Assessing the Effect of Wetland Connectivity on In-Stream Nitrate Concentrations Using an Integrated Wetland-River Network

Heather Wall

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science
In
Biological Systems Engineering

Jonathan A. Czuba, PhD, Chair
Daniel L. McLaughlin, PhD
Durelle T. Scott, PhD

May 12, 2025
Blacksburg, Virginia

Keywords: nitrate removal, wetlands, watershed, rivers, spatial distribution

Assessing the Effect of Wetland Connectivity on In-Stream Nitrate Concentrations Using an Integrated Wetland-River Network

Heather Wall

ABSTRACT

Nitrate degrades water quality from the local to continental scale. Excess nitrate can cause eutrophication and hypoxia. A hypoxic zone exists in the Gulf of Mexico that causes organism death and displacement. One of its primary sources of nitrate is the agriculture region in the Mississippi River Basin. The location for this study is the Minnesota River Basin which is part of the Mississippi River Basin. Excess fertilizer from corn and soybean production runs off the landscape to the waterways. Nitrate is naturally removed from waterbodies through denitrification with wetlands being the most effective due to their long residence time. Wetland restoration has the potential to remove nitrate from small portions of the landscape. The question is how to place wetlands to most effectively remove nitrate across a watershed. There are multiple models and studies looking at this and they need an integrated wetland-river network; however, the common sources of information for wetlands and waterways are the National Wetlands Inventory and the National Hydrography Dataset, respectively. These were not inherently made to be compatible. A method was created to combine these two datasets, which outputs an integrated wetland-river network complex that can be used for further analysis. For this method, floodplain wetlands were first removed and both flowlines and wetlands were assigned unique identifiers and information on how they connect to allow routing through the wetland-river network. A reliable way to predict nitrate is important as sampling and laboratory analysis to determine nitrate concentrations is expensive and time consuming. Two common metrics related to nitrate concentrations are percent wetland area and interception fraction. Interception fraction takes into consideration the spatial placement of wetlands by comparing the area of the watershed intercepted by wetlands to the total watershed area. These metrics were calculated from the integrated wetland-river network and compared to nitrate samples from June 2014 and 2015. Interception fraction was found to be a better predictor of nitrate concentration than percent wetlands and that the percent of non-intercepted cropland and the size of the wetland that intercepts most of the watershed were not useful in understanding the residuals. This work will allow for the more rapid prediction of nitrate watershed exports to inform where wetlands should be restored to reduce nitrate exports into waterways.

Assessing the Effect of Wetland Connectivity on In-Stream Nitrate Concentrations Using an Integrated Wetland-River Network

Heather Wall

GENERAL AUDIENCE ABSTRACT

Nitrogen degrades water quality from the local to continental scale. Excess nitrogen can cause organism death and displacement as seen in the Gulf of Mexico. One of its primary sources of this nitrogen is the agriculture region in the Mississippi River Basin. The location for this study is the Minnesota River Basin which is part of the Mississippi River Basin. Excess fertilizer from corn and soybean production runs off the landscape to the waterways. Nitrogen is naturally removed from waterbodies. Rivers remove nitrate slowly, but wetlands are more effective due to their capacity to hold water for longer periods of time. Multiple studies have found that wetland restoration has the potential to remove nitrogen from small portions of the landscape. The question is how to place wetlands to most effectively remove nitrogen across a watershed. Multiple models and studies have begun looking at this. They all need an integrated wetland river network; however, the common sources of information for wetlands and waterways are the National Wetlands Inventory and the National Hydrography Dataset, respectively. These were not inherently made to be compatible. An automated method was created to combine these two datasets, which outputs an integrated wetland-river network that can be used for further analysis. A reliable way to predict nitrogen is important as sampling and laboratory analysis to determine nitrogen concentrations is expensive and time consuming. Two common metrics related to nitrate concentrations are percent wetland area and interception fraction. Interception fraction takes into consideration the spatial placement of wetlands by comparing the area of the watershed intercepted by wetlands to the total watershed area. These metrics were calculated from the integrated wetland-river network and compared to nitrate samples from June 2014 and 2015 . Interception fraction was found to be a better predictor of nitrate concentration than percent wetlands and that the percent of non-intercepted cropland and the size of the wetland that intercepts most of the watershed were not useful in understanding the residuals. This work will allow for the more rapid prediction of nitrate watershed exports to inform where wetlands should be restored to reduce nitrate exports into waterways.

Acknowledgements

Thank you to my advisor, Dr. Jon Czuba for your assistance, support, guidance, and belief in my ability to do well in the research and for helping me find a topic that would let me continue to use my GIS and Python skills. Thank you to my committee members Dr. Durelle Scott and Dr. Danial McLaughlin for your support and interest. Thank you to Dr. Amy Hansen for taking the time digging up notes on your previous research to help me with mine. Thank you to Reilly Oare for being a homework buddy and for being a good friend. Thank you to Katie Wardiniski for introducing me to the field of research and providing a wonderful first time in the field despite the cold, rain, and snow. Thank you to my friends and family, who may not completely understand what I am researching, but still love and support me.

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1 Impact of Nitrogen

River, lake, and coastal water quality has been degraded by excess nitrate at local, regional, and continental scales. The nitrogen cycle describes the process through which nitrogen moves through the environment, changing from inorganic to organic forms to be used by different organisms. This cycle has been affected by humans through anthropogenic methods, such as the Haber-Bosh process, that have increased the availability of nitrogen for biological use (Vitousek et al., 1997, Martínez-Espinosa et al., 2021). One of the primary uses of anthropogenically created reactive nitrogen is agricultural practices due to the development of synthetic fertilizers. Excess fertilizer is transported in runoff to waterways, providing a major source of nitrogen to river systems (Randell and Mulla, 2001, Hansen et al., 2018).

Nitrogen is a primary cause of eutrophication in waterbodies (Martínez-Espinosa et al., 2020) which can lead to hypoxic conditions (Vitousek et al., 1997) such as seen in the Gulf of Mexico (Cheng et al., 2020). This hypoxic zone varies in size with the amount of nitrate carried by freshwater that is discharged by the Mississippi River. This zone can cause displacement and mortality of fish and can kill burrowing organisms if it remains too long. After the hypoxic zone is gone, secondary production is reduced (Rabalais and Turner, 2019). The Gulf of Mexico hypoxic zone's main source of nitrogen is from agriculture in the Upper Mississippi River Basin (Rabalais and Turner, 2001; David et al., 2010). The land in this watershed is primarily used for agriculture to produce corn and soybeans (David et al., 2010). The Hypoxia Task Force is working to reduce nitrogen and phosphorous entering the Gulf of Mexico, with a goal of 48% reduction by 2035 for both based on the average nutrient loading from 1980 to 1996 (EPA, 2023). The interim goal of 20% reduction by 2025 has been met for nitrogen but phosphorous has increased by 3% (EPA, 2023).

1.2 Importance of Wetlands

Historically wetlands covered about 221 million acres of land in the conterminous United States (Dahl and Allord, 1996). To support a rapidly growing population, many wetlands were ditched, drained, and removed to make way for agriculture (Dahl and Allord, 1996). In Minnesota most of the land within the Minnesota River Basin has been converted to cropland (Frans, 2013). This facilitates rapid movement of water from agriculture fields to streams due to artificial systems that work to optimize crop growth through excess water removal (Foufoula-Georgiou et al., 2013).

Wetlands are known to remove nitrate through assimilation and denitrification. Denitrification is the process in which nitrate is converted to atmospheric nitrogen, playing a major role in the nitrogen cycle (Knowles, 1982). Wetlands are known to be more effective than streams at removing nitrate due to longer residence times. Rivers transport more nutrients within higher flow conditions. This causes a rapid increase in nutrients that the system is unable to retain (Wollheim et al., 2018). Wetland restoration is generally agreed upon in literature to have the potential to reduce nitrate within a small portion of the landscape (Hansen et al., 2018). A study in northeast Spain found that wetlands can be used as a buffer zone in agriculture regions to remove nitrates (Darwiche-Criado et al., 2015). In North Carolina wetlands are being explored to reduce nitrogen released from agriculture production to assist in improving the quality of shellfish waters (Messer et al., 2017). Additionally, a study on the Ipswich River in Massachusetts showed that lakes and wetlands along the network increase the amount of nutrients removed compared to a network with none (Wollheim et al., 2018).

1.3 Wetlands and River Networks in Modeling

Wetlands are known to remove nitrate, but how they can be placed to most effectively remove nitrate is still being explored. Different models have been developed to explore how wetland location within a watershed can impact downstream nitrate concentrations. The Nitrate Network Model is a model that estimates the export of nitrate from the watershed and computes nitrate-nitrogen and organic carbon concentrations within a wetland-river network (Czuba et al., 2018). The goal for this model is for it to be used to determine the specifications of the wetland and where to restore it to optimize environmental benefits in the watershed (Czuba et al., 2018). Another model with similar goals is the AgRiver model which integrates the Nitrate Network Model, Management Option Simulation Model, and the Soil and Water Assessment Tool (Hansen et al., 2021). All of these models work to determine the best way to improve water quality. Another model, InVEST®, is for looking at how ecosystem changes can impact people around the world (Natural Capital Project, 2025).

One key factor needed for these models to work well for assessing nitrate removal is an integrated river and wetland network. Geospatial wetland data is most easily acquired from the

National Wetland Inventory. Wetlands in the National Wetland Inventory are acquired from high altitude imagery and classified based on vegetation, geography, and visible hydrology (U.S. Fish and Wildlife Service, 2019). Flowline or river data is most easily acquired from the National Hydrography Dataset Plus Version 2. These data are prepared through the use of digital elevation models and imagery (EPA, 2019). These two datasets are not directly compatible. This is clearly seen through the riverine wetlands in the National Wetland Inventory that represent rivers and streams, but the flowlines from the National Hydrography Dataset do not match them (Figure 1). There is limited waterbody representation within the National Hydrography Dataset that represents lakes, therefore, if the network is broken in any way to accommodate wetlands, new network topology would have to be created to properly route water and nitrate through the network. To calculate any watershed scale relationships related to nitrate, a network must incorporate wetlands since they play an important role in nutrient removal.

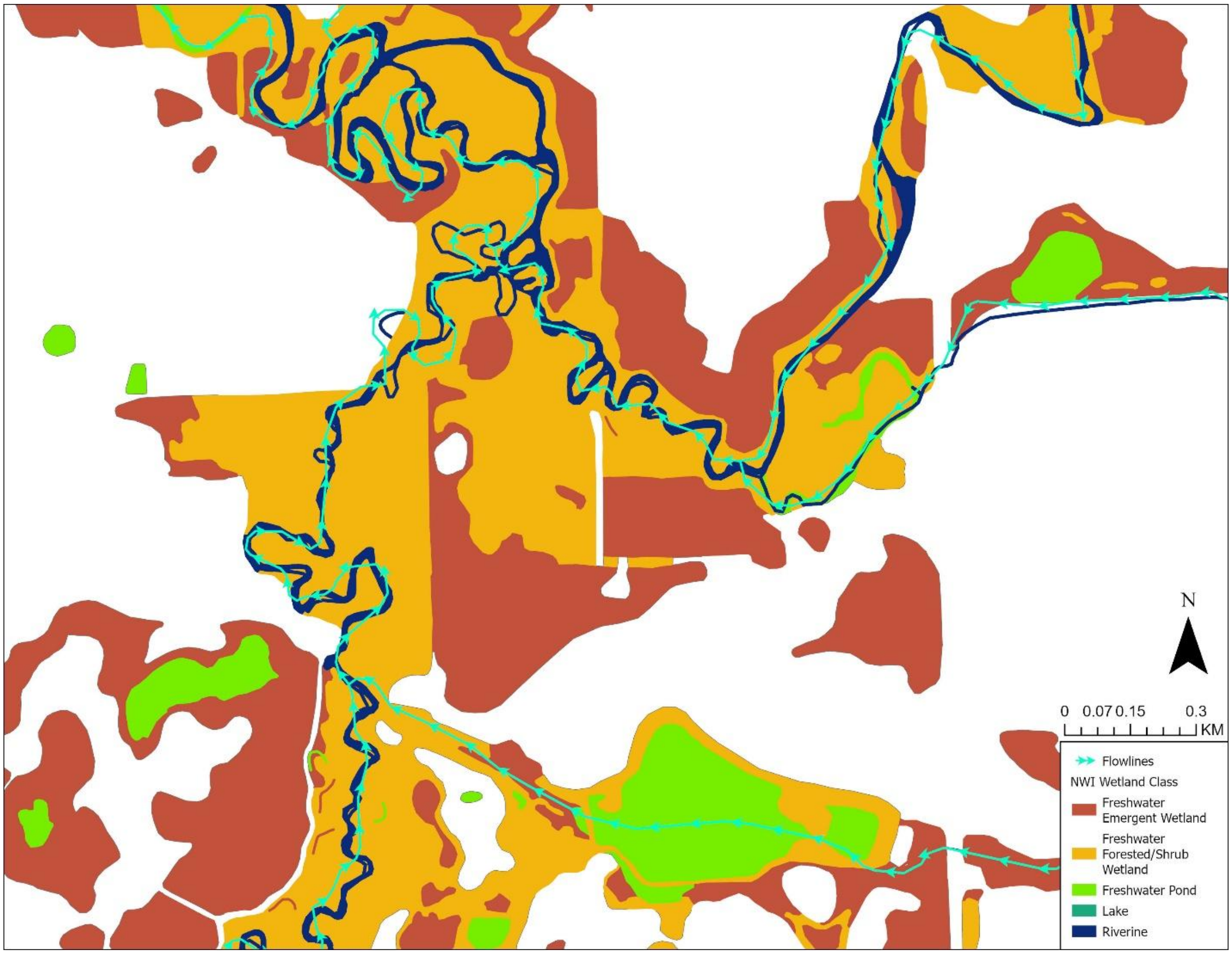


Figure 1. A map demonstrating the incompatibility of the National Hydrography Dataset Plus Version 2 and the National Wetland Inventory. The riverine wetland classification that represents the river and streams in the National Wetland Inventory (dark blue) does not line up with the flowlines that represent the river and streams in the National Hydrography Dataset (light blue).

1.4 Wetland Connectivity Metrics

Hansen et al. (2018) discussed how wetlands contribute to the removal of nitrate from river networks at a watershed scale. The authors found that wetland restoration is five times more effective than the retirement of land to pasture. They also found that ephemeral wetlands play a role, but only during high flows. However, with increasing frequency and magnitude of precipitation predicted by climate models (Marvel et al., 2023), the effect of ephemeral wetlands on nitrate removal is likely to increase (Hansen et al., 2018).

The percentage of wetland area in a watershed is a key predictive metric of in-stream nitrate concentrations (Hansen et al., 2018), however it was found to have a lot of variability (Figure 2a). Accounting for the positioning of the wetland with the upstream watershed, Hansen et al. (2018) developed a metric called interception fraction (Figure 2b). Interception fraction is defined as the fraction of the watershed drainage area intercepted by the wetlands divided by the total watershed area. Using this metric, they looked at a small subset of sites that had similar crop cover, permanent wetland cover, and drainage area. The selected sites (shown in Figure 2a) had similar crop covers ranging from 69-84% and similar amounts of non-ephemeral wetlands ranging from 2.3-3.0% (Hansen et al., 2018). For the five sites used in this analysis, they found that the interception fraction was a better metric than percentage of wetlands; however, they did not look at this metric for their full dataset. This analysis showed that the positioning of wetlands plays a key role in how much nitrate they can remove from the watershed. The purpose of this study is to extend the metric of interception fraction to all of Hansen's original data points, verifying the utility of the approach, and consider if other network-based metrics show promise in better predicting nitrate than percent wetlands.

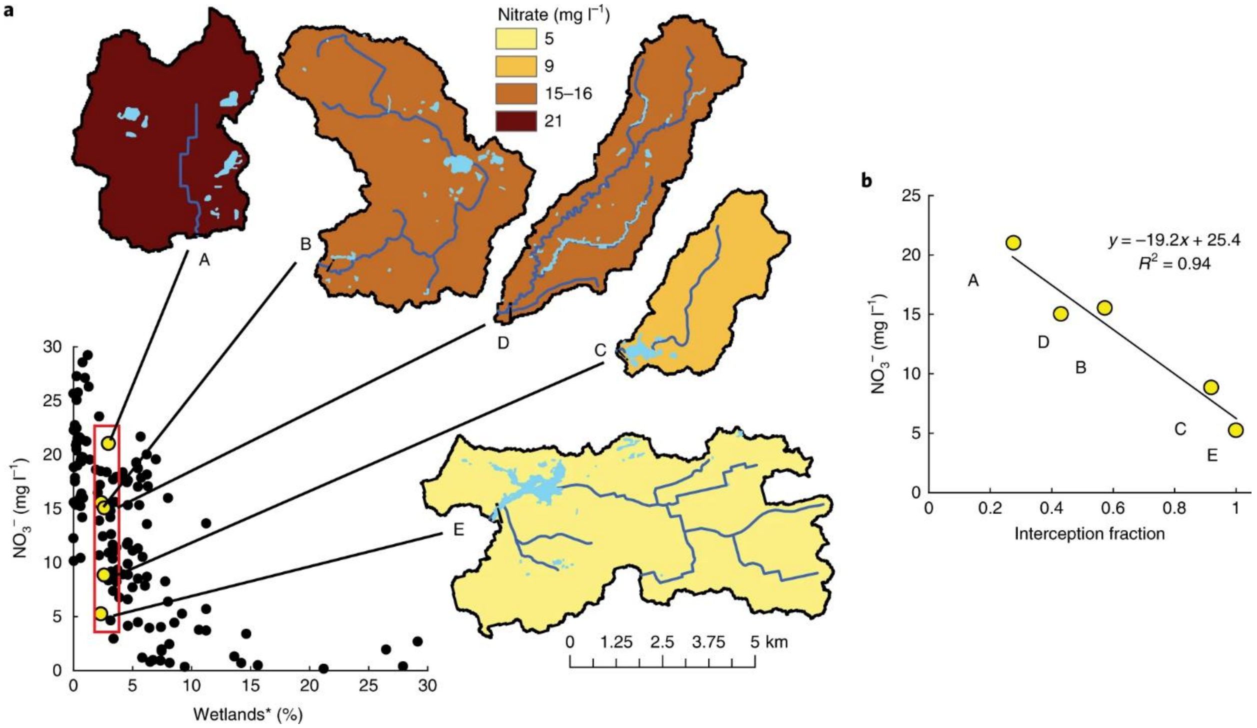


Figure 2. Effect of wetland spatial patterning on riverine nitrate (reproduced from Hansen et al. 2018, Figure 4). Panel b shows the interception fraction as a predictor of nitrate.

2. Objectives

1. Develop a method for automating the creation of river and wetland network to be used in modeling frameworks
2. Expand upon the previous analysis done to assess the relationship between interception fraction and nitrate concentrations and determine if it explains more than the percent wetland area to nitrate relationship.
3. Analyze additional network-based metrics to further understand how the spatial context of wetlands affects in-stream nitrate concentrations, focusing on the connectivity of the flow paths through a distribution of wetlands on the landscape

3. Methods

3.1 Study Area

The study area is in the Minnesota River Basin in southern Minnesota, USA (Figure 3). The wetlands of interest in this region are isolated and flow-through wetlands. In previous work by other researchers, nitrate concentration data were collected using grab samples gathered from over 200 sites during 2013 to 2016 (Dolph et al., 2017). This dataset consists of sample locations, their corresponding drainage areas, and water chemistry, stable isotopes, sediment characteristics, and stream discharge for each site. The data were collected specifically in the LeSueur, Cottonwood, and Chippewa Basins (Figure 3).

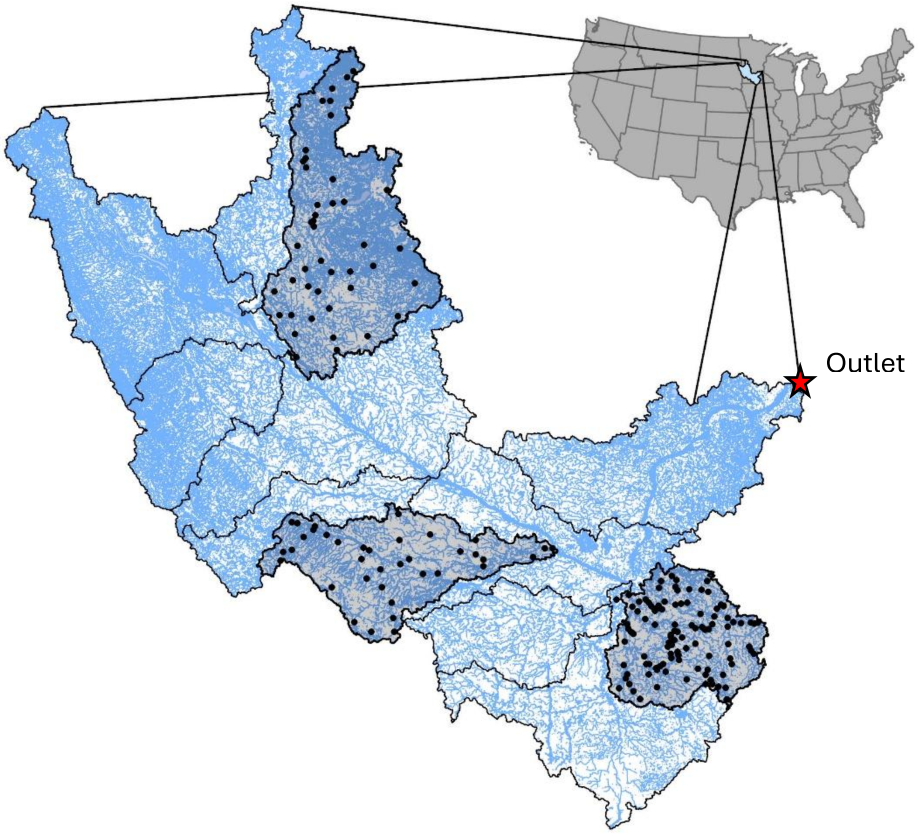


Figure 3. The Minnesota River Basin and the three highlighted watersheds where the data were collected; from top to bottom: Chippewa, Cottonwood, and Le Sueur River Basins. The black dots are the sample locations.

The landcover in the Minnesota River Basin primarily consists of cultivated crops with a 73% coverage (Figure 4). Within the three basins used in this analysis, LeSueur and Cottonwood had percent coverage of crops of 84% and Chippewa had the least crop cover at 66%. Figure 4 demonstrates the extent to which the Minnesota River Basin is cropland using landcover data from the Multi-Resolution Land Characteristics Consortium (MRLC, 2023).

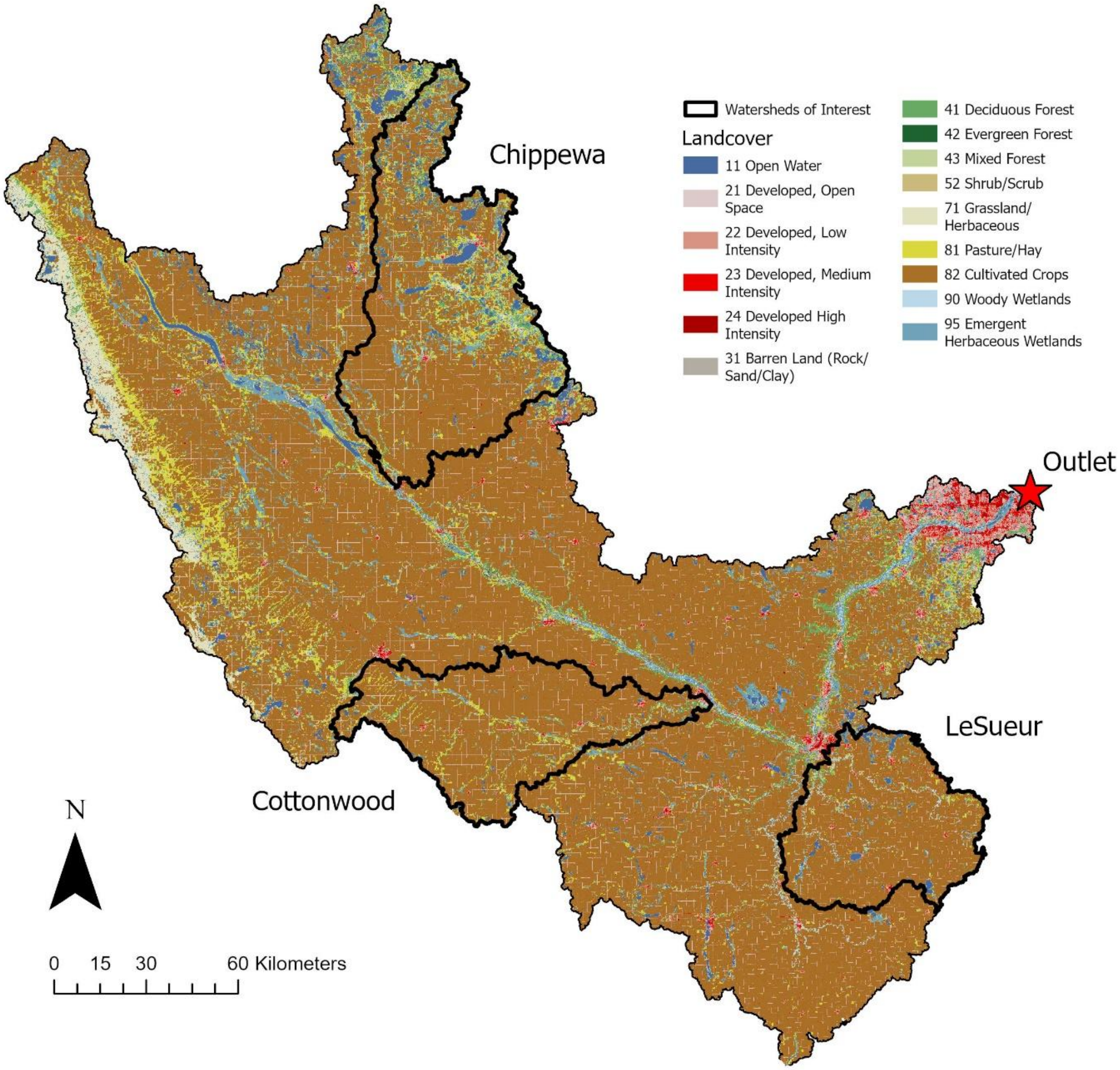


Figure 4. A map showing the landcover in the Minnesota River Basin.

3.2 Data Sources

The wetland data were obtained from the National Wetland Inventory (U.S. Fish and Wildlife Service, 2024). The stream data including the flow accumulation raster and flow direction raster were obtained from the National Hydrography Plus Version 2 dataset (EPA, 2018). The nitrate data and sample locations were obtained from data provided by Dolph et al., 2017. Landcover data was procured from the Multi-Resolution Land Characteristics Consortium with the download of the 2023 Land Cover (CONUS) dataset (MRLC, 2023). ArcGIS Pro 3.4.3 was primarily used in this study and any tools referenced further on are relevant to this version.

3.3 Data Processing

It was necessary to clean the data before use, since the nitrate data provided by Dolph et al., 2017 contained more information than what was used in Hansen et al., 2018. Sites that were not used in Czuba et al. (2018) in the Le Sueur basin were removed. Any sites missing nitrate (NO3) data were removed. There were seven sites removed because they were not located in the Minnesota River Basin. If sites contained “bdl” in their nitrate data, it was replaced with the minimum detection limit of 0.010 mg/L NO3. Nitrate data from the dataset were used from samples collected in June 2015 for LeSueur and June 2014 for Cottonwood and Chippewa. This nitrate data represents baseflow conditions during June for the region. However, the 2014 samples were collected after a large storm event. Therefore, for consistency, all the sample points collected in 2014 were selected and another analysis was run with them using the same methods. It should be noted that the LeSueur basin has the most samples, numbering 70 in June 2015, Chippewa has the second highest number of samples at 38 in June 2014, and Cottonwood has the least number of samples at 30 in June 2014 once these edits were made. However, when the 2014 June data is used for LeSueur in the June 2014 only analysis, it only has 36 data points.

3.4 Creating Watersheds

When the watersheds for each sampling site were created it was found that the site locations, as indicated by a point shapefile in ArcGIS, were not always on the point of highest flow accumulation which indicated a major flow path. These sites were moved to the nearest main flow path flow accumulation point by hand. Some sites were in wetlands where the largest flow accumulation value was not clear. These sites were not moved and not used in the analysis.

To create watersheds for each of the sites, the watershed tool in ArcGIS was used; however, if the sites were all incorporated in a single use of the tool, it would create incremental drainage areas since the site’s watersheds overlap. To get around this a script was created that would loop through each point in the shapefile, creating a pour point using the snap pour point tool and then just creating one watershed at a time. After all the watersheds were created, they were merged into one shapefile for easier use.

3.5 Creating an Integrated Wetland-Riverine Network

An integrated wetland-riverine network is a network of wetlands and flowlines that has attributes assigned to both wetlands and flowlines to allow for the seamless inclusion of wetland features into a network. Once it has been created, it can be used to calculate the percent wetland area of the watershed, interception fraction, and it can be used within more complex models to determine nitrate exports in watersheds.

Only wetlands that were not in the floodplain and that are not classified as riverine by the NWI are used in the calculation of interception fraction (Figure 5). The goal is to keep only the isolated and flow-through wetlands within the Minnesota River Basin. The riverine wetlands were simply removed by using Select by Attributes and exporting all the features that were not riverine. Determining what a floodplain wetland was more complex. To do this the National Hydrography Dataset Plus Version 2 flowlines were buffered 120 meters and this buffer was used to clip the NWI wetlands. The 120 meter buffer was picked because it was divisible by the DEM resolution and visually encompassed most wetlands that were adjacent to the flowlines. A ratio was then created for the wetlands that compared how much of each wetland's area was in the 120-meter buffer to the total wetland area. If the wetland was a Lake or Freshwater Pond or adjacent to one, it was kept. Next, after dividing the total wetland area by the area in the buffer if the ratio was less than or equal to 5, they were kept. The cut off value of 5 was selected since it visually removed the majority of the floodplain wetlands. When looking at a histogram for this value, even with a logarithmic transformation, there were no distinct peaks.

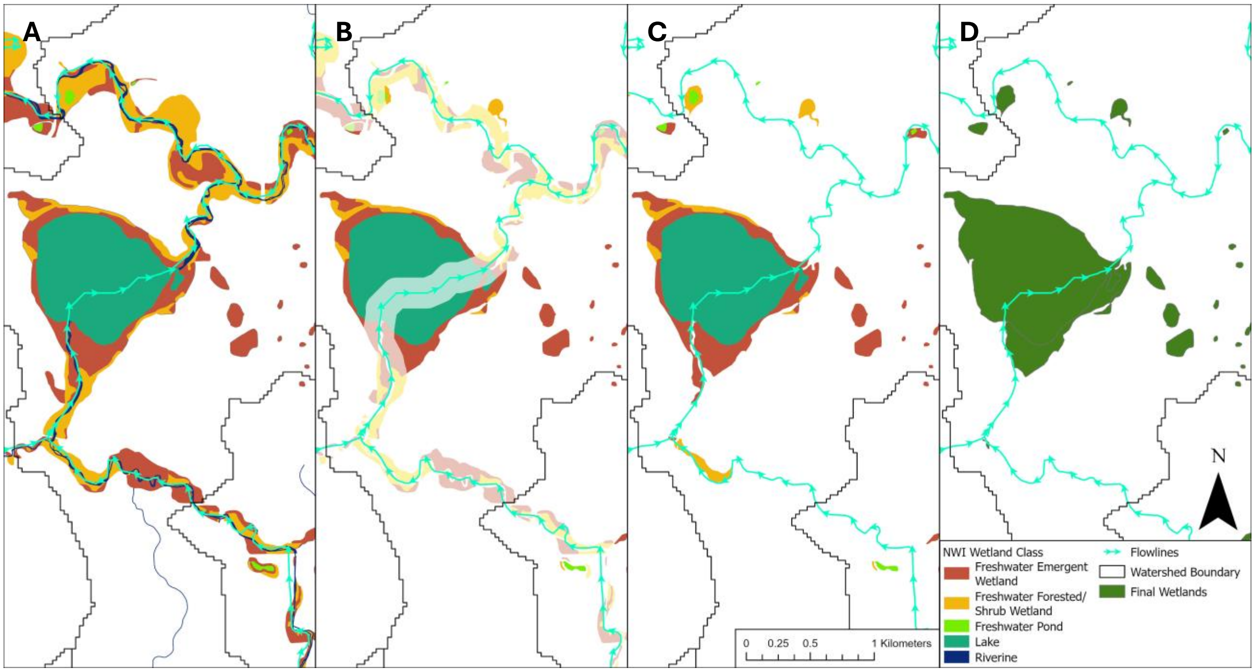


Figure 5. Maps showing how the wetlands were processed to remove floodplain wetlands. A is the original NWI network overlayed with flowlines and site watershed boundaries. B represents the 120 buffer that was used to create a ratio of wetland area within the buffer to outside of the buffer to ensure that large wetland complexes as seen above are not removed from the network. C shows the results with no floodplain wetlands. D shows the final result after a wetland drainage area to wetland area ratio was created and used to remove the last remaining floodplain wetlands. All maps in this figure are at the same scale and orientation.

Small wetlands that intercepted a majority of the flow were removed. This was done by examining the wetland area to drainage area ratio in a histogram for each basin. Figure 6 shows LeSueur's ratio of wetland area to wetland drainage area which was the only one with a second peak. Wetlands that were considered small if their wetland area to drainage area ratio was less than or equal to 0.0001. This value is the inverse natural log of the value shown in Figure 6. This helps to ensure that any remaining floodplain wetlands get removed. Additional wetlands were removed if they had the special modifier "A" using the Erase tool before calculating interception fraction. The "A" modifier is part of the Cowardin wetland classification and can be found in National Wetland Inventory wetland polygon attribute table. It is defined as temporarily flooded were surface water is only present for short periods with the water table located well below ground surface for most of the growing season (U.S. Fish and Wildlife Service, 2019).

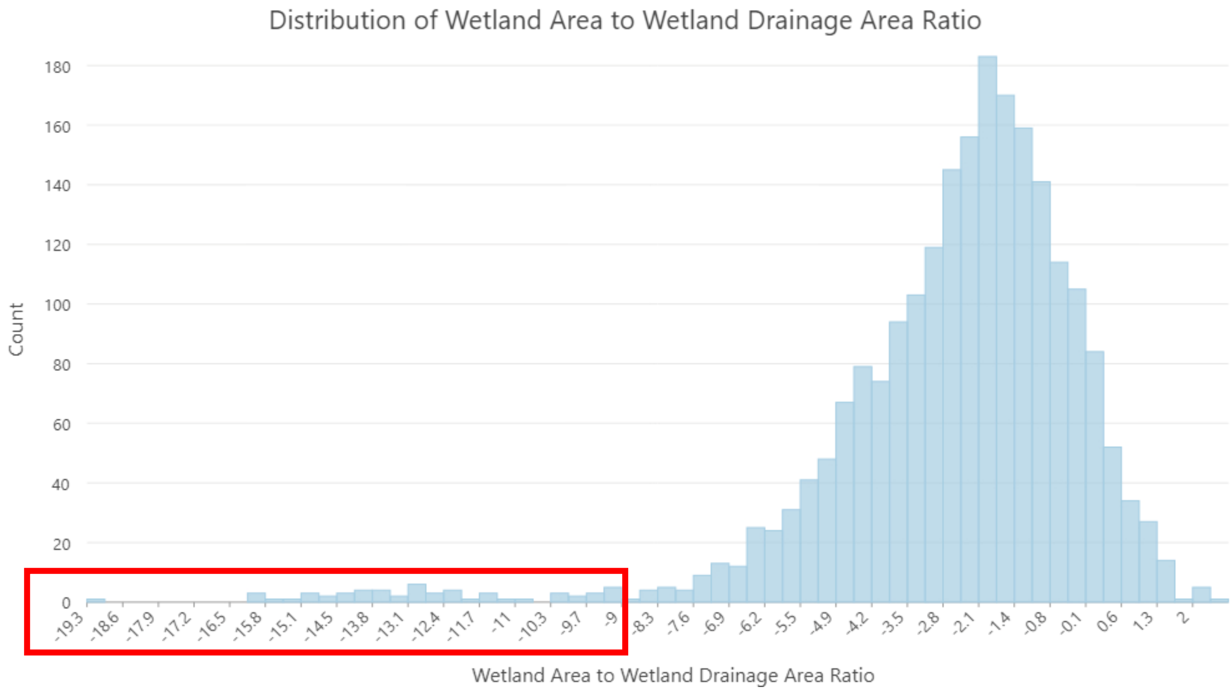


Figure 6. A histogram showing the distribution of wetland area to drainage area ratio in the LeSueur basin.

The flowlines provided by the National Hydrography Dataset were not used. Instead, new flowlines were made from the flow accumulation raster from the National Hydrography Dataset using Raster Calculator, keeping all values greater than or equal to 500. This value was chosen because it created a network most similar in extent to the NHD network without too many extra lines. Lines that extended beyond the National Hydrography Dataset were removed and slope was added to each line based on the nearest National Hydrography Dataset flowline. The Digital Elevation Model that came with the National Hydrography Dataset had watershed boundaries and flow paths burned into it making the extraction of channel slope from it not possible. The flowlines within wetlands were removed and the lines were then split to ensure that they were no longer than 5 kilometers. To route the network, points were created at the start and ends of each link and flow accumulation values were assigned to them. The full details on processing the wetland and river network dataset within the ArcGIS environment are described in detail in Appendix A.

3.6 Spatial Analysis

Percent wetland is the area of the watershed that is wetland divided by the watershed area. This was calculated by summing the areas of the wetlands within each site's watershed and dividing by the total watershed area. The wetlands used in the percent wetland calculation are the isolated and flow-through wetlands from the integrated wetland-riverine network.

Interception fraction is the area of the watershed intercepted by wetlands divided by the total watershed area (Figure 7). The wetlands used in the interception fraction calculation are the isolated and flow-through wetlands from the integrated wetland-riverine network. Within this section all of the italicized text are tools found and used in ArcGIS Pro 3.4.3. The first step to determining the area of the watershed intercepted by the wetlands is to find the point of highest flow accumulation within each wetland. The flow accumulation value was converted to points and *Summarize Within* was used on wetlands that contain flow accumulation points to determine where the outfall was by finding the maximum point. *Summarize Nearby* was used on the wetlands that had no values after *Summarize Within* was used. *Summarize Nearby* used with a 30 meter buffer to include small wetlands that did not intersect the flow accumulation points. Since the flow accumulation raster was 30×30 meters the points that were created had 30 meter gaps in between them so it was possible for a wetland to not be on one of the points. Once the outfall points for the wetlands were identified, any that had a flow accumulation value of zero were removed, since no flow goes to that cell. If a wetland had duplicate maximum flow accumulation values, the *Delete Identical* tool was used to ensure that it only had one outlet. These wetlands occurred mostly with low flow accumulation values. The *Snap Pour Point* tool and the *Watershed* tool were used to delineate the wetland watersheds. The area for each wetland was calculated and then rejoined to the wetland outlet points. The *Summarize Within* tool was used for each site watershed on those points to determine the total wetland watershed drainage area for each site. The wetland watershed area was then divided by the watershed area to determine the interception fraction.

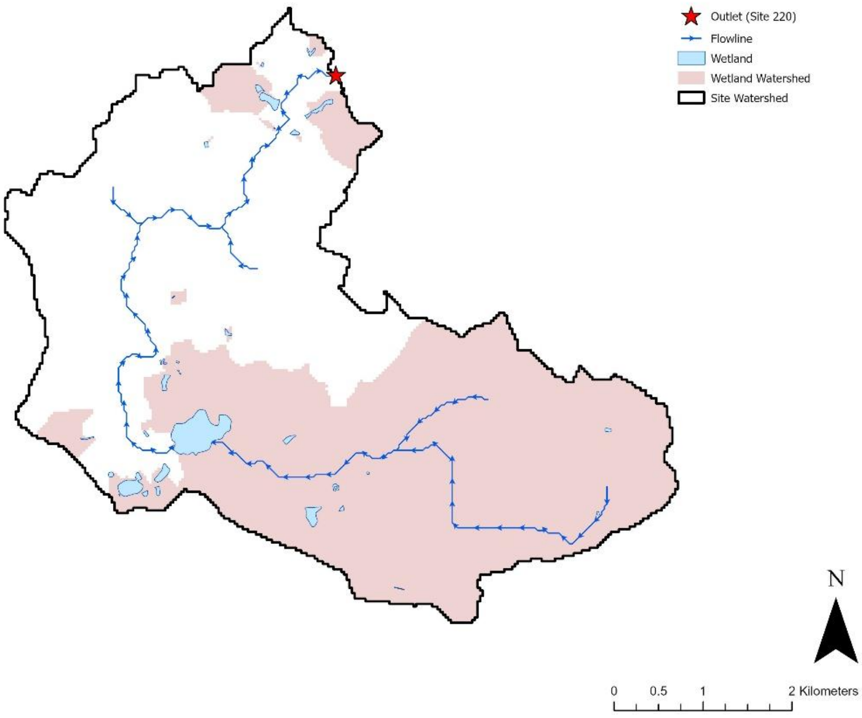


Figure 7. A visual example of what interception fraction looks like. The light red is the area of the watershed that is intercepted by wetlands. The interception fraction for this watershed is 53%. It had a median nitrate concentration of 15.7 mg/L at the outlet for June 2015.

An additional metric was calculated to account for wetland size and location. This metric aims to assess if the size of the wetland at the furthest downstream point within the watershed has an effect on interception fraction. A small wetland would remove less nitrate than a large wetland with a longer residence time. A watershed with a small wetland at the outlet may have a higher nitrate export than one with a large wetland at the outlet, but they would both have the same interception fraction value since all of the watershed is intercepted. This was done by finding the wetland with the largest drainage area for each site's watershed by spatially joining the drainage area value previously found when determining the wetland's drainage area to the wetlands. Then *Summarize Within* was used to locate the wetland with the largest drainage area, the one that intercepts the most flow. This value was then converted to square kilometers by multiplying by the raster grid size of 30×30 meters and the conversion factor for square meters to square kilometers. Then the wetland area was divided by the drainage area of the wetland to get a ratio of wetland area to drainage area.

Another metric was calculated to account for crop coverage in the non-intercepted areas of each sample's watershed. This was accomplished by finding the area that is not intercepted using the *Symmetrical Difference* tool and running *Zonal Statistics as Table* to determine the number of cultivated crop pixels in each sample's watershed. The pixel size of the landcover data is 30×30 meters, so the pixel count was converted to square kilometers using that value and the conversion factor for square meters to square kilometers. Then the non-intercepted crop area was divided by the watershed area to determine the percent of non-intercepted crop land for each sample.

An exponential line of best fit was applied to the percent wetland area versus nitrate relationship. This line of best fit was selected because it had the best r-squared value when compared to other best-fit lines. A linear best-fit line was applied to the comparison of interception fraction versus nitrate because it had the best r-squared value when compared with other best-fit lines. Residuals were then calculated for the interception fraction versus nitrate relationship by calculating a predicted nitrate value for each site's interception fraction value. The difference between the predicted and the actual sample value for the nitrate is the residual. Next to determine if the metrics described above, wetland area to maximum drainage area and non-intercepted crop cover, had any impact on the relationship between interception fraction and nitrate, the values for each site were plotted against the site's residuals in separate graphs for each metric.

This analysis was repeated for two different subsets of the data from the month of June. The first set of data mimicked the data used to create Figure 2b. In Figure 2b, the data used for LeSueur was from 2015 and the data used for Cottonwood was from 2014. Chippewa was not used in Figure 2b, but since data was not collected in 2015 for Chippewa, the 2014 data for it was used. This was done to stay consistent with the original work done in Hansen et al. (2018). The second set of data was only from 2014 for all three watersheds. This choice was made because it was noted that there was a storm event in 2014 before sampling (Hansen et al., 2018) which may have affected the sampled nitrate concentrations.

Interception fraction was calculated for every link in the integrated wetland-riverine network in the Minnesota River Basin. These values were then used to create a predictive map of nitrate concentration within the Minnesota River Basin using the best fit lines created for each dataset that represent the relationship between interception fraction and nitrate. Then these values were averaged within the LeSueur, Cottonwood, and Chippewa and the root mean square error was calculated to explore how these values vary across the Minnesota River Basin.

4. Results

4.1 LeSueur 2015, Cottonwood and Chippewa 2014

The following results include the LeSueur sample data from 2015 and Cottonwood and Chippewa from 2014 to be consistent with the values show in Hansen et al (2018). Figure 8 shows the percentage of wetland area within each site's watershed area compared to nitrate concentration. There is a decrease in nitrate as the percentage of wetlands in the watershed increases. These values are consistent with the findings of Hansen et al. (2018) as seen in Figure 2. This suggests that the wetlands used in both studies are consistent. Additionally, once a watershed has more than 5% wetlands, most of the nitrate is less than 5 mg/L. LeSueur tends to have higher concentrations compared to the Cottonwood and Chippewa basins and the r-squared value for the best fit curve in Figure 8 is 0.47.

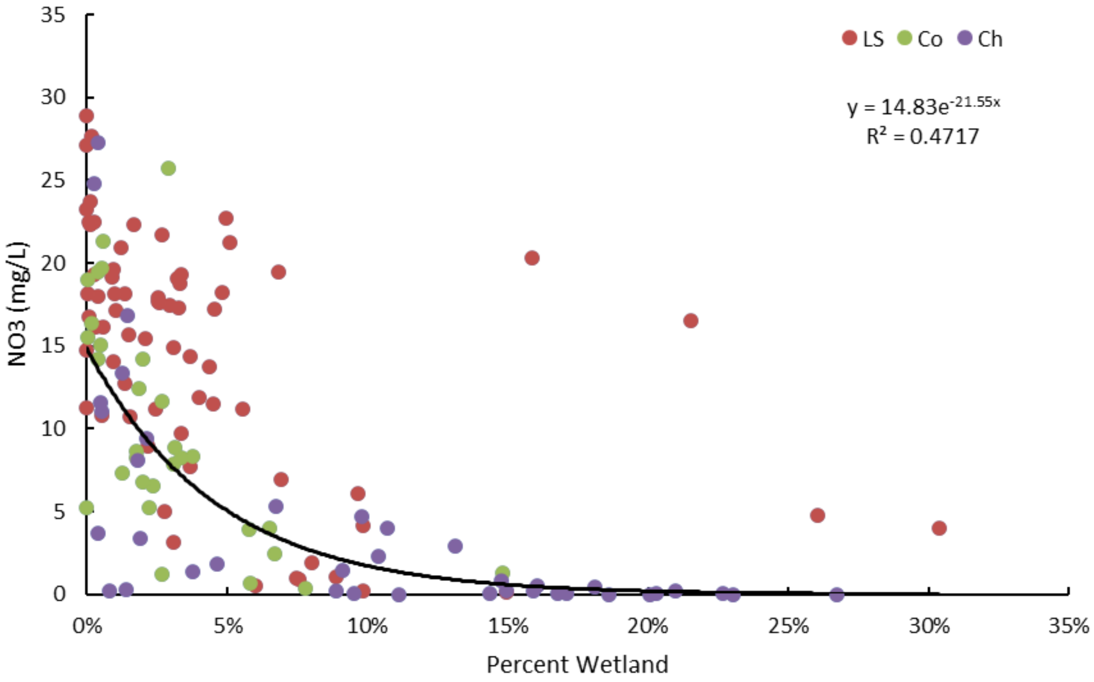


Figure 8. Percent wetland in comparison to NO3 for all of the sites in LeSueur from June 2015, Cottonwood from June 2014, and Chippewa from June 2014.

Figure 9 is a reproduction of Figure 2b from Hansen et al. 2018 using the methods to calculate interception fraction as described above. The same sites were used in Figure 9 and Figure 2b. Since some of the sites in Figure 9 had multiple NO3 samples taken during the time frame used, the median NO3 value was used to obtain a single value. When compared to Figure 2b, the main difference between the two is the interception fraction values, with Site C changing the most shifting from about 90% interception in Figure 2b to almost 100% interception in Figure 9. Figure 9 does show that interception fraction is a good indicator of nitrate and confirms that the methods to create the network were consistent.

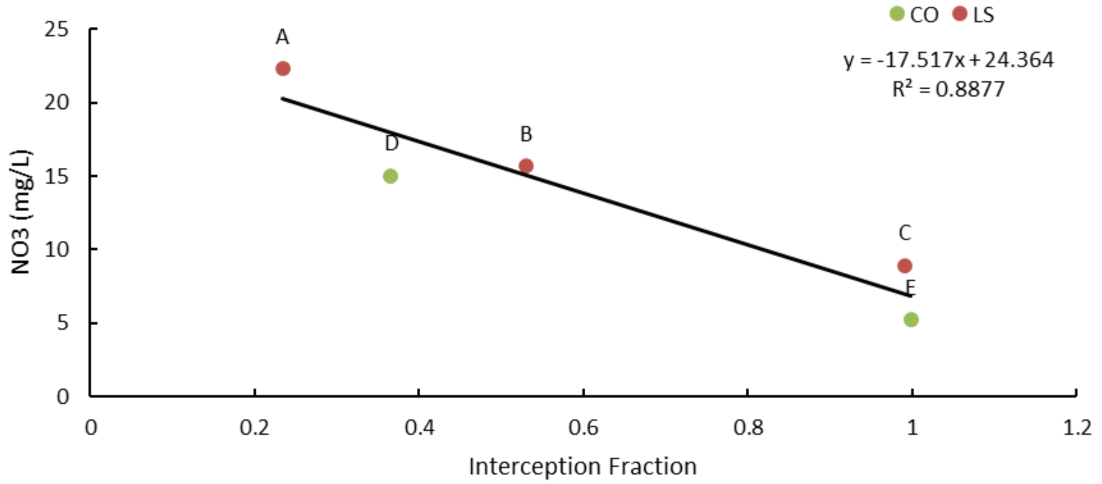


Figure 9. Recreation of Figure 2b from Hansen et al. 2018 using the methods described above.

Interception fraction was calculated for each site. Figure 10 shows a basin wide visual of how much of each watershed is intercepted by wetlands. Chippewa has the most area intercepted by wetlands with 69% of the total watershed area being intercepted, followed by LeSueur with 47% intercepted, and then Cottonwood with 27% intercepted. Figure 10 helps to visualize how much land these wetlands actually receive water from. All maps in this Figure 10 are at the same scale and orientation.

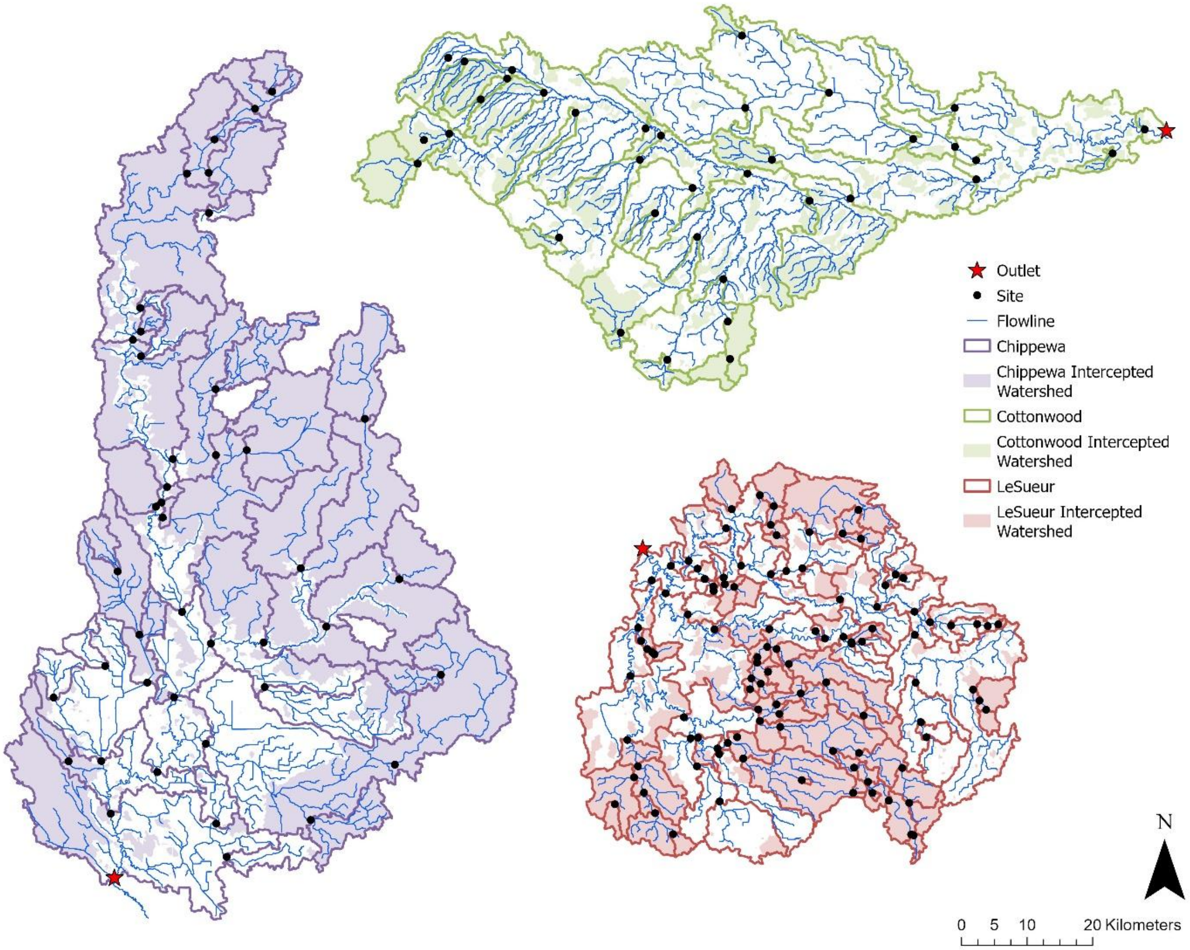


Figure 10. A map of the three watersheds that shows total interception fraction for each.

Figure 11 shows the relationship between interception fraction and NO3 and includes all the data from June 2015 in LeSueur and June 2014 in Cottonwood and Chippewa. It should be noted that the LeSueur basin has the most samples, numbering 70, Chippewa has the second highest number of samples at 38, and Cottonwood has the least number of samples at 30. The majority of Chippewa's samples are concentrated around an interception fraction of 1 with low nitrate values, while Cottonwood and LeSueur have a greater variety of results. There is a clear relationship between interception fraction and nitrate. These results are reflected in the amount of watershed area that is intercepted by wetlands in each watershed as seen in Figure 10.

Figure 11 also shows that interception fraction is a better metric than percent wetland area. The difference is apparent through the r-squared values. The r-squared value in Figure 11 is 0.55 while the value in Figure 8 is 0.47 indicating that interception fraction is an improvement over percent wetland. However, these results in Figure 11 with more data points were not as good as the results in Figure 9.

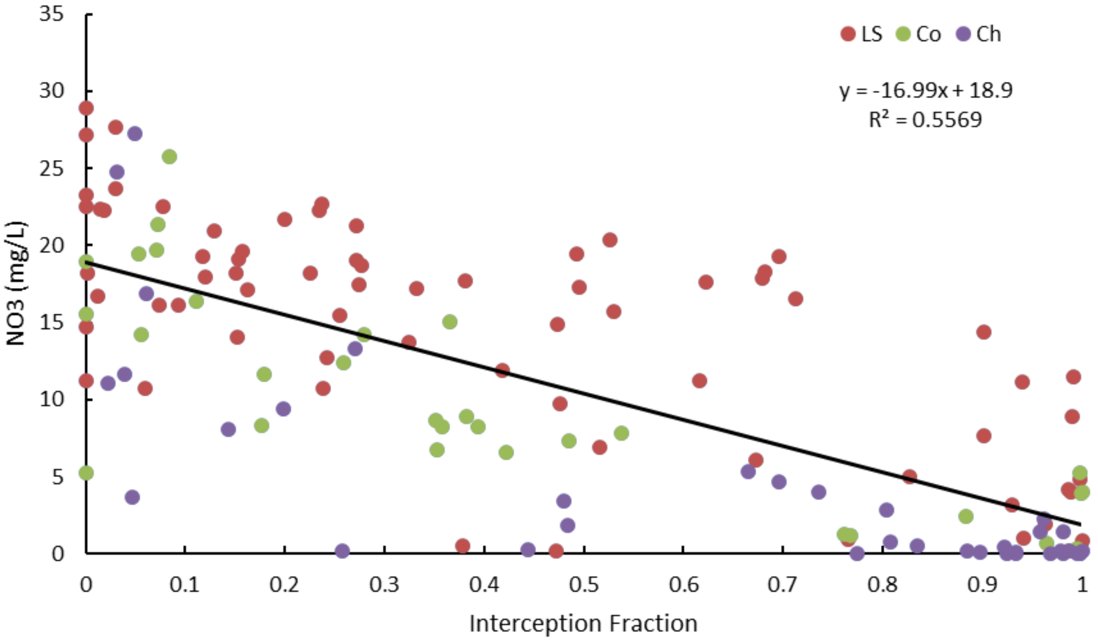


Figure 11. Interception fraction compared to NO3 for all of the sites in LeSueur from June 2015, Cottonwood from June 2014, and Chippewa from June 2014.

Figure 12 shows the residuals from the line of best fit in Figure 11. This indicates no obvious trends. The Cottonwood and Chippewa values tended to be underestimated while LeSueur tended to be overestimated. Cottonwood had the least amount of variation.

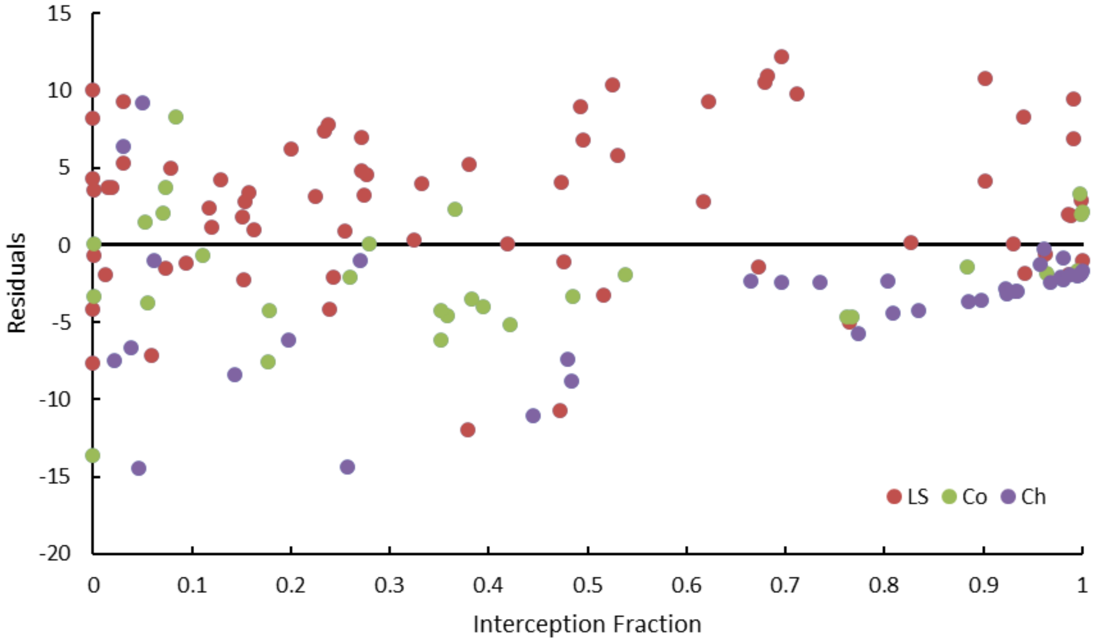


Figure 12. Residuals from the line of best fit in Figure 11.

To determine if the wetland size in relation to its position would have an impact on interception fraction the wetland with the maximum drainage area for each site was selected. Then a ratio of wetland area to maximum drainage area was calculated and plotted against the residuals found in Figure 13. Figure 13 shows that there is no effect of the wetland size in relation to its location on the network to interception fraction.

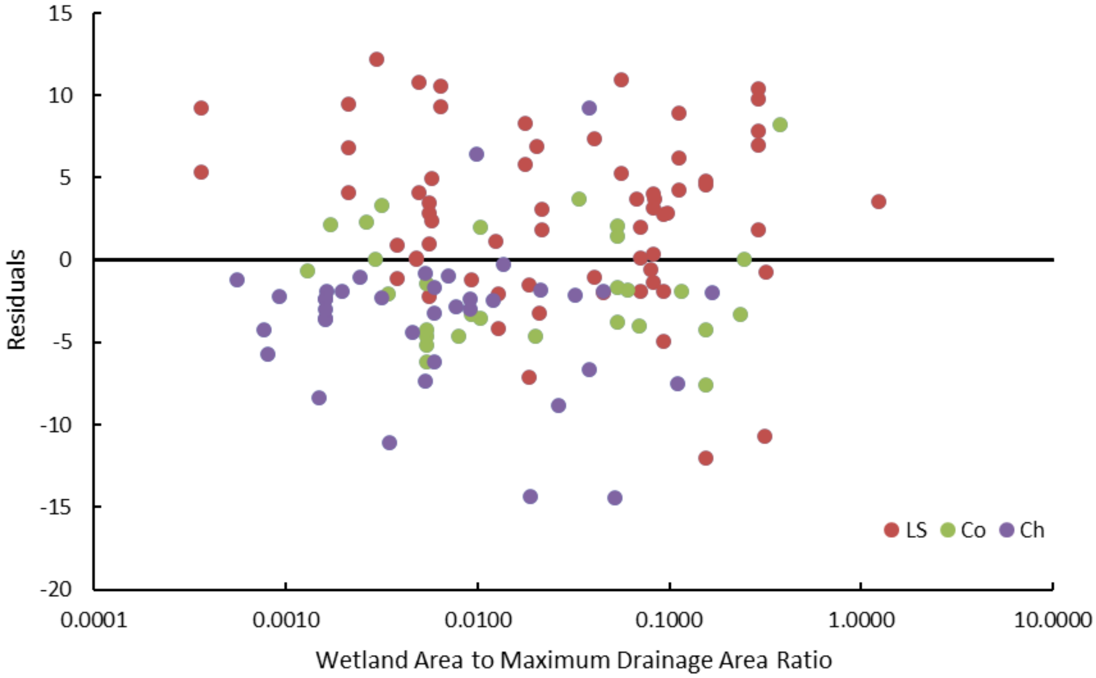


Figure 13. Residuals (from Figure 12) compared to the wetland drainage area divided by the wetland surface area for the wetland with the largest drainage area in each watershed.

Figure 14 shows the percentage of non-intercepted area that is crops compared to the residuals from Figure 12. This was done to see if crop coverage would have an impact on interception fraction. Figure 14 shows that there is no relationship between crop coverage in the non-intercepted areas for each site and interception fraction.

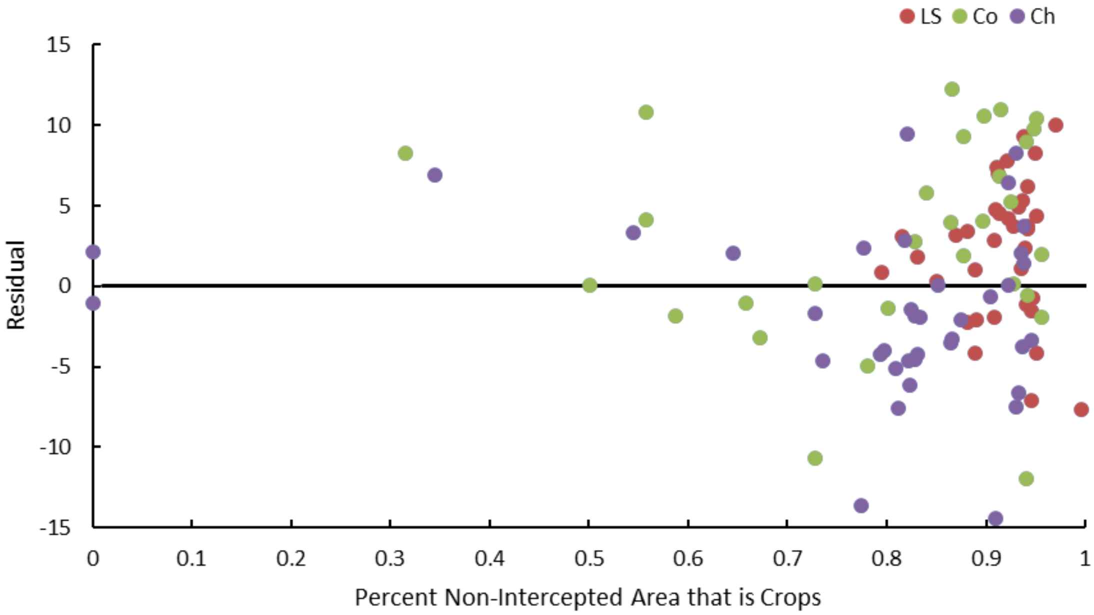


Figure 14. Residuals (from Figure 12) compared the percent of non-intercepted land for each site that is crop land.

Figure 15 shows the interception fractions for each link in the Minnesota River Basin. Brown means the link has a low interception fraction and blue means it has a high interception fraction. The average interception fractions for LeSueur is 47%, Cottonwood is 27%, and Chippewa is 69%. Overall the entire Minnesota River Basin has an average interception fraction value of 55%.

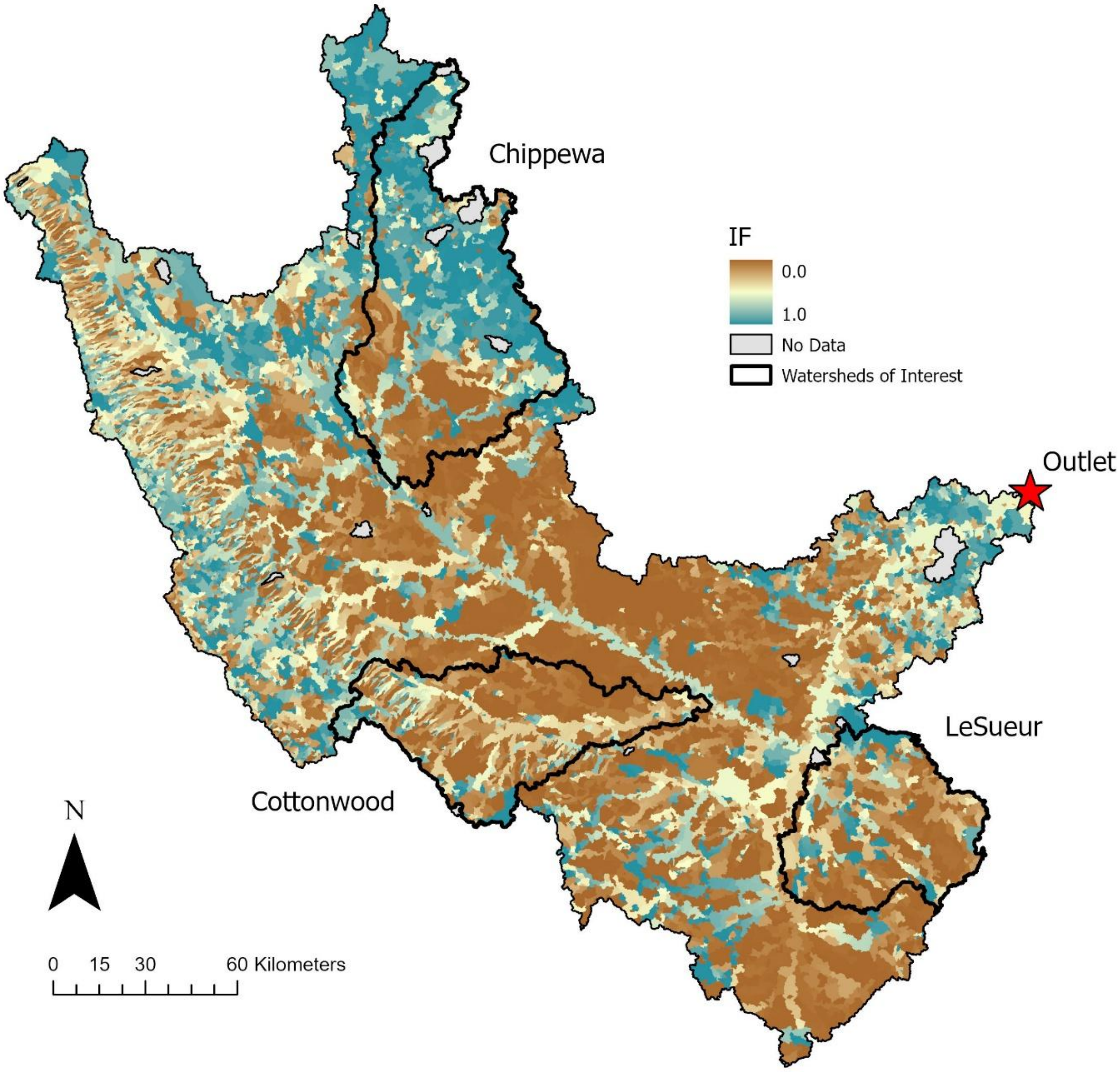


Figure 15. A map showing the interception fractions for all of the links in the Minnesota River Basin.

Figure 16 shows predicted nitrate values within the Minnesota River Basin. The estimated nitrate values that are shown were calculated using the linear regression equation in Figure 11. Areas that are dark green have high predicted nitrate values. Areas that are pale green have low predicted nitrate values. The major river sections are predicted to have less nitrate input from the landscape while areas on smaller tributaries may have much more. Each flow path link and wetland have a watershed therefore if the link was a wetland the interception fraction would have been 1 which would provide a low nitrate value of 1.91 mg/L. The upper range of nitrate is

18.9 mg/L which occurs when there are no wetlands within a links watershed. These nitrate concentration values occur at 100% interception and 0% interception respectively.

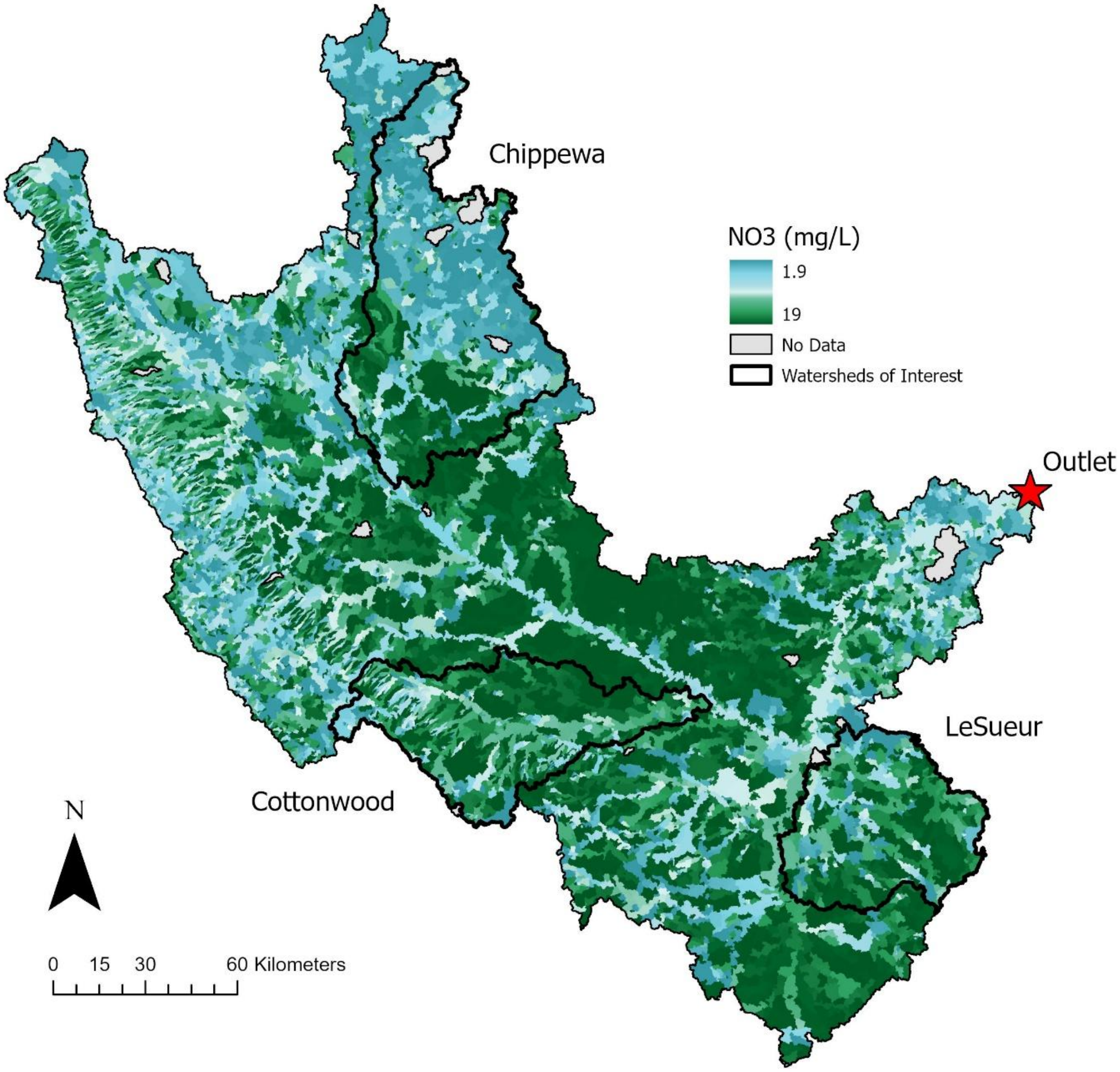


Figure 16. A map showing the nitrate concentration values as estimated using the linear equation in Figure 11 in the Minnesota River Basin.

4.2 LeSueur, Cottonwood, and Chippewa 2014

The following results are for samples collected in all three watersheds in June 2014. The 2015 results for LeSueur in section 4.2 were analyzed to be consistent with the data used in Figure 2. Since the samples in LeSueur taken in 2015 tended to be higher than the 2014 samples in Chippewa and Cottonwood the 2014 samples for LeSueur were also analyzed to see if the different conditions that the samples were collected in 2015 and 2014 influenced the relationships.

Figure 17 shows the percentage of wetlands in the LeSueur, Cottonwood, and Chippewa watersheds compared to nitrate concentration. As the percentage of wetlands increase the nitrate concentration tends to decrease. This relationship is better than the one in its companion plot in Figure 8 as seen through the r-squared values. The r-squared value for this curve is 0.60 which is better than the r-squared value in Figure 8 of 0.47. It still looks like LeSueur has higher nitrate concentrations than the Cottonwood and Chippewa, but there is less of a difference than what is seen in Figure 8.

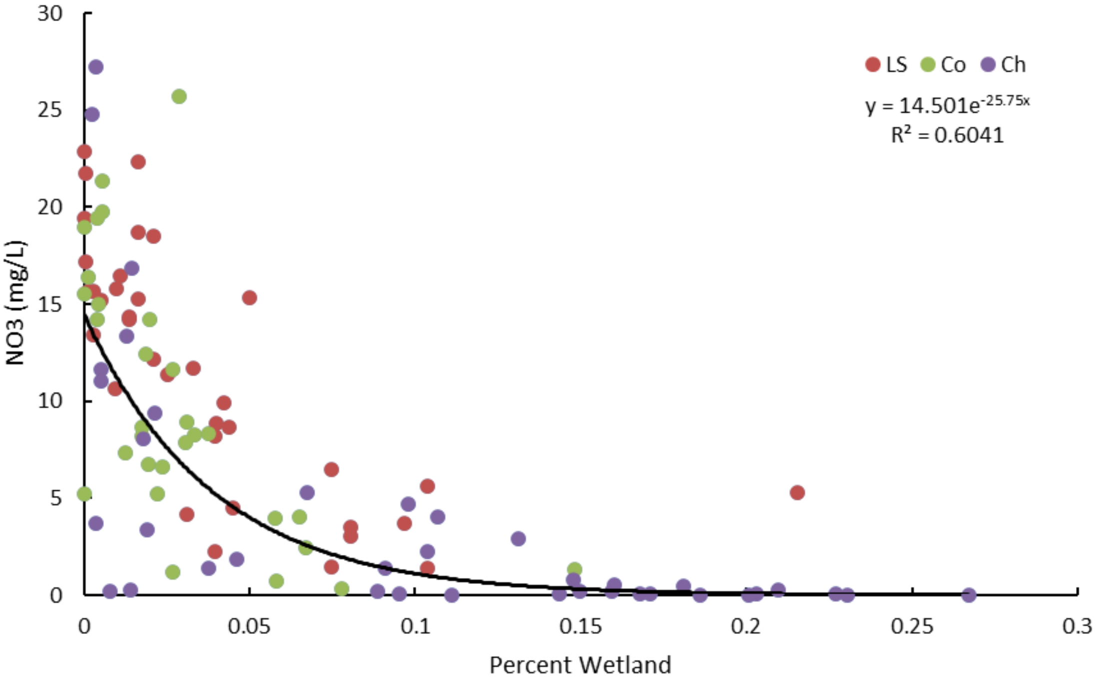


Figure 17. Percent wetland compared to nitrate for the LeSueur, Cottonwood, and Chippewa basins using only the samples from 2014.

Using just the samples from June 2014 for nitrate concentration in all three watersheds, Figure 18 shows the relationship between NO3 and Interception Fraction. It has a better relationship than the one shown in Figure 11, clearly indicated by the r-squared values. Figure 18 has an r-squared value of 0.66 which is greater than Figure 11's r-squared value of 0.55. The interception fraction to nitrate concentration relationship in Figure 18 is still better than the percent wetland to watershed ratio shown in Figure 17. There are no watersheds that particularly stand out in Figure 18. However, these results with more data points are still not as good as the results in Figure 9, but they are closer than previously shown in Figure 11.

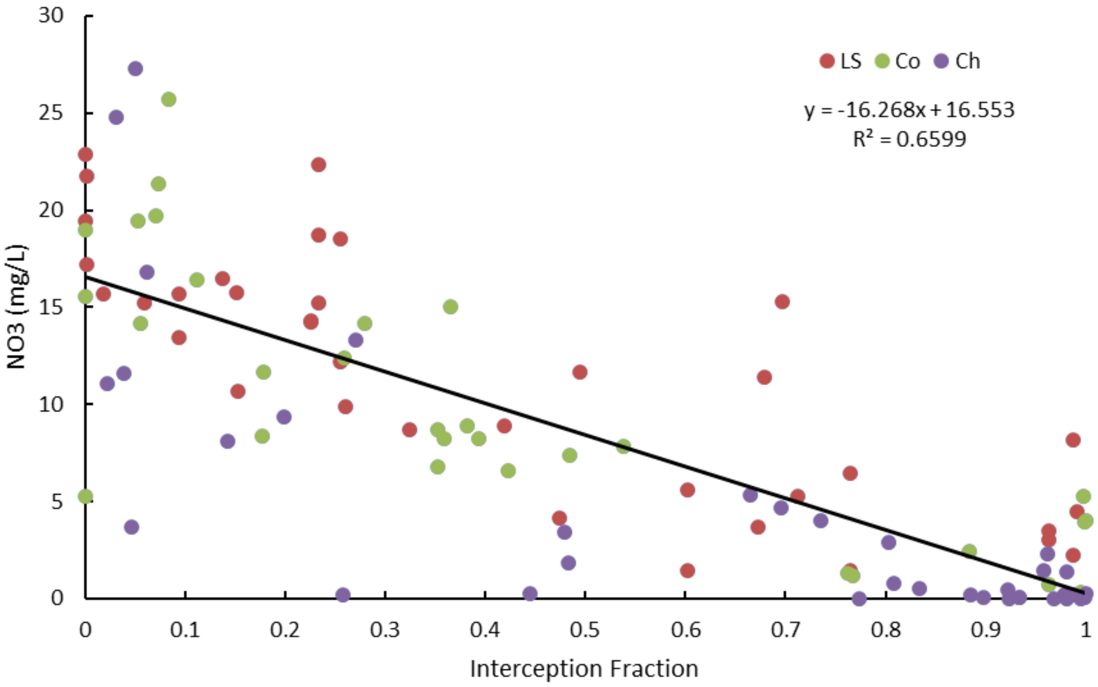


Figure 18. Interception fraction compared to NO3 for all of the sites in LeSueur, Cottonwood, and Chippewa from samples taken in June 2014.

Figure 19 shows the residuals from the line of best fit in Figure 18. It indicates no obvious trends. Chippewa tends to be underestimated, while the fit for Cottonwood and LeSueur has more overestimation. This varies from Figure 12 in which Cottonwood was overestimated.

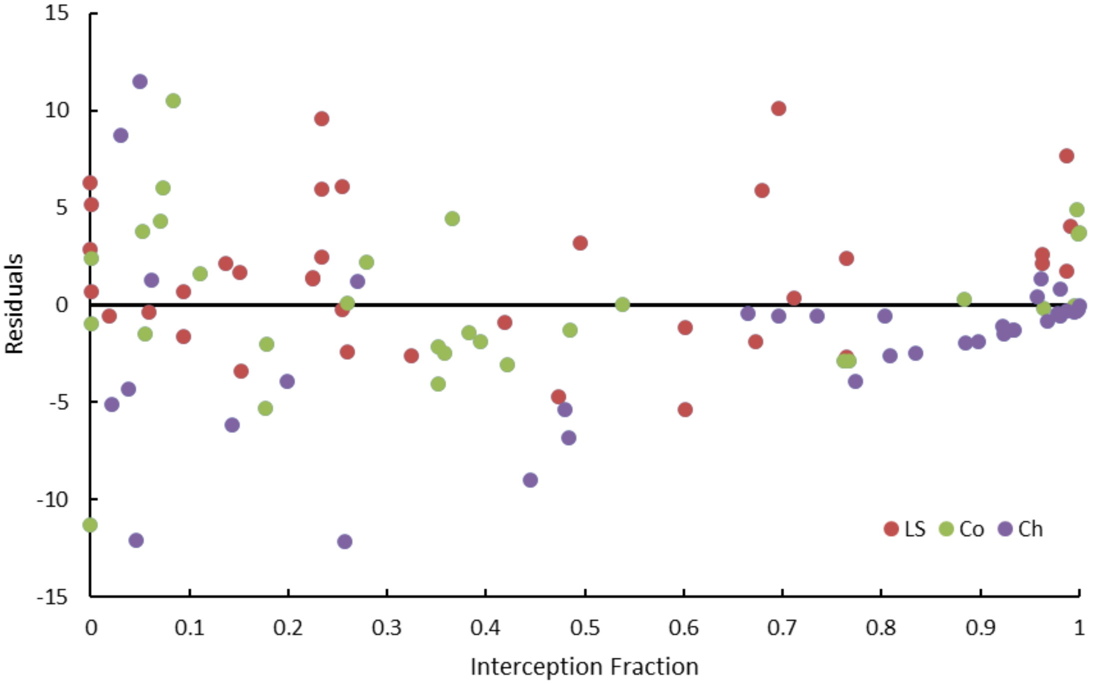


Figure 19. Residuals from the line of best fit in Figure 18.

Figure 20 shows no relationship between the wetland area to maximum drainage area ratio and the residuals from Figure 19. The only difference it has from Figure 13 is the LeSueur data points.

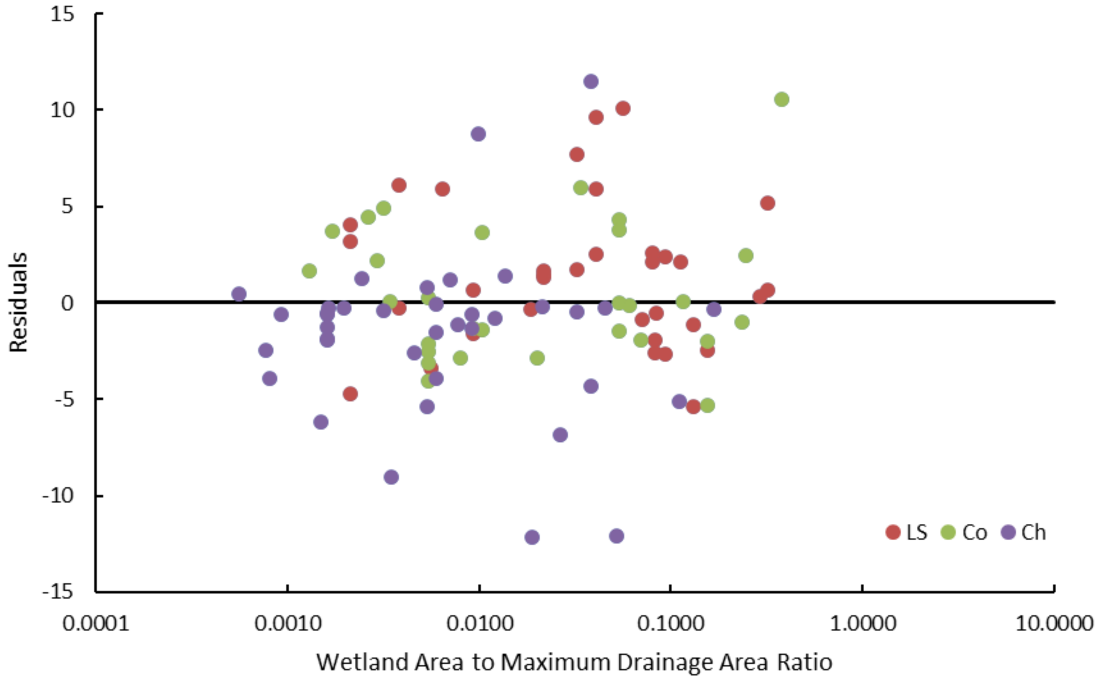


Figure 20. Residuals (from Figure 19) compared to the wetland drainage area divided by the wetland surface area for the wetland with the largest drainage area in each watershed.

To determine if crop coverage would have an impact on interception fraction, the crop cover in the non-intercepted areas of each site was plotted against the residuals from Figure 19 (Figure 21). Figure 21 shows that there is no relationship between crop coverage in the non-intercepted areas for each site and interception fraction. LeSueur and Cottonwood are very similar in their point locations in Figure 21, with Chippewa showing a little more variation.

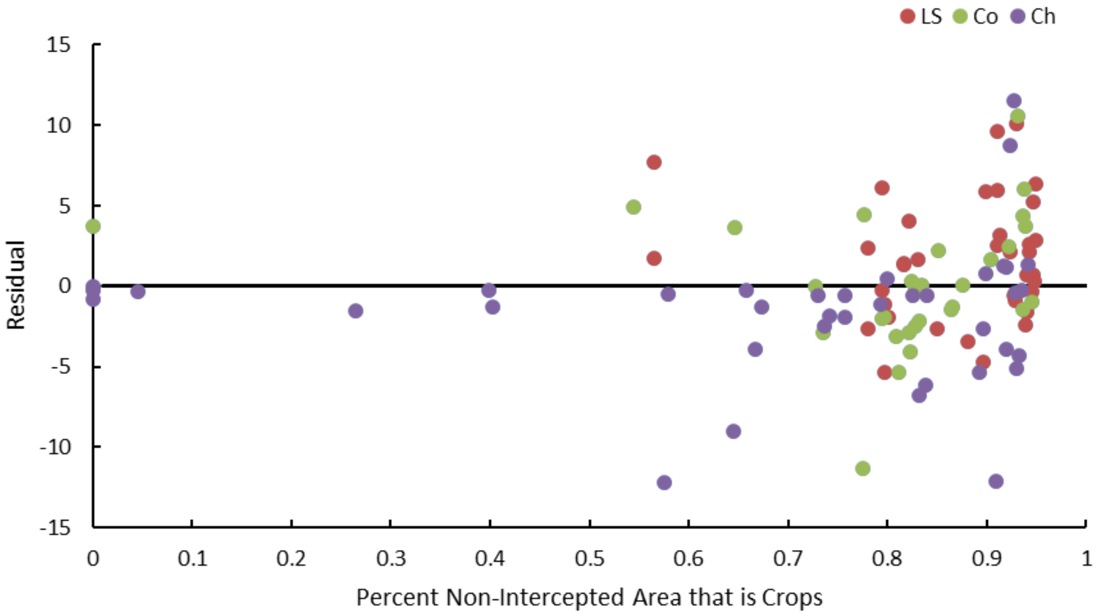


Figure 21. Residuals (from Figure 19) compared to the percent of non-intercepted land for each site that is crop land.

The same interception fraction values from Figure 15 were used to calculate the nitrate concentrations in Figure 22. The difference between Figure 16 and Figure 22 is the line of best fit equation used. Figure 22 used the line of best fit from Figure 18. Areas that are dark green have high predicted nitrate values and areas that are light green have low predicted nitrate values. Figure 22 follows the same trends as Figure 16 with the main flow paths having less nitrate output and the smaller tributaries have more. The nitrate values range from 0.29 mg/L to 16.5 mg/L due to the regression equation.

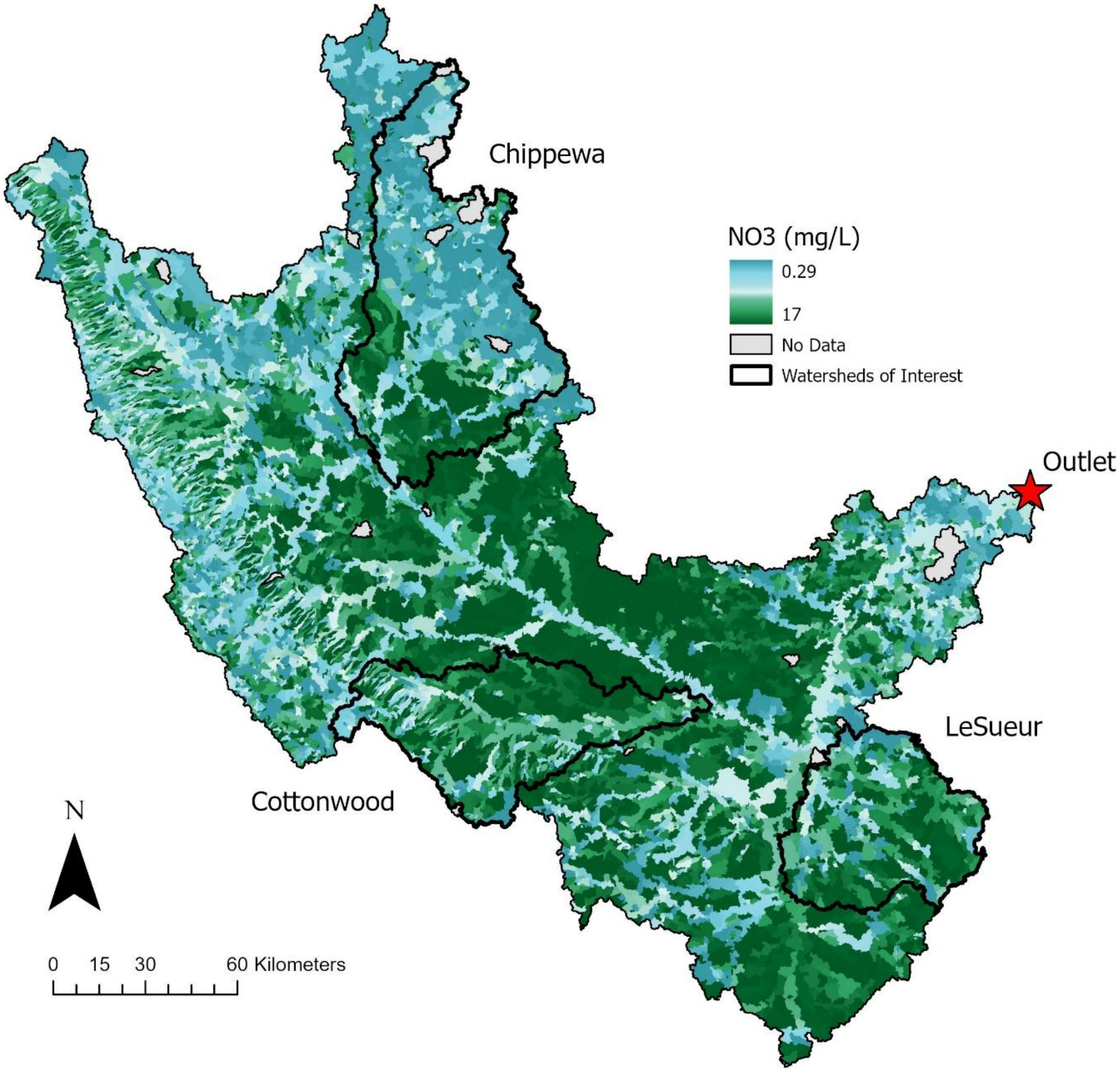


Figure 22. A map showing the nitrate concentration values as estimated using the linear equation in Figure 18 in the Minnesota River Basin.

4.3 Comparison of Predicted and Actual Nitrate Values

To clearly see the effect the different sampling years had on the data, Table 1 shows the average predicted and actual nitrate concentration values for the Minnesota River Basin, LeSueur, Cottonwood, and Chippewa. The predicted values for nitrate were higher than the sample nitrate concentration values for all watersheds except for LeSueur when the 2015 data was used. Chippewa is the most overpredicted in both iterations while the predicted values for the LeSueur are the closest to the sampled data in both. There is an closer agreement in the relationship between the predicted and sampled NO3 when sample data from only the same year, 2014, is used. The predicted average nitrate values for the Minnesota River Basin deviate slightly more from the average sample values when all the data from the same year is used. These results suggest that LeSueur has the highest sample average nitrate concentrations. However, when averaging across all of the links with the predicted values, Cottonwood has slightly higher nitrate concentrations than LeSueur. The root mean square errors for each watershed were higher when the 2015 data were used than when only the 2014 data were used.

Table 1. Comparison of actual and predicted nitrate values for the three watersheds that data were sampled in and an overall average for the Minnesota River Basin. It also shows the root mean square error (RMSE) for each watershed and the number of samples.

Year Collected	June 2014, 2015				June 2014			
Watershed	NO3 (mg/L)	NO3 Samples (mg/L)	RMSE (mg/L)	n	NO3 (mg/L)	NO3 Samples (mg/L)	RMSE (mg/L)	n
Minnesota River Basin	10.7	12.5	-	-	8.4	10.4	-	-
LeSueur	14.6	14.0	5.9	70	11.5	11.9	4.1	36
Cottonwood	10.0	14.9	4.5	30	10.0	12.8	4.0	30
Chippewa	4.1	8.8	5.6	38	4.1	6.9	4.6	38

5. Discussion

5.1 Combining Wetlands and Flowlines

Watersheds represent complex systems that require multiple datasets to analyze spatially. These datasets often come from multiple sources that were not made to be compatible with each other. Models such as the Nitrate Network Model (Czuba et al., 2018) or AgRiver (Hansen et al., 2021) are trying to predict and optimize the placement of wetlands to make wetland restoration optimal for the environment and help reduce the costs of the restoration. To do this, wetland and flow path data are required. Wetland data is typically derived from high quality imagery (U.S. Fish and Wildlife Service, 2019) while the flow paths are acquired from digital elevation models (EPA, 2019). In the United States comprehensive wetland data are most easily obtained from the National Wetland Inventory and flow path data are found readily available in the National

Hydrography Dataset Plus Version 2. The National Wetland Inventory is managed by the U.S. Fish and Wildlife Service, while the National Hydrography Dataset is under the Environmental Protection Agency.

In creating an integrated wetland-riverine network floodplain wetlands were removed because the models that would use this network do not use floodplain wetlands. When working to understand how nitrate is transported under these conditions, the integrated wetland-river network does not include floodplain wetlands because the inundation dynamics of these floodplains are not incorporated into the models that use it. There is no way to adequately assess the role of floodplain wetlands. Especially, in the case of calculating metrics such as interception fraction that looks at how much of a watersheds flow to its outlet is intercepted by wetlands that are active almost all year round. If floodplain wetlands were included, all the water to a watershed would be intercepted. In the National Wetland Inventory there is no indicator or classification for a floodplain wetland. Its classification is hierarchical with five main systems of which the Minnesota River Basin contains Lacustrine, Riverine, and Palustrine Systems (Federal Geographic Data Committee, 2013). None of these layers in the hierarchical system indicate if the wetland is located within a floodplain. The modifier that gets closest is the water regime modifier for nontidal wetlands. It provides a glimpse of the hydrologic regime of the wetland by classifying it by how much flooding it receives during the growing season (Federal Geographic Data Committee, 2013). The classification ranges from permanently flooded to intermittently flooded. However, these are not specific enough to determine if a wetland is flooded. Other issues can arise, such as if wetlands are misclassified. On the northwest side of the Minnesota River Basin there are some palustrine systems that were very long and snaky and looked like riverine systems. By looking at aerial imagery and National Hydrography Dataset flowlines, these wetlands were clearly misclassified and needed to be removed since riverine systems represent the rivers and streams, not wetlands.

5.2 Interception Fraction

Interception fraction, as introduced by Hansen et al. (2018; Figure 2b), is shown as a potentially better predictor of nitrate than percent wetland coverage. However, this was originally shown with only five data points specifically selected to have similar land cover, watershed size, and ephemeral wetland count. The reproduced graph of Hansen's work (Figure 9) shows similar trends to Figure 2b; however, the r-squared value was not as high. The main difference between the graphs was the interception fraction values, with the one with the most significant difference being D. More of the watershed was intercepted by wetlands in D. This was most likely due to the different ways in which the floodplain wetlands were removed. The results showed that a wetland's position on the network impacts its effectiveness.

Furthering Hansen's work was done by calculating interception fractions for all of the data points in the LeSueur, Cottonwood, and Chippewa watersheds within the Minnesota River Basin where samples had previously been taken (Figure 10, Figure 11). Figure 10 shows the amount of land within each watershed that is intercepted by wetlands. Figure 11 shows the interception values for each site each time a sample was taken in June of 2015 for LeSueur and for June 2014 for Cottonwood and Chippewa. It should be noted though that June 2014 was sampled on the falling limb of an event that had streamflows with a probability of exceedance less than 1% for all of the basins (Hansen et al., 2018). In contrast the streamflow measure in June 2015 was standard for the time of year. (Hansen et al., 2018). This difference in streamflow could account for the lower nitrate values for the samples taken in the Cottonwood and Chippewa compared to LeSueur since the large event will carry larger amounts of nutrients and will flush the nitrate quickly out of the system (Tank et al., 2021). The falling limb may have had less nitrogen in it since it may have been carried off in the peak flows. The r-squared value from the percent wetland nitrate relationship in Figure 10 is 0.47 which is worse than the r-squared value of 0.55 for the interception fraction nitrate relationship in Figure 11. This shows that interception fraction is a better metric for predicting nitrate than percent wetland.

Use of the 2015 data for LeSueur was to remain consistent with the data used in Figure 2 from Hansen et al (2018). However, the storm event that was noted to occur before the sampling in 2014 may have had an impact on the nitrate concentrations, so the same analysis was run on the 2014 data for all of the sites including LeSueur (Figure 18). The resulting data showed a better relationship than when the 2015 data was used for LeSueur (Figure 17). This again shows that interception fraction is a better metric for predicting nitrate than percent wetland. It also shows that using nitrate data that is collected under similar flow conditions is important to this relationship. The storm that occurred before data collection in 2014 most likely had an impact on the nitrate values as seen in the shift in the LeSueur nitrate concentrations between the two years.

The residuals were calculated for interception fraction as a predictor of nitrate and to two other potential metrics. These metrics were a wetland area to maximum drainage area ratio and percent of cropland not intercepted by wetlands. This shows that wetland size at the point of most flow interception has no impact on the predictive relationship of interception fraction to nitrate (Figure 20). The percentage of crop cover that is not intercepted by wetlands also had no impact (Figure 21). The majority of the Minnesota River Basin is cultivated crops (Figure 4) with the LeSueur and Cottonwood having 84% coverage and the Chippewa having 66% coverage. There is little variation in the landcover between these watersheds, which explains the lack of response to crop cover as a potential predictor variable. The variation that is seen in the Chippewa points also makes sense since it has less crop cover.

5.3 Predicted Nitrate Values in the Minnesota River Basin

Figure 16 and Figure 22 show the nitrate values predicted for the Minnesota River Basins based on the interception fraction shown in Figure 14 and their respective linear regression equations as shown in Figure 11 and Figure 18. Table 1 compares the average nitrate values with the predicted nitrates for each of the watersheds discussed in this study. It seems likely that the data collected in the LeSueur is most strongly influencing the interception fraction metric as a predictor of nitrate. When the average is calculated for 2015 in LeSueur and then again with only the 2014 data, the percent difference for LeSueur remains less than 5%. LeSueur does have more samples in it in 2015, but it only differs in agriculture landcover from Chippewa. LeSueur and Cottonwood had the same agriculture landcover. Chippewa may be different from its predicted values because it has a very different distribution of wetlands compared to LeSueur. It has the most wetland cover compared to the Cottonwood and LeSueur basins. When looking at the whole Minnesota River Basin the NO3 sample value was calculated using all the samples taken in LeSueur, Cottonwood, and Chippewa in the respective years. It decreases when the 2014 data was used for all the watersheds and the predicted value differed even more from the estimated one. However, this probably does not mean much since the nitrate sampled does not cover every watershed in the Minnesota River Basin. Overall, this indicates that the nitrate samples must be gathered under similar flow conditions for interception fraction to be considered as a predictor of nitrate. This map could also be used in place of traditional methods of visualizing nitrate concentrations across a watershed. This approach provides a rapid assessment of where nitrate might be high or low across a basin, instead of manually collecting data points across watersheds or building an entire model. For example, the Hypoxia Task Force could use a map like this to locate areas with high nitrate concentrations to identify locations where wetlands could be restored most effectively, or inversely it could inform where wetlands need to be conserved, since they are effectively removing nitrate.

5.4 Limitations and Future Work

The present analysis has a number of limitations. Relative to the automated extraction method, this approach has multiple numbers used as cut off values that will differ across different locations. These numbers were configured for the wetlands in the upper Midwest, but wetlands vary across the landscape and will look different in locations with different underlying topography. Additionally, this method takes advantage of the easily accessible, open access datasets provided by the Environmental Protection Agency and the U.S. Fish and Wildlife Service. In other countries, other datasets or information may be needed to create the necessary wetland and river information. Currently, floodplain wetlands are excluded, but in the future, these could be incorporated for better analysis once there is a better understanding of their inundation at different flows. The use of floodplain wetlands would require adjustments to the methods since they would need to be classified as floodplain instead of being removed. Including the floodplains might allow for data to be used that has been collected under different flow conditions. This analysis is also limited by the data collected at only certain places at certain times of the year. This data was collected primarily during the growing season with differing amounts of data in each watershed.

Future work for this study would include comparing these results to the results of the Nitrate Network Model or other similar models. Especially since this work indicated that the wetland that intercepted most of the watershed was not additionally predictive of nitrate concentrations. Floodplain wetland interactions could also be explored. Additionally, this study focused on nitrate; however, there were additional nitrogen samples collected in the dataset such as total nitrogen and nitrite, comparison of interception fraction to these could be valuable. This method will also need some work to be made more user friendly. Additionally, a method could be created to use this data to assess where wetlands could be placed to better intercept flow on the landscape.

6. Conclusions

Water quality has been degraded by excessive nitrate inputs. With the discovery of the Haber-Bosh process more nitrogen has become available for use with one of the primary uses being for agricultural production. Excess fertilizer is transported in runoff to waterways leading to eutrophication which can cause hypoxic conditions such as seen in the Gulf of Mexico. The Gulf of Mexico hypoxic zone’s main source of nitrogen is the agricultural region in the Upper Mississippi River Basin which contains the Minnesota River Basin, the location of this study. Wetlands are effective at removing nitrate from waterbodies; however, millions were removed to make way for agriculture. Wetland restoration is known to have the potential to reduce nitrate on small portions of the landscape. The next step is to understand how the isolated and flow-through wetlands within a watershed can work to reduce nitrate. There are multiple models that work to do this such as the Nitrate Network Model and the AgRiver Model. The key to the success of these models is developing an integrated wetland-river network that links wetlands to flowlines from different open access sources. A method was created to make analysis of wetland-river complexes easier and more accessible. A metric that looks at the potential for wetland impact is interception fraction as developed by Hansen et al. (2018). Interception fraction is the area a wetland intercepts divided by the watershed area of interest. It works to account for the positioning of wetlands within an upstream watershed to understand the nitrate concentration sampled at the outlet. In this study, it was found that interception fraction is a better predictor of nitrate than percent wetland area. The wetland size at the point of most watershed intercepted and percent of non-intercepted areas as crop land does not impact interception as a predictor of nitrate. After testing additional metrics against interception fraction, interception fraction was expanded to the entire Minnesota River Basin to estimate nitrate concentration with an r-squared value of 0.66. With this applied to the whole Minnesota River Basin the average nitrate concentration of the Minnesota River Basin is 10.4 mg/L. This helps identify more specifically where wetlands are present to remove nitrate and where they need to be restored to help achieve the Hypoxia Task Force goals. This map could also be used to see where wetlands could be targeted for conservation. Other factors that could impact nitrate concentrations at the sample sites would be landcover and timing after weather events. Future work would include potentially considering how floodplain wetlands may play a role or looking at different forms of nitrogen.

References

Cheng, F. Y., Van Meter, K. J., Byrnes, D. K., & Basu, N. B. (2020). Maximizing US nitrate removal through wetland protection and restoration. *Nature*, *588*(7839), 625-630. https://doi.org/10.1038/s41586-020-03042-5

Czuba, J. A., Hansen, A. T., Foufoula-Georgiou, E., & Finlay, J. C. (2018). Contextualizing wetlands within a river network to assess nitrate removal and inform watershed management. *Water Resources Research*, *54*(2), 1312-1337. https://doi.org/10.1002/2017WR021859

Dahl, T. E., and Allord, G. J., (1996). History of wetlands in the conterminous United States. *National Water Summary on Wetland Resources, U.S. Geol. Surv. Water Supply Pap., 2425.* U.S. Geol. Surv., Reston, VA. http://water.usgs.gov/nwsum/WSP2425/history.html.

David, M. B., Drinkwater, L. E., & McIsaac, G. F. (2010). Sources of nitrate yields in the Mississippi River Basin. *Journal of Environmental Quality*, *39*(5), 1657-1667. https://doi.org/10.2134/jeq2010.0115

Dolph C, Hansen A, Kemmit K, Janke B, Winikoff S, Boardman E, Baker A, Rorer M & Finlay J. (2017). Characterization of streams and rivers in the Minnesota River Basin Critical Observatory: nutrients, sediment, hydrology and biology from field sites in the Le Sueur, Chippewa and Cottonwood River Basins, 2013-2016. University of Minnesota. University of Minnesota Digital Conservancy. https://doi.org/10.13020/D6FH44

EPA. (2018). NHDPlus Upper Mississippi Data. Environmental Protection Agency. Washington, D.C. https://www.epa.gov/waterdata/nhdplus-upper-mississippi-data-vector-processing-unit-07

EPA. (2023). Mississippi River/Gulf of Mexico Watershed Nutrient Task Force 2023 Report to Congress. *United States Environmental Protection Agency.* https://www.epa.gov/system/files/documents/2023-11/10305_2023-htf-report-to-congress_508.pdf

Federal Geographic Data Committee. (2013). Classification of wetlands and deepwater habitats of the United States. FGDC-STD-004-2013. Second Edition. Wetlands Subcommittee, Federal Geographic Data Committee and U.S. Fish and Wildlife Service, Washington, DC. https://www.fws.gov/sites/default/files/documents/Classification-of-Wetlands-and-Deepwater-Habitats-of-the-United-States-2013.pdf

Foufoula-Georgiou, E., Takbiri, Z., Czuba, J. A., & Schwenk, J. (2015). The change of nature and the nature of change in agricultural landscapes: Hydrologic regime shifts modulate ecological transitions. Water Resources Research, 51(8), 6649-6671. https://doi.org/10.1002/2015WR017637

Frans, C., Istanbulluoglu, E., Mishra, V., Munoz-Arriola, F., & Lettenmaier, D. P. (2013). Are climatic or land cover changes the dominant cause of runoff trends in the Upper Mississippi River Basin?. Geophysical Research Letters, 40(6), 1104-1110. https://doi.org/10.1002/grl.50262

Hansen, A. T., Dolph, C. L., Foufoula-Georgiou, E., & Finlay, J. C. (2018). Contribution of wetlands to nitrate removal at the watershed scale. *Nature Geoscience*, *11*(2), 127-132. https://doi.org/10.1038/s41561-017-0056-6

Hansen, A. T., Campbell, T., Cho, S. J., Czuba, J. A., Dalzell, B. J., Dolph, C. L., ... & Wilcock, P. (2021). Integrated assessment modeling reveals near-channel management as cost-effective to improve water quality in agricultural watersheds. Proceedings of the National Academy of Sciences, 118(28), e2024912118. https://doi.org/10.1073/pnas.2024912118

Knowles, R. (1982). Denitrification. Microbiological reviews, 46(1), 43-70.

Martínez-Espinosa, C., Sauvage, S., Al Bitar, A., Green, P. A., Vörösmarty, C. J., & Sánchez-Pérez, J. M. (2021). Denitrification in wetlands: A review towards a quantification at global scale. *Science of the total environment*, 754, 142398. https://doi.org/10.1016/j.scitotenv.2020.142398

Marvel, K., Su, W., Delgado, R., Aarons, S., Chatterjee, A., Garcia, M. E., ... & Vose, R. S. (2023). Ch. 2. Climate trends. *Fifth National Climate Assessment*. Washington, DC, USA. https://doi.org/10.7930/NCA5.2023.CH2

Messer, T. L., Burchell II, M. R., Birgand, F., Broome, S. W., & Chescheir, G. (2017). Nitrate removal potential of restored wetlands loaded with agricultural drainage water: A mesocosm scale experimental approach. *Ecological Engineering*, *106*, 541-554. https://doi.org/10.1016/j.ecoleng.2017.06.022

MRLC. (2023). Data. Multi-Resolution Land Characteristics Consortium. https://www.mrlc.gov/data?f%5B0%5D=region%3Aconus

Natural Capital Project. (2025). InVEST 3.15.0. 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

Rabalais, N. N., & Turner, R. E. (2019). Gulf of Mexico hypoxia: Past, present, and future. *Limnology and Oceanography Bulletin*, 28(4), 117-124. https://doi.org/10.1002/lob.10351

Rabalais, Nancy N., and R. Eugene Turner. "Hypoxia in the northern Gulf of Mexico: description, causes and change." Coastal hypoxia: consequences for living resources and ecosystems 58 (2001): 1-36. https://doi.org/10.1029/CE058p0001

Randall, G. W., & Mulla, D. J. (2001). Nitrate nitrogen in surface waters as influenced by climatic conditions and agricultural practices. *Journal of Environmental Quality*, *30*(2), 337-344. https://doi.org/10.2134/jeq2001.302337x

Tank, J. L., Speir, S. L., Sethna, L. R., & Royer, T. V. (2021). The case for studying highly modified agricultural streams: farming for biogeochemical insights. Limnology and Oceanography Bulletin, 30(2), 41-47.

U.S. Fish and Wildlife Service. (2019). Wetlands Mapper Documentation and Instructions Manual. U.S. Fish and Wildlife Service, Ecological Services, National Standards and Support Team Madison, WI 53711-1061. https://www.fws.gov/sites/default/files/documents/Wetlands-Mapper-Documentation-Manual-May-2019.pdf

U.S. Fish and Wildlife Service. (2024). National Wetlands Inventory. U.S. Department of the Interior, Fish and Wildlife Service, Washington, D.C. https://www.fws.gov/page/national-wetlands-inventory-frequently-asked-questions#citing-digital-data-and-maps

Vitousek, P. M., Aber, J. D., Howarth, R. W., Likens, G. E., Matson, P. A., Schindler, D. W., ... & Tilman, D. G. (1997). Human alteration of the global nitrogen cycle: sources and consequences. *Ecological applications*, *7*(3), 737-750. https://doi.org/10.1890/1051-0761(1997)007[0737:HAOTGN]2.0.CO;2

Wollheim, W. M., Bernal, S., Burns, D. A., Czuba, J. A., Driscoll, C. T., Hansen, A. T., ... & Wohl, E. (2018). River network saturation concept: factors influencing the balance of biogeochemical supply and demand of river networks. Biogeochemistry, 141, 503-521. https://doi.org/10.1007/s10533-018-0488-0

Appendix A. Creating a Combined Wetland Flow Path Network

Editing Wetland Data from the National Wetland Inventory

Removing Floodplain Wetlands

Step 1: Downloading and Prepping Data

National Wetland Inventory Data Download
https://fwsprimary.wim.usgs.gov/wetlands/apps/wetlands-mapper/

Select "Get Data" and download by HUC. This ensures seamless integration if study area crosses state boundaries.

Add the NWI shapefiles to a map and merge them into one shapefile.

MRB is used to denote that it's the Minnesota River Basin.

Merge
Input Datasets: *NWI shapefiles, names will vary*
Output Dataset: "MRB_NWI"

Step 2: Remove the Riverine Wetlands

Select by Attributes
Input Rows: "MRB_NWI"
Selection Type: New Selection
Where WETLAND_TY Does Not Equal Riverine

Export Features
Input Features: "MRB_NWI"
Output Feature Class: "MRB_NWI_nR" Where "*nR*" means "*no Riverine*"

Add Field
Layer: "MRB_NWI_nR"
Name: Wetland_ID, Alias: Wetland_ID, Type: Double

Calculate Field
Field Name: "Wetland_ID"
Select "Sequential ID()" from the Helpers, scroll all the way down to

The next few steps are Minnesota River Basin Specific

If the National Wetland Inventory wetlands contain similar misclassified wetlands this section may be relevant (Figure A1). The idea is that they will have a distinctive length to width ratio and that can be used to identify and remove them with little impact on the rest of the data that we want to keep.

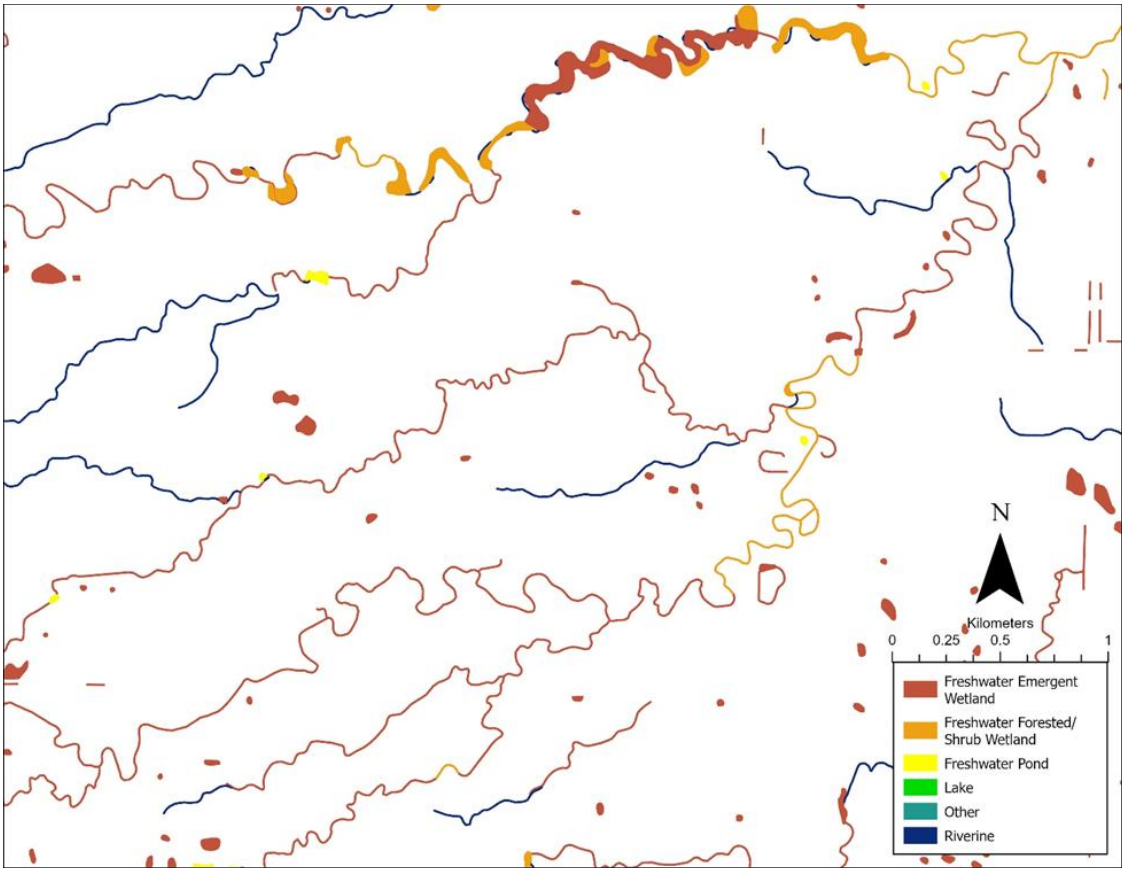


Figure A1. Misclassified wetlands in the National Wetland Inventory along the west side of the Minnesota River Basin

Step 3: Create Centerlines

For each of the remaining wetland types use *Select by Attributes* and then run *Polygon to Centerline.* Doing this will help minimize connector polylines that the *Polygon to Centerline* tool creates when creating centerlines.

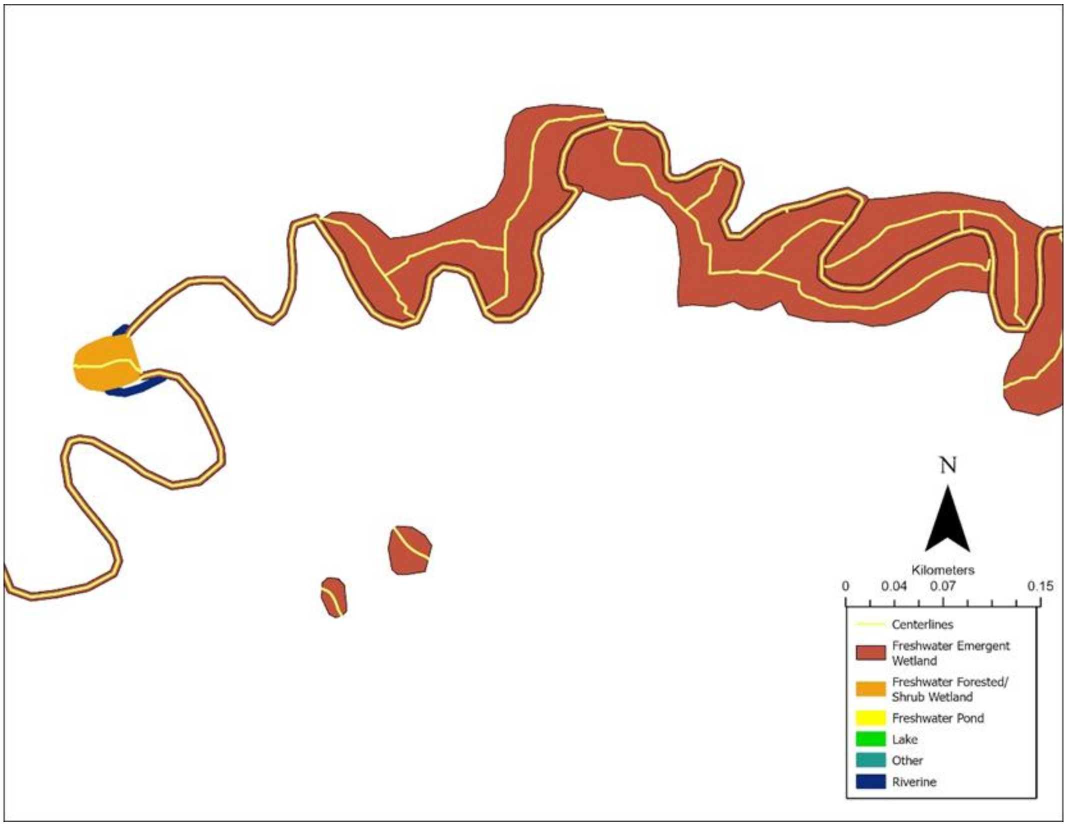


Figure A2. Some of the results of the Polygon to Centerline tool showing that it does not give just one centerline, but it should be sufficient to determine which polygons are supposed to riverine.

Polygon to Centerline
Input Features: “MRB_NWI_nR”
Output Feature Class: “MRB_NWI_nR_Lake” or “MRB_NWI_nR_Pond” or “MRB_NWI_nR_PFO” or “MRB_NWI_nR_PEM”
Connecting Features: LEAVE BLANK

The output feature class for *Polygon to Centerline* will vary based on which wetland type you have selected. Make sure to do this tool separately for each wetland type and then merge the shapefiles together at the end.

Merge
Input Datasets: “MRB_NWI_nR_Lake” and “MRB_NWI_nR_Pond” and “MRB_NWI_nR_PFO” and “MRB_NWI_nR_PEM”
Output Dataset: “MRB_NWI_nR_Centerlines”

Next add a field to calculate the length of the centerlines.

Add Field
Layer: “MRB_NWI_nR_Centerlines”

Name: LengthKM, Alias: LengthKM, Type: Double

Calculate Geometry
Input Features: "MRB_NWI_nR_Centerlines"
Field: LengthKM, Property: Length (Geodesic)
Length Unit: Kilometers

Step 3: Calculate Length to Width Ratio

Summarize Within
Input Polygons: "MRB_NWI_nR"
Input Summary Features: "MRB_NWI_nR_Centerlines"
Output Feature Class: "MRB_NWI_nR_SW"
Keep all input polygons
Summary Fields: "LengthKM", Sum
Group Field: leave empty

Add Field
Name: AreaSQKM, Alias: AreaSQKM, Type: Double
Name: WidthKM, Alias: WidthKM, Type: Double
Name: LW_Ratio, Alias: LW_Ratio, Type: Double

Calculate Geometry
Input: "MRB_NWI_nR_SW"
Field: AreaSQKM
Geometry Attributes: Area (Geodesic)
Length Unit: Square Kilometers

Calculate Field
Input: "MRB_NWI_nR_SW"
Field: WidthKM
Expression: AreaSQKM/LengthKM

Calculate Field
Input: "MRB_NWI_nR_SW"
Field: LW_Ratio
Expression: LengthKM/WidthKM

Step 4: Selecting and Removing Misclassified Wetlands

Select by Attributes
Where LW_Ratio is greater than or equal to 20
And WidthKM is less than 0.025

Doing this removes most the areas that we think are supposed to be riverine however some PEM wetlands in the northern section are removed.

The Following Should be Applicable to Outside of the MRB

Step 5: Downloading the NHDPlus Flowlines

NHDPlus Data Download:
https://www.epa.gov/waterdata/get-nhdplus-national-hydrography-dataset-plus-data

Select “07 Upper Mississippi” under List of Areas by Number (for this project location)
07 Upper Mississippi is broken into three parts. The Minnesota River Basin is in 07c.

Download the files with the following in their names
“…NHDSnapshot…”
“…FdrFac…”
“…NHDPlusAttributes…”

Link to NHDPlus User Guide:
https://www.epa.gov/system/files/documents/2023-04/NHDPlusV2_User_Guide.pdf

Step 6: Prepping NHD Data

Add Join
Input Table: NHD_Flowlines
Input Field: COMID
Join Table: PlusFlowlineVAA.dbf
Join Field: COMID

Select by Attributes
Where FLOWDIR is equal to Uninitialized
Or TerminalFl is equal to 1
Or Divergence is equal to 2

See NHD User Manual for Explanation

Remove these flowlines, by switching the selection and using Export Features
Export Features
Input Features: “NHD_Flowlines”
Output Feature Class: “NHD_Flowlines_Cleaned”

Create Field
Name: Link_ID, Alias: Link _ID, Type: Double

Calculate Field

Field Name: “Link_ID”
Select “Sequential ID()” from the Helpers, scroll all the way down to

Step 7: Buffer the NHD Flowlines
Buffer
Input Features: “NHD_Flowlines_Cleaned”
Output Feature Class: “NHD_Flowlines_Cleaned_B120m”
Distance: 120 Meters
Side Type: Full
End Type: Round
Method: Geodesic (shape preserving)
Dissolve Type: No Dissolve

Step 8: Clip NWI Wetlands
Clip
Input Feature or Dataset: “MRB_NWI_nR”
*C*lip Features: “NHD_Flowlines_Cleaned_B120m”
Output Features or Dataset: “MRB_NWI_nR_ClipB120m”

Create Field
Name: AREA_SQKM_B120m, Alias: AREA_SQKM_B120m, Type: Double

Calculate Geometry
Input: “MRB_NWI_nR_ClipB120m”
Field: AreaSQKM_B120m
Geometry Attributes: Area (Geodesic)
Length Unit: Square Kilometers

Step 9: Selecting Floodplain Wetlands
Join Field
Input Table: “MRB_NWI_nR”
Input Field: “Wetland_ID”
Join Table: “MRB_NWI_nR_ClipB120m”
Join Field: “Wetland_ID”
Transfer Method: Select transfer fields
Transfer Fields: AreaSQKM_B120m

Create Field
Name: FLD_Ratio, Alias: FLD_Ratio, Type: Double

Calculate Field
Input Table: “MRB_NWI_nR”
Field Name: FLD_Ratio

Expression: AreaSQKM / AreaSQKM_B120m

Select by Attributes
Where WETLAND_TY is not equal to Lake
Or WETLAND_TY is not equal to Freshwater Pond
Or FLD_Ratio is less than or equal to 5

The wetlands directly adjacent to a freshwater pond or lake also need to be selected. To do this run *Select by Attributes* on the NWI shapefile that has had no modification for freshwater ponds and lakes and use the selected features in the *Select by Location* tool to with the Share Boundaries selection tool on the MRB_NWI_nR layer and add those wetlands to the selection.

Step 10: Switch the Selection and Export Features
Export Features
Input Features: “MRB_NWI_nR”
Output Feature Class: “MRB_NWI_nR_nFLD”

Step 11: Removing Last Few Floodplain Wetlands
This is being done to remove small wetlands with really large drainage areas. This means that they are in the floodplain.

Zonal Statistics as Table
Input Raster or Feature Zone Data: “MRB_NWI_nR_nFLD”
Zone Field: “Wetland_ID” (must be an integer or string, may need to modify field type)
Input Value Raster: “FAC”
Output Table: “MRB_wetland_MAX_FAC”
Statistics Type: Maximum

Join Field
Input Table: “MRB_NWI_nR_nFLD”
Input Field: “Wetland_ID”
Join Table: “MRB_wetland_MAX_FAC”
Join Field: “Wetland_ID”
Transfer Method: Select transfer fields
Transfer Fields: MAX

Create Field
Name: Wetland_DA, Alias: Wetland_DA, Type: Double
Name: WDA_Ratio, Alias: WDA_Ratio, Type: Double

Calculate Field
Input Table: “MRB_NWI_nR_nFLD”
Field Name: Wetland_DA
Expression: MAX * 30 * 30 / 100000 (converts 30m by 30m pixel to SKQM)

Calculate Field
Input Table: “MRB_NWI_nR_nFLD”
Field Name: WDA_Ratio
Expression: Wetland_DA / Area_SQKM

Select by Attributes
Where WDA_Ratio is less than than 2000

Export Features
Input Features: “MRB_NWI_nR_nFLD”
Output Feature Class: “MRB_Wetlands_nFLD”

Step 12: Create Multipart Wetland Features

This is done to ensure that wetlands that are separated by a riverine feature, that should probably be considered a single wetland complex are considered a single wetland.

Buffer
Input Features: “MRB_Wetlands_nFLD”
Output Feature Class: “MRB_Wetlands_nFLD_15m”
Distance [Value or Field]: 15
Linear Unit: Meter
Side Type: Full
End Type: Round
Method: Planar
Dissolve Type: No Dissolve

Dissolve
Input Features: “MRB_Wetlands_nFLD_15m”
Output Feature Class: “MRB_Wetlands_nFLD_15m_dissolve”
Leave Dissolve Fields and Statistics Field Blank
Do no create multipart features

Calculate Field
Input Table: “MRB_Wetlands_nFLD_15m_dissolve”
Field Name: “Wetland_ID”
Field Type: Double
Expression: Sequential_ID()

Spatial Join
Target Features: “MRB_Wetlands_nFLD”
Join Features: “MRB_Wetlands_nFLD_15m_dissolve”
Output Feature Class: “MRB_Wetlands_nFLD_multi”
Join Operation: Join one to one

Dissolve
Input Features: “MRB_Wetlands_nFLD_multi”
Output Feature Class: “MRB_Wetlands_nFLD_multi_dis”
Dissolve Fields: “Wetland_ID”
Leave Statistics Field Blank
Create Multipart Features

Step 13: Calculating Percent Emergent

Select by Attributes
Input Rows: “MRB_NWI”
Selection Type: New Selection
Where WETLAND_TY is equal to Freshwater Emergent Wetland

Calculate Geometry
Input Features: “MRB_NWI”
Use Selected
Field (Existing or New): PEM_AreaSQKM
Property: Area (geodesic)
Area Unit: Square Kilometers

Summarize Within
Input Polygons: “MRB_Wetlands_nFLD_multi_di”
Input Summary Features: “MRB_NWI”, use selected
Output Feature Class: “MRB_Wetlands_pEM”
Keep all input polygons
Summary Fields: PEM_AreaSQKM, Sum

Calculate Geometry
Input Features: “MRB_Wetlands_pEM”
Field (Existing or New): Wetland_AreaSQKM
Property: Area (geodesic)
Area Unit: Square Kilometers

Calculate Field
Input Table: “MRB_Wetlands_pEM”
Field Name (Existing or New): pEM
Field Type: Double

Expression Type: Python
Expression: PEM_AreaSQKM_SUM / Wetland_AreaSQKM * 100

Creating Flowline Wetland Network

Step 1: Use the NHD FAC Raster to Create Flowlines

Raster Calculator
Map Algebra Expression: “MRB_FAC” >= 500
Output Raster: “MRB_FAC_500”

Reclassify
Input Raster: “MRB_FAC_500”
Reclass field: Value
Reclassification:

Value	New
0	NODATA
1	1
NODATA	NODATA

Output Raster: “MRB_FAC_500_stream”

Stream to Feature
Input stream raster: “MRB_FAC_500_stream”
Input flow direction raster: “fdr” (from NHD)
Output polyline feature: “MRB_FAC_500_stream_line”

Step 2: Remove Lines that do not Relate to the NHD

Remove lines that do not intersect with the NHD within 30 meters. Create points on dangling vertices and select flowlines that intersect with those points. Out of those lines remove from the selection the ones that do not have their centers within 30 meters of the NHD flowlines. Then remove the ones from the selection that are longer than 0.15 kilometers.

<u>Selecting the new flowlines the correspond to the NHD</u>
Select by Location
Input Features: “MRB_FAC_500_stream_line”
Relationship: Intersect
Selecting Features: “NHD_Flowlines_Cleaned”
Search Distance: 30 Meters
Selection Type: New selection

Export Features
Input Features: “MRB_FAC_500_streamline”
Output Feature Class: “MRB_FAC_500_streamline_NHDi”

Some lines might need to be added to the selection by hand if the FAC flowlines deviate from the NHD flowlines. See image below

Figure A3. A map showing that there may exist flowlines that do not agree with the NHD despite being made directly from the flow accumulation provided with the NHD flowlines. This needs to be selected by hand.

Feature Vertices to Points
Input Features: “MRB_FAC_500_stream_line”
Output Feature Class: “MRB_FAC_500_stream_line_NHDi_dangling”
Point Type: Dangling vertex

Select by Location
Input Features: “MRB_FAC_500_streamline_NHDi”
Relationship: Intersect
Selecting Features: “MRB_FAC_500_streamline_NHDi_dangling”
Search Distance: leave blank
Selection Type: New selection
Check invert selection

Export Features
Input Features: “MRB_FAC_500_streamline_NHDi”
Output Feature Class: “MRB_FAC_500_streamline_NHDi_Dr”

Feature Vertices to Points

Input Features: “MRB_FAC_500_streamline_NHDi_Dr”
Output Feature Class: “MRB_FAC_500_streamline_NHDi_dangling2”
Point Type: Dangling vertex

Select by Location
Input Features: “MRB_FAC_500_streamline_NHDi_Dr”
Relationship: Intersect
Selecting Features: “MRB_FAC_500_streamline_NHDi_dangling2”
Search Distance: 30 Meters
Selection Type: New selection

Select by Attributes
Input Features: “MRB_FAC_500_streamline_NHDi_Dr”
Selection Type: Remove from the current selection
Where LengthKM is greater than or equal to 0.15

Delete the selected values.

Step 3: Unsplit the features

The unsplit line tool can be finicky make sure to check your features as you go to ensure they unsplit correctly. Use the MODE of the slope to get the slope that is found most frequently for that line segment

Unsplit Line
Input Features: “MRB_FAC_500_streamline_NHDi_Dr”
Output Feature Class: “MRB_FAC_500_streamline_NHDi_Dr_unsplit”
Leave dissolve and statistics fields blank

Step 4: Determine Slope

The basic idea for this is to determine which NHD flowline is most similar to the flowline that we created. If there are multiple NHD flowlines after summarized within is preformed, select only those lines generate points along the lines at a smaller interval and summarize within again. Repeat until each line has a slope. If a line has multiple potential slopes repeat the process Buffer, Generate Points Along Line and Summarize Within increasing the frequency of the points on just those lines or select the slope by hand.

Buffer
Input Features: “MRB_FAC_500_streamline_NHDi_Dr_unsplit”
Output Feature Class: “MRB_FAC_500_streamline_NHDi_Dr_buffer30m”
Distance: 30 Meters
Side Type: Full
End Type: Flat
Method: Planar

Dissolve Type: No Dissolve

Generate Points Along Line
Input Features: “NHD_Flowlines_Cleaned”
Output Feature Class: “NHD_gpal_10m”
Point Placement: By Distance
Distance: 10 Meters
Do not include end points or add accumulated distance and sequence fields
Distance Method: Geodesic

Summarize Within
Input Polygons: “MRB_FAC_500_streamline_NHDi_Dr_buffer30m”
Input Summary Features: “NHD_gpal_10m”
Output Feature Class: “MRB_FAC_500_B30_SW”
Keep all input polygons
Summary Fields: COMID, both Maximum and Minimum
Group Field: COMID
Add minority and majority attributes
Add group percentages
Output Grouped Table: “COMID_Summary”

Join all resulting Summarize Within to the flowlines shapefile that was buffered and then create a new field that all of the COMID will be added into. Then using that field add the slope from the NHD lines.

Join Field
Input Table: “MRB_FAC_500_streamline_NHDi_Dr_unsplit”
Input Field: the unique ID field used
Join Table: “MRB_FAC_500_B30_SW”
Join Field: the unique ID field used
Transfer Method: Select transfer fields
Transfer Fields: Majority COMID

Step 5: Split the features and remove wetlands

Project
Input Dataset or Feature Class: “MRB_FAC_500_streamline_NHDi_Dr_unsplit”
Output Dataset or Feature Class: “MRB_FAC_500_unsplit_project”
Output Coordinate System*: GCS_North_American_1983
*must be the same coordinate system as the wetland shapefile

Export Features
Input Features: “MRB_FAC_500_streamline_NHDi_Dr”
Output Feature Class: “MRB_FAC_500_streamline_NHDi_Dr_splitfeatures”

Split Features
This tool modifies the Target Features
Cutting and Target Features must be in the same projection
The Cutting features must be selected; select the ones that intersect the stream network
Cutting Features: “MRB_Wetlands”
Target Features: “MRB_FAC_500_streamline_NHDi_Dr_splitfeatures”

Select by Location
Input Features: “MRB_FAC_500_streamline_NHDi_Dr_splitfeatures”
Relationship: Intersect
Selecting Features: “MRB_Wetlands”
Search Distance:
Selection Type: New selection

Switch the selection and then export features

Export Features
Input Features: “MRB_FAC_500_streamline_NHDi_Dr_sf_noWetland”
Output Feature Class: “MRB_FAC_500_streamline_NHDi_Dr_sf_noWetland”

Step 5: Break the features into segments no longer than 5 kilometers

This step will be repeated multiple times until the segments are no longer than 5 kilometers. Length will be calculated for each line segment using calculate geometry. Then all the lengths greater than 5 kilometers will be selected and midpoints will be created along these links. Split the network at these points. Keep doing this until there are no lengths longer than 5 kilometers. For the Minnesota River Basin this was done three times.

Project
Input Dataset or Feature Class: “MRB_FAC_500sl_NHDi_Dr_sf_nW_unsplit”
Output Dataset or Feature Class:
“MRB_FAC_500sl_NHDi_Dr_sf_nW_unsplit_projected”
Output Coordinate System: NAD 1983 UTM Zone 15N

Calculate Geometry Attributes
Input Features: “MRB_FAC_500sl_NHDi_Dr_sf_nW_unsplit_projected”
Geometry Attributes: Field: “LengthKM”, Property: Length (geodesic)
Length Unit: Kilometer

Select by Attributes:
Input Rows: “MRB_FAC_500sl_NHDi_Dr_sf_nW_unsplit_projected”
Selection Type: New selection
Where LengthKM is greater than 5

Feature Vertices to Points
Input Features: “MRB_FAC_500sl_NHDi_Dr_sf_nW_unsplit_projected”
Output Feature Class: “MRB_FAC_500sl_NHDi_Dr_sf_nW_unsplit_projected_mdpt”
Point Type: Midpoint

Split Line at Point
Input Features: “MRB_FAC_500sl_NHDi_Dr_sf_nW_unsplit_projected”
Point Features: “MRB_FAC_500sl_NHDi_Dr_sf_nW_unsplit_projected_mdpt”
Output Feature Class:
“MRB_FAC_500sl_NHDi_Dr_sf_nW_unsplit_projected_mdpt_splt”

Once the links are properly shortened, run the project tool to put the stream network back into a geodesic coordinate system.

Project
Input Dataset or Feature Class:
“MRB_FAC_500sl_NHDi_Dr_sf_nW_unsplit_projected_mdpt_splt”
Output Dataset or Feature Class: “stream_network_noIDs”
Output Coordinate System: NAD_1983_Albers

Add four new fields: “GridID_Up”, “GridID_Down”, “Inconclusive”, “ID_new”
“GridID_Up”, “GridID_Down” => Field Type: Double
“Inconclusive” => Field Type: Text
“ID_new” => Field Type: Double, Calculate Field and use Sequential Number

Step 6: Create nodes to determine downstream vs upstream for each link

Feature Vertices to Points
Input Features: “stream_network_noIDs”
Output Feature Class: “stream_network_noIDs_startvertex”
Point Type: Start vertex

Feature Vertices to Points
Input Features: “stream_network_noIDs”
Output Feature Class: “stream_network_noIDs_endvertex”
Point Type: End vertex

Select by Location
Input Features: “stream_network_noIDs_endvertex”
Relationship: Intersect
Selecting Features: “stream_network_noIDs_startvertex”

Merge

Input Datasets: “stream_network_noIDs_endvertex”,
“stream_network_noIDs_startvertex”
Output Dataset: “stream_network_nodes”

Extract Multi Values to Points
Input point features: “stream_network_nodes”
Input rasters: “MRB_FAC”
Output field name: “FAC”

Add Spatial Join
Target Features: “stream_network_nodes”
Join Features: the wetland features
Check Keep All Target Features
Match Option: Intersect
Check Permanently Join fields
Under the fields tab, remove all fields except the unique wetland IDs, this is the only field needed

Create a new field. Select points that intersect the wetlands and put them in that field. Then invert the selection and using the sequential number helper starting at 200000 for the nodes that do not intersect wetlands to clearly differentiate between the two.

Step 7: Run code to assign upstream and downstream links
Names and file locations will need to be changed (Figure A4)

```
import arcpy
from arcpy.sa import *

import pandas as pd
from arcgis.features import GeoAccessor, GeoSeriesAccessor

arcpy.env.workspace = r"C:\Users\cvill\Desktop\Grad School\Nitrate\GIS\Nitrate_Analysis_2"

#reading in shapefiles
Point_shapefile = "mrb_fac_500_st_Feature_Merge3"
Line_shapefile = "mrb_fac_500_st_Split_Project1"

points = pd.DataFrame.spatial.from_featureclass(Point_shapefile)
lines = pd.DataFrame.spatial.from_featureclass(Line_shapefile)

fid = lines["ID_new"].tolist()
print(fid)

Point_shapefile = "mrb_fac_500_st_Feature_Merge3"
Line_shapefile = "mrb_fac_500_st_Split_Project1"

points = pd.DataFrame.spatial.from_featureclass(Point_shapefile)
lines = pd.DataFrame.spatial.from_featureclass(Line_shapefile)

fid = lines["ID_new"].tolist()
print(fid)

#assigning link ids
for i in range(len(fid)):
    current_ID = fid[i]
    where_clause = '"ID_new" = ' + str(current_ID)

    arcpy.management.SelectLayerByAttribute(Line_shapefile,"NEW_SELECTION",where_clause)

    arcpy.management.SelectLayerByLocation(Point_shapefile,"INTERSECT",Line_shapefile)

    p = pd.DataFrame.spatial.from_featureclass(Point_shapefile)

    if str(p["FAC"].iloc[0]) == "<NA>" or str(p["FAC"].iloc[1]) ==  "<NA>":
        lines["inconclusive"].iloc[current_ID-1] = "no RasterValue"
    elif p["FAC"].iloc[0] == p["FAC"].iloc[1]:
        lines["inconclusive"].iloc[current_ID-1] = "same RasterValue"
    else:
        if p["FAC"].iloc[0] >= p["FAC"].iloc[1]:
            lines["grid_id_down"].iloc[current_ID-1] = p["Grid_ID"].iloc[0]
            lines["GridID"].iloc[current_ID-1] = p["Grid_ID"].iloc[1]
        elif p["FAC"].iloc[0] <= p["FAC"].iloc[1]:
            lines["Grdi"].iloc[current_ID-1] = p["Grid_ID"].iloc[0]
            lines["grid_id_down"].iloc[current_ID-1] = p["Grid_ID"].iloc[1]

    print(current_ID)

#exporting to shapefile
lines.spatial.to_featureclass("MRB_Flowlines_Links_sameR_new2")
```

Figure A4. Code needed for step 7.

Step 8: Assign upstream and downstream points to each link to links that have the same raster value

Select only the lines that have “SameRasterValue” in the inconclusive column. Generate Points Along Lines on only those lines and keep only the start and end points. Run the code (Figure A5).

```
#short Links
lines = pd.DataFrame.spatial.from_featureclass("MRB_Flowlines_Links_sameR_new1_removeisolated_inwet_dis_short")

fid = lines["id_new"].tolist()
print(fid)

for i in range(len(fid)):
    current_ID = fid[i]
    where_clause = '"id_new" = ' + str(current_ID)

    arcpy.management.SelectLayerByAttribute("MRB_Flowlines_Links_sameR_new1_removeisolated_inwet_dis_short","NEW_SELECTION",where_clause)

    arcpy.management.SelectLayerByLocation("MRB_Flowlines_Li_SpatialJoin1","INTERSECT","MRB_Flowlines_Links_sameR_new1_removeisolated_inwet_dis_short")

    p = pd.DataFrame.spatial.from_featureclass("MRB_Flowlines_Li_SpatialJoin1")

    if str(p["ORIG_LEN"].iloc[0]) == "<NA>" or str(p["ORIG_LEN"].iloc[1]) ==  "<NA>":
        lines["inconclusive"].iloc[current_ID-1] = "no RasterValue"
    elif p["ORIG_LEN"].iloc[0] == p["ORIG_LEN"].iloc[1]:
        lines["inconclusive"].iloc[current_ID-1] = "same RasterValue"
    else:
        if p["ORIG_LEN"].iloc[0] >= p["ORIG_LEN"].iloc[1]:
            lines["grid_id_down"].iloc[current_ID-1] = p["Grid_ID"].iloc[0]
            lines["GridID"].iloc[current_ID-1] = p["Grid_ID"].iloc[1]
        elif p["ORIG_LEN"].iloc[0] <= p["ORIG_LEN"].iloc[1]:
            lines["GridID"].iloc[current_ID-1] = p["Grid_ID"].iloc[0]
            lines["grid_id_down"].iloc[current_ID-1] = p["Grid_ID"].iloc[1]

    print(current_ID)
```

Figure A5. Code needed determine linkage in step 8.

Step 9: Remove isolated flowlines and flowlines that connect the same wetland

To remove flowlines that connect the same wetland

Select by Attributes
Input Rows: “MRB_Flowline_Links” (final output of code)
Selection Type: New selection
Where GridID is equal to grid_id_down

To remove flowlines within holes in the watershed, generate a watershed at the outlet and then do an inverse select by location selection on the lines within the polygon.

Step 10: Determine flowline downstream of wetland

Select Layer by Location
Input Features: “MRB_Flowline_Links”
Relationship: Intersect
Selecting Features: “MRB_Wetlands_pEM”
Search Distance: 0.5 Meters (If for some reason there is a really small gap between the polygon and lines this will grab it)
Selection Type: New selection

Feature Vertices To Points
Input Features: “MRB_Flowline_Links”
Output Feature Class: “MRB_Flowline_inWet_start”
Point Type: Start vertex

Spatial Join
Target Features: “MRB_Flowline_inWet_start”
Join Features: “MRB_Wetlands_pEM”
Output Feature Class: “MRB_Flowline_inWet_start_SJ”
Join Operation: Join one to one
Keep all Target Features
Match Option: Intersect
Search Radius: 0.5 Meters
Only need to keep the unique wetland ID field

Extract Multi Values to Points
Input point features: “MRB_Flowline_inWet_start_SJ”
Input rasters: “fac”, “fac”

Summary Statistics
Input Table: “MRB_Flowline_inWet_start_SJ”
Output Table: “MRB_Flowline_inWet_start_SJ_maxFAC”
Statistics Fields: fac
Statistic Type: Maximum
Case Fields: Wetland_ID

Join Field
Input Table: “MRB_Flowline_inWet_start_SJ”
Input Field: “Wetland_ID”
Join Table: “MRB_Flowline_inWet_start_SJ_maxFAC”
Join Field: “Wetland_ID”
Transfer Method: Select transfer fields
Transfer Fields: MAX_fac
Index Join Fields: Do not add indexes

Select Layer by Attributes
Input Rows: “MRB_Flowline_inWet_start_SJ”
Selection Type: New selection
Where MAX_fac is equal to fac

Join Field
Input Table: “MRB_Wetlands_pEM”
Input Field: “Wetland_ID”

Join Table: “MRB_Flowline_inWet_start_SJ”
Join Field: “Wetland_ID”
Transfer Method: Select transfer fields
Transfer Fields: “GridID” (the unique link ID)
Index Join Fields: Do not add indexes

Downstream IDs have been assigned to wetlands.
Export Features
Input Features: “MRB_Flowline_inWet_start_SJ”
Use current selection
Output Feature Class: “MRB_Wetland_downstreamPoints”

Step 11: Assign downstream point at junctions
Calculate Field
Input Table: “MRB_Flowline_Links”
Field Name: “Count_Junc”
Field Type: Double
Expression Type: Python
Expression: 1

Summary Statistics
Input Table: “MRB_Flowline_Links”
Output Table: “MRB_Flowline_Links_junc_count”
Field: “Count_junc”
Statistic Type: “Sum”
Case Fields: “GridID”

Join Field
Input Table: “MRB_Flowline_Links”
Input Field: “GridID”
Join Table: “MRB_Flowline_Links_junc_count”
Join Field: “GridID”
Transfer method: Select transfer fields
Transfer Fields: “SUM_Count_junc”

Select by Attributes
Input Rows: “MRB_Flowline_Links”
Selection Type: New selection
Where Sum_Count_junc is greater than or equal to 2

Generate Points Along Lines
Input Features: “MRB_Flowline_Links”
Use selected

Output Feature Class: “MRB_Flowline_Links_junc”
Point Placement: By distance
Distance: 15 meters
Include end points
Add accumulated distance and sequence fields
Distance Method: Geodesic

Extract Multi Values to Points
Input point features: “MRB_Flowline_Links_junc”
Input rasters: “fac”, “fac”

Summary Statistic
Input Table: “MRB_Flowline_Links_junc”
Output Table: “MRB_Flowline_Links_startpoint”
Field: “fac”
Statistic Type: “Maximum”
Case Fields: “GridID”

Join Field
Input Table: “MRB_Flowline_Links_junc”
Input Field: “GridID”
Join Table: “MRB_Flowline_Links_startpoint”
Join Field: “GridID”
Transfer method: Select transfer fields
Transfer Fields: “MAX_fac”

This next select by attributes will make sure the start point for a flowline that flows into a junction is not the same as they other flowlines flowing into the junction. Inverting the selection will select all the points that don’t have the same start as the starting point.

Select Layer by Attributes
Input Rows: “MRB_Flowline_Links_junc”
Selection Type: New selection
Where MAX_fac is equal to fac
Invert the Selection

Summary Statistic
Input Table: “MRB_Flowline_Links_junc”
Output Table: “MRB_Flowline_Links_newMAX”
Field: “fac”
Statistic Type: “Maximum”
Case Fields: “GridID”

Join Field
Input Table: “MRB_Flowline_Links_junc”
Input Field: “GridID”
Join Table: “MRB_Flowline_Links_newMAX”
Join Field: “GridID”
Transfer method: Select transfer fields
Transfer Fields: “MAX_fac”

Select Layer by Attributes
Input Rows: “MRB_Flowline_Links_junc”
Selection Type: New selection
Where MAX_fac is equal to fac (MAX_fac needs to be the new one, will most likely have a 1 in the name)

Export Features
Input Features: “MRB_Flowline_Links_junc”
Use current selection
Output Feature Class: “MRB_Flowline_Links_juncpoints”

Step 12: Determine incremental watershed sizes for each link and wetland

Select Layer by Attributes
Input Rows: “MRB_Flowline_Links”
Selection Type: New selection
Where Sum_Count_junc is equal to 1

Feature Vertices To Points
Input Features: “MRB_Flowline_Links”
Use selected
Output Feature Class: “MRB_Flowline_Links_endpoints”
Point Type: End vertex

Merge
Input Datasets: “MRB_Flowline_Links_endpoints”
“MRB_Flowline_Links_juncpoints”
“MRB_Wetland_downstreamPoints”
Output Dataset: “MRB_downstreampoints”

If the wetland and flowlines have different ID fields make sure to combine them before running the following commands.

Snap Pour Point
Input raster or feature pour point data: “MRB_downstreampoints”

Pour point field: GridID
Input accumulation raster: “fac”
Output raster: “MRB_spp”

Watershed
Input D8 flow direction raster: “fdr”
Input raster or feature pour point data: “MRB_spp”
Pour point field: “Value”
Output raster: “MRB_inc_watershed”

Raster to Polygon
Input raster: “MRB_inc_watershed”
Field: “Value”
Output polygon features: “MRB_inc_watershed_polygon”
Uncheck simplify polygons
Check create multipart features

Calculate Geometry
Input Features: “MRB_inc_watershed_polygon”
Geometry Attributes
Field: incDA_SQKM
Property: Area (geodesic)
Area Unit: Square Kilometers

Step 13: Determine watershed size for each link

Extract Multi Values to Points
Input point features: “MRB_downstreampoints”
Input rasters: “fac”, “fac”

Calculate Field
Input Table: “MRB_downstreampoints”
Field Name (Existing or New): “TotalDA_SQKM”
Field Type: Double
Expression Type: Python
Expression: “fac” * 30 * 30 / 100000 (converts pixel count to square kilometer)

Step 14: Rejoin data to wetlands and flowlines

Use the Join Field tool and join the “TotalDA_SQKM” from the “MRB_downstreampoints” to the wetland and flowline shapefiles based on the unique link ID which should have been carried through. Use the Join Field tool and join the “incDA_SQKM” from the “MRB_inc_watershed_polygon” to the wetland and flowline shapes based on the unique link ID.

At this point if just the wetlands on the network are wanted just run a select by locations based on the flowlines.

Appendix B. Calculating Interception Fraction

Step 1: Create site watersheds by running the watershed tool separately for each sampling site.

Step 2: Estimate wetland outlet using the flow accumulation raster

Raster to Point
Input raster: “fac”
Field: value
Output point features: “fac_point”

Summarize Within
Input Polygons: wetland feature class in use
Input Summary Features: “fac_point”
Output Feature Class: “wetland_fac_max_SW”
Uncheck keep all input polygons
Summary Fields: grid_code
Statistic: Maximum

Spatial Join
Target Features: “fac_points”
Join Features: “wetland_fac_max_SW”
Output Feature Class: “wetland_fac_max_SW_SJ”
Join Operation: Join one to one
Match Option: Intersect

Select by Attribute
Input Rows: “wetland_fac_max_SW_SJ”
Selection Type: New selection
Where grid_code is equal to Maximum gride_code

Export Features
Input Features: “wetland_fac_max_SW_SJ”
Output Feature Class: “wetland_outfalls_LW”

Delete Identical
Input Dataset: “wetland_outfalls_LW”
Field(s): unique wetland identifier field

Delete rows where Maximum grid_code is equal to zero

Select by Location
Input Features: wetland feature class in use
Relationship: Intersect
Selecting Features: “wetland_fac_max_SW”
Invert the selection

Summarize Nearby
Input Features: wetland feature class in use
Use current selection
Input Summary Features: “fac_point”
Output Feature Class: “wetland_fac_max_SN”
Distance Measurement: Straight line
Distances: 30
Distance Units: Meters
Uncheck keep all input polygons
Summary Fields: grid_code
Statistic: Maximum
Uncheck add shape summary attributes

Spatial Join
Target Features: “fac_points”
Join Features: “wetland_fac_max_SN”
Output Feature Class: “wetland_fac_max_SN_SJ”
Join Operation: Join one to one
Match Option: Intersect

Select by Attribute
Input Rows: “wetland_fac_max_SN_SJ”
Selection Type: New selection
Where grid_code is equal to Maximum gride_code

Export Features
Input Features: “wetland_fac_max_SN_SJ”
Output Feature Class: “wetland_outfalls_SN”

Delete Identical
Input Dataset: “wetland_outfalls_SN”
Field(s): unique wetland identifier field

Delete rows where Maximum grid_code is equal to zero

Step 3: Create drainage areas for the wetlands

Merge
Input Datasets: “wetland_outfalls_SW”, “wetland_outfalls_SN”
Output Dataset: “wetland_outfalls”

Snap Pour Point
Input raster or feature pour point data: “wetland_outfalls”
Pour point field: unique wetland identifier field
Input accumulation raster: “fac”
Output raster: “wetland_outfalls_SPP”

Watershed
Input D8 flow direction raster: “fdr”
Input raster or feature pour point data: “wetland_outfalls_SPP”
Pour point field: Value
Output raster: “wetland_watershed”

Raster to Polygon
Input raster: “wetland_watershed”
Field: Value
Output polygon features: “wetland_watershed_poly”
Uncheck simplify polygons
Create multipart features

Calculate Geometry
Input Features: “wetland_watershed_poly”
Geometry Attributes Field (Existing or New): wetland_DA_SQKM
Property: Area (geodesic)
Area Unit: Square Kilometer

Step 4: Calculate interception fraction

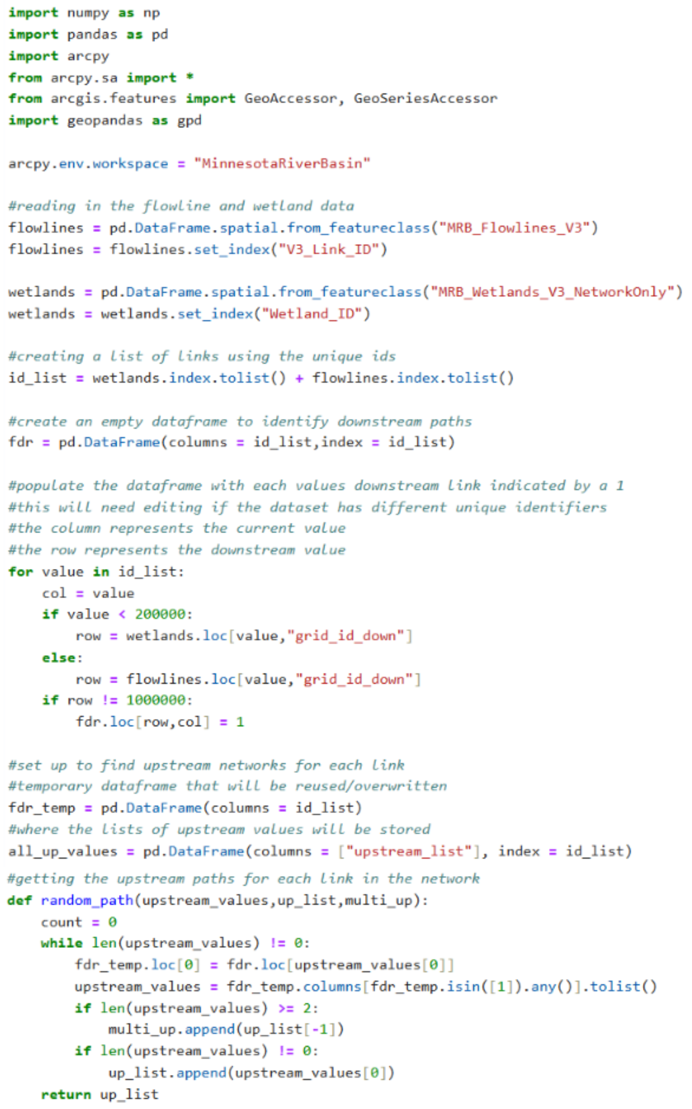
Join Field
Input Table: “wetland_outfalls”
Input Field: unique wetland identifier field
Join Table: “wetland_watershed_poly”
Join Field: gridcode
Transfer Method: Select transfer fields
Transfer Fields: wetland_DA_SQKM

Summarize Within
Input Polygons: site watersheds merged into one shapefile
Input Summary Features: “wetland_outfalls”
Output Feature Class: “interception_fraction”
Uncheck keep all input polygons

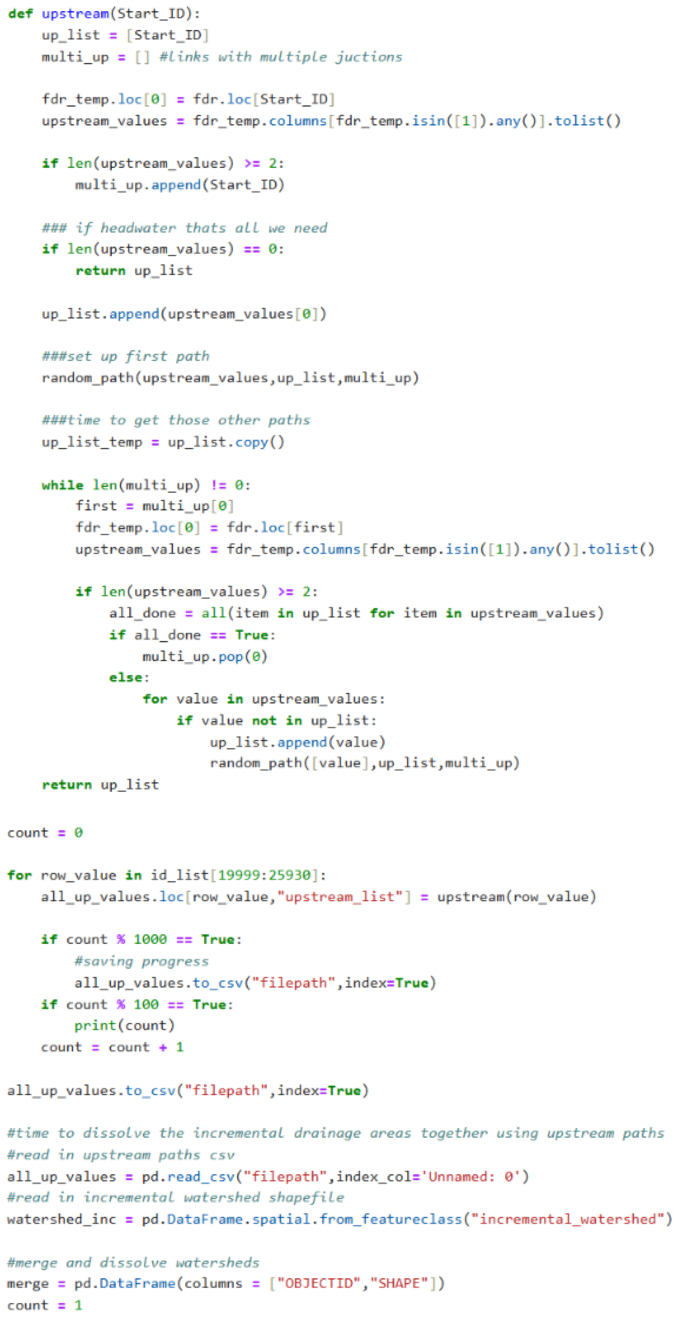
Summary Fields: wetland_DA_SQKM
Statistic: Sum

Calculate Field
Input Table: “interception_fraction”
Field Name (Existing or New): IF
Field Type: Double
Expression Type: Python
Expression: SUM wetland_DA_SQKM / Site_Watershed_Area

Appendix C. Code used to determine total drainage areas for each link of the Minnesota River Basin



```python
import numpy as np
import pandas as pd
import arcpy
from arcpy.sa import *
from arcgis.features import GeoAccessor, GeoSeriesAccessor
import geopandas as gpd

arcpy.env.workspace = "MinnesotaRiverBasin"

#reading in the flowline and wetland data
flowlines = pd.DataFrame.spatial.from_featureclass("MRB_Flowlines_V3")
flowlines = flowlines.set_index("V3_Link_ID")

wetlands = pd.DataFrame.spatial.from_featureclass("MRB_Wetlands_V3_NetworkOnly")
wetlands = wetlands.set_index("Wetland_ID")

#creating a list of links using the unique ids
id_list = wetlands.index.tolist() + flowlines.index.tolist()

#create an empty dataframe to identify downstream paths
fdr = pd.DataFrame(columns = id_list,index = id_list)

#populate the dataframe with each values downstream link indicated by a 1
#this will need editing if the dataset has different unique identifiers
#the column represents the current value
#the row represents the downstream value
for value in id_list:
    col = value
    if value < 200000:
        row = wetlands.loc[value,"grid_id_down"]
    else:
        row = flowlines.loc[value,"grid_id_down"]
    if row != 1000000:
        fdr.loc[row,col] = 1

#set up to find upstream networks for each link
#temporary dataframe that will be reused/overwritten
fdr_temp = pd.DataFrame(columns = id_list)
#where the lists of upstream values will be stored
all_up_values = pd.DataFrame(columns = ["upstream_list"], index = id_list)
#getting the upstream paths for each link in the network
def random_path(upstream_values,up_list,multi_up):
    count = 0
    while len(upstream_values) != 0:
        fdr_temp.loc[0] = fdr.loc[upstream_values[0]]
        upstream_values = fdr_temp.columns[fdr_temp.isin([1]).any()].tolist()
        if len(upstream_values) >= 2:
            multi_up.append(up_list[-1])
        if len(upstream_values) != 0:
            up_list.append(upstream_values[0])
    return up_list
```



```
def upstream(Start_ID):
    up_list = [Start_ID]
    multi_up = [] #links with multiple juctions

    fdr_temp.loc[0] = fdr.loc[Start_ID]
    upstream_values = fdr_temp.columns[fdr_temp.isin([1]).any()].tolist()

    if len(upstream_values) >= 2:
        multi_up.append(Start_ID)

    ### if headwater thats all we need
    if len(upstream_values) == 0:
        return up_list

    up_list.append(upstream_values[0])

    ###set up first path
    random_path(upstream_values,up_list,multi_up)

    ###time to get those other paths
    up_list_temp = up_list.copy()

    while len(multi_up) != 0:
        first = multi_up[0]
        fdr_temp.loc[0] = fdr.loc[first]
        upstream_values = fdr_temp.columns[fdr_temp.isin([1]).any()].tolist()

        if len(upstream_values) >= 2:
            all_done = all(item in up_list for item in upstream_values)
            if all_done == True:
                multi_up.pop(0)
            else:
                for value in upstream_values:
                    if value not in up_list:
                        up_list.append(value)
                        random_path([value],up_list,multi_up)
    return up_list

count = 0

for row_value in id_list[19999:25930]:
    all_up_values.loc[row_value,"upstream_list"] = upstream(row_value)

    if count % 1000 == True:
        #saving progress
        all_up_values.to_csv("filepath",index=True)
    if count % 100 == True:
        print(count)
    count = count + 1

all_up_values.to_csv("filepath",index=True)

#time to dissolve the incremental drainage areas together using upstream paths
#read in upstream paths csv
all_up_values = pd.read_csv("filepath",index_col='Unnamed: 0')
#read in incremental watershed shapefile
watershed_inc = pd.DataFrame.spatial.from_featureclass("incremental_watershed")

#merge and dissolve watersheds
merge = pd.DataFrame(columns = ["OBJECTID","SHAPE"])
count = 1
```

```python
for value in id_list[:1]:
    read_current = all_up_values.loc[value,"upstream_list"][1:-1].split(", ")
    current = [float(element) for element in read_current]
    water_temp = pd.DataFrame(columns = ["gridecode","SHAPE"])
    for i in range(len(current)):
        if current[i] >= 200000:
            water_id = flowlines.loc[current[i],"watershed_ID"]
            water_id_index = watershed_inc.index[watershed_inc["gridcode"]== water_id].to_list()
            if len(water_id_index) != 0:
                water_temp.loc[len(water_temp)] = [watershed_inc.loc[water_id_index[0],"gridcode"],watershed_inc.loc[water_id_index[0],"SHAPE"]]
        else:
            water_id = current[i]
            water_id_index = watershed_inc.index[watershed_inc["gridcode"]== water_id].to_list()
            if len(water_id_index) != 0:
                water_temp.loc[len(water_temp)] = [watershed_inc.loc[water_id_index[0],"gridcode"],watershed_inc.loc[water_id_index[0],"SHAPE"]]
    if value == 327.1 or value == 327.2:
        print(value,"-","Trouble")
    else:
        arcpy.management.Dissolve(water_temp, "water_temp_dissolve")
        dissolve_temp = pd.DataFrame.spatial.from_featureclass("water_temp_dissolve")
        dissolve_temp.loc[0,"OBJECTID"] = value
        merge.loc[len(merge)] = [dissolve_temp.loc[0,"OBJECTID"],dissolve_temp.loc[0,"SHAPE"]]

    if count % 100 == 0:
        merge.spatial.to_featureclass("filepath")
    if count % 100 == 0:
        print(count)
    count = count + 1
merge.spatial.to_featureclass("filepath")
```

Figure A6. Code needed to delineate watersheds for each link in the network.