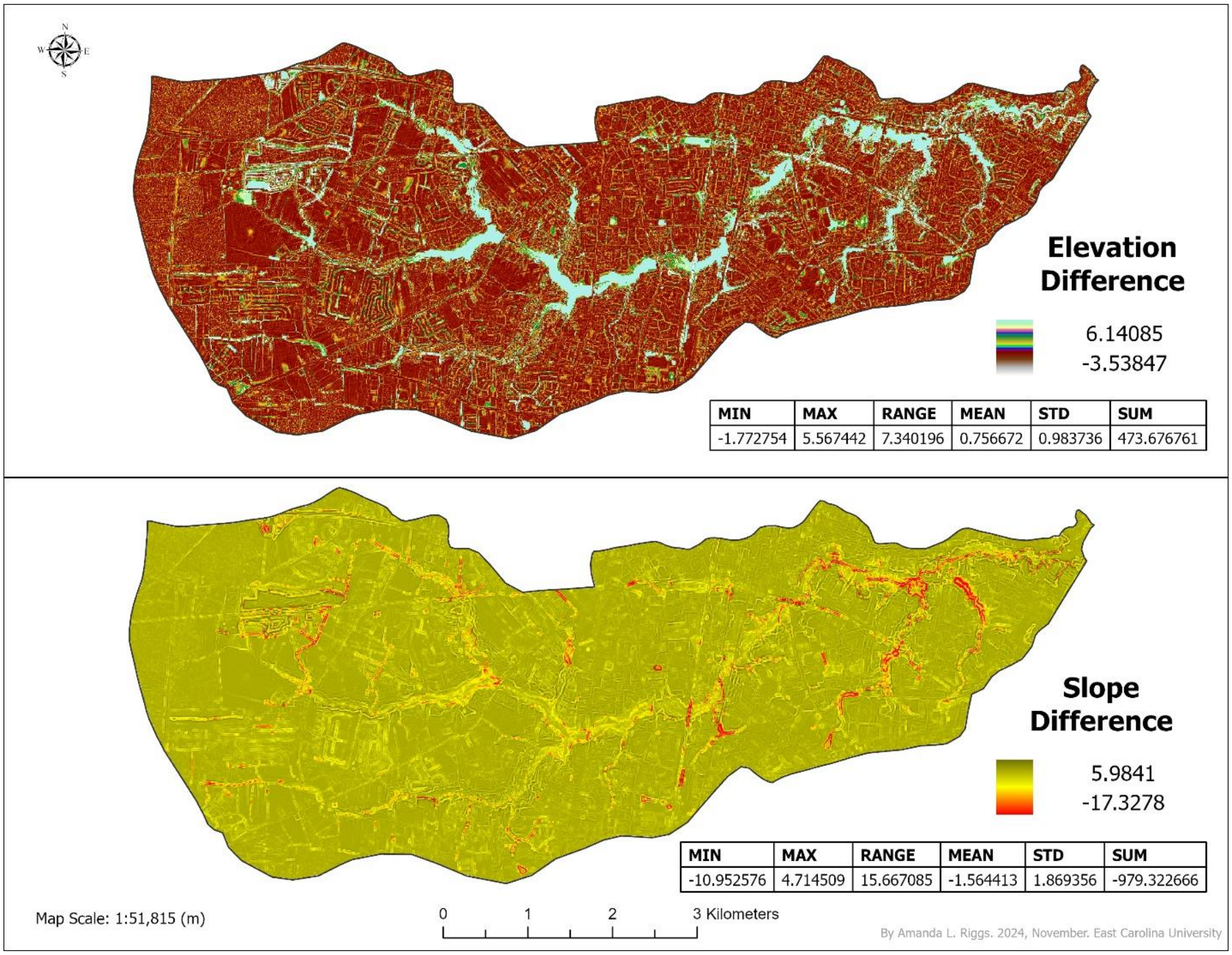
The alignment of the Flow Direction and Flow Accumulation outputs confirms that the hydro-conditioned DEM is hydrological sound, with continuous and connected stream networks accurately routing water flow, demonstrating that core hydrological processes are functioning correctly, despite discrepancies with the reference dataset. Similarly, the generation of contours validated that hydro-conditioning removed triangulation artifacts and smoothed the terrain without introducing errors, ensuring natural water flow representation (Terziotti, 2023). Together, these validations confirm that the DEM modifications preserved hydrological accuracy while addressing topographic artifacts.

The Overlap Ratio (OR) and Intersection over Union (IoU) metrics provide additional insights into the alignment between the predicted and reference stream networks. The OR of 0.09 indicates that only 9% of the reference stream network overlaps with the predicted streams, highlighting potential gaps in the predicted dataset or generalization in the reference. Similarly, the IoU of 0.02 reflects minimal spatial agreement, emphasizing large areas of false positives or negatives in stream delineation (Chen et al., 2024).

These metrics suggest that the predicted network captured details such as smaller tributaries absent from the generalized reference, but they also underscore challenges in stream alignment due to restrictive or permissive thresholds. While the low OR and IoU indicate room for refinement, they also reveal the limitations of the reference dataset (Chen et al., 2024). The processed stream network, derived from high-resolution DEM data, remains a more localized and accurate representation of the study area's hydrology.

A second stream network generated with a higher threshold (30,000 pixels) to exclude smaller streams produced no significant improvement in alignment, as the Kappa value (0.329) remained unchanged. This lack of improvement highlights the limitations of the reference data and suggests that both thresholds resulted in similar patterns of over- and under-prediction.

Finally, reflecting on Chen et al.'s (2024) recommendation to evaluate the impact of hydrological processing on elevation, maps were included displaying both elevation and slope differences, accompanied by their respective statistics as seen in Figure 10.



MIN	MAX	RANGE	MEAN	STD	SUM
-1.772754	5.567442	7.340196	0.756672	0.983736	473.676761

MIN	MAX	RANGE	MEAN	STD	SUM
-10.952576	4.714509	15.667085	-1.564413	1.869356	-979.322666

Figure 10. Maps displaying both the elevation and slope differences along with their respective statistics.

The elevation differences show a range from -1.77 m to 5.57 m, with a mean difference of 0.756 m, indicating consistent adjustments to align the terrain with hydrological flow requirements. Similarly, the slope differences range from -10.95 to 4.71 degrees, with a mean of -1.56 degrees, highlighting areas where gradient modifications were necessary to ensure accurate water routing. These results align with Chen et al.'s emphasis on examining how terrain adjustments preserve hydrological and topographic integrity, demonstrating that the hydro-conditioning process effectively balanced elevation modifications with functional hydrological improvements but did alter the original terrain of the original DEM.

Future improvements could include refining the reference dataset using higher-resolution or localized data to better capture smaller streams. Enhanced validation metrics, such as root mean square error (RMSE) or advanced spatial alignment measures, may also provide more comprehensive assessments. Additionally, integrating land use data could help refine thresholds and improve stream delineation in complex terrains, and after refinement, further steps could include automating the workflow. These future improvements would not only enhance the accuracy and reliability of stream delineation but also streamline the process, enabling more precise and efficient analyses in diverse environmental contexts.

5 References

ArcGIS Enterprise. (2024). Flow direction function. *ESRI*. https://enterprise.arcgis.com/en/image/latest/raster-analytics/flow-direction-raster-function.htm

ArcGIS REST Services. (2024, April 2). FindExistingLocationsOutput [Data]. *ESRI*. https://services2.arcgis.com/sJvSsHKKEOKRemAr/arcgis/rest/services/Roads_Greenville/FeatureServer

Boyd, Doug. (2015, October 01). Cleaner by nature project slows filters stormwater begore it reaches stream. News Services. East Carolina University. https://news.ecu.edu/2015/10/01/cleaner-by-nature/

Census Bureau. (2022). TIGER/Line shapefile, 2022, nation, U.S., current County and equivalent national [Data]. *U.S. Department of Commerce, U.S. Census Bureau, Geography Division, Spatial Data Collection and Products Branch*. https://www2.census.gov/geo/pdfs/maps-data/data/tiger/tgrshp2022/TGRSHP2022_TechDoc.pdf

Chen, Q., Mudd, S. M., Attal, M., & Hancock, S. (2024, July 29). Extracting an accurate river network: Stream burning re-visited. Remote sensing of environment. 312 114333. https://doi.org/10.1016/j.rse.2024.114333

Crawford, C. (2022, July 28). Hydro-flattening of river shorelines in lidar based DEM production. ESRI. https://www.esri.com/arcgis-blog/products/arcgis-pro/3d-gis/hydro-flattening-of-river-shorelines-in-lidar-based-dem-production/

ESRI. (2023, July 27). Create a watershed model using hydrology in ArcGIS Pro. ArcGIS Pro documentation. ESRI. https://support.esri.com/en-us/knowledge-base/how-to-create-a-watershed-model-using-hydrology-in-arcg-000023169

Greenville North Carolina. Nd. About Greenville. Greenville North Carolina. https://www.greenvillenc.gov/our-city/about-greenville

Hazen and Sawyer. (2016, August). City of Greenville Greens Mill Run watershed master plan. Hazen No. 31187. Hazen. greenvillenc.gov/home/showpublisheddocument/13142/636204126553570000

Lehner, B., Verdin, K., & Jarvis, A. (2008). New global hydrography derived from spaceborne elevation data [Data]. *Eos Transactions*, *AGU*. *89*(10). Pp. 93-94.

National Hydrography. (2023). Watershed Boundary Dataset (WBD) [Data]. *U.S. Department of Agriculture, Natural Resources Conservation Service*. Retrieved from National Hydrography Dataset | U.S. Geological Survey

Natural Capital Project. 2024. InVEST 0.0. *Stanford University, University of Minnesota, Chinese Academy of Sciences, The Nature Conservancy, World Wildlife Fund, Stockholm Resilience Centre and the Royal Swedish Academy of Sciences*. https://naturalcapitalproject.stanford.edu/software/invest

NGP Standards and Specifications. (2024). Lidar based specification appendix 2: Hydro-flattening reference. Lidar base specification 2024 rev. A. *U.S. Geological Survey*. Lidar Base Specification Appendix 2: Hydro-flattening Reference | U.S. Geological Survey

Nicolau, A.P., Dyson, K., Saah, D., Clinton, N. Cardille, J.A., Crowley, M.A., Saah, D., Clinton, N.E. (eds) (2024). Accuracy Assessment: Quantifying Classification Quality. *Cloud-Based Remote Sensing with Google Earth Engine. Springer, Cham*. https://doi.org/10.1007/978-3-031-26588-4_7

Hardy, A. V. (N.d). General nature of the county. *Pitt County, NC.gov*. https://www.pittcountync.gov/DocumentCenter/View/6424/General-Nature-of-the-County?bidId=

OCM Partners. (2024). 2016 Coastal North Carolina NAIP Digital Ortho Photo Imagery from 2010-06-15 to 2010-08-15 [Data]. *NOAA National Centers for Environmental Information*. https://www.fisheries.noaa.gov/inport/item/49531.

OpenAI. (2024). ChatGPT (November 15 version) [Large language model]. https://chat.openai.com/chat

Stoker, K. & Miller, B. (2022, February 15). The accuracy of and consistency of 3D elevation program data: A systematic analysis. *Environmental Monitoring and Mapping Using 3D Elevation Program Data. 14*(4). 940. https://www.mdpi.com/2072-4292/14/4/940

United States Geological Survey (2021). United States Geological Survey 3D Elevation Program 1/3 arc-second Digital Elevation Model [Data]. *Distributed by OpenTopography*. https://doi.org/10.5069/G98K778D. Accessed: 2024-11-14

U.S. Census Bureau. (2022, August. Updated 2024, January 27). TIGER/Line shapefile, 2022, County, Pitt County, NC, area hydrology [Data]. *U.S. Department of Commerce*. TIGER/Line Shapefile, 2022, County, Pitt County, NC, Area Hydrography - Catalog (data.gov)