Processing USGS NHDPlus HR Files for Compatibility with the Landlab NetworkModelGrid and NetworkSedimentTransporter Components (Version 2)

Landlab is a Python-based modeling environment that allows geoscientists to simulate interactions between a variety of processes on Earth's surface. Since its debut in the early 2010s, Landlab's use in peer-reviewed studies has gradually increased, as it has progressively become the community standard software for creating landscape evolution models. Although Landlab can simulate the evolution of real-world terrains derived from raster and shapefile data, it has mostly been used in theoretical studies that use artificial grids to evaluate conceptual landscape dynamics. However, it has the potential to become an extremely valuable tool for applied sciences, as it has many applications for river and sediment management, slope stability, flood-risk assessment, etc.

One potential application of the Landlab toolkit is to model the past, present, or future behavior of real-world rivers. The USGS NHDPlus HR is a seamless geospatial model of US streams and waterways built at a 1:24,000 scale or better. Given its high spatial resolution, seamless topology, and free accessibility online, the USGS NHDPlus HR is an ideal dataset for Landlab to simulate the dynamics of US rivers. However, the raw NHDPlus HR products must be preprocessed before Landlab can properly import and implement them as model grids.

This procedure describes how to prepare files from the USGS NHDPlus HR dataset for compatibility with the Landlab NetworkModelGrid and SedimentNetworkTransporter components. The instructions below explain how to download files from the NHDPlus HR, process them using GIS, and import / implement them in Landlab. This version of the instructions contains several manual steps with GIS; future versions will aim to automate the procedure.

Requirements

These instructions describe how to prepare files from the USGS NHDPlus HR datasets for compatibility with the Landlab NetworkModelGrid and SedimentNetworkTransporter components. The procedure requires:

- QGIS (Freely available online. Click here and follow the directions to download the most recent version. The example in these instructions uses QGIS 3.20 Odense)
- The following QGIS plugins: "Check for disconnected islands" and "Locate points along lines" (Freely available online, see Section 2)
- A polygon shapefile delineating the basin you want to import into Landlab (The example in these instructions uses the "Soque_River.shp" file in the "Associated_Files" folder)
- Jupyter Notebooks (Contained in the freely available Anaconda Individual Edition. Click here to download)
- The Landlab python toolkit (Freely available online. Click here and follow the directions to download and install the toolkit.)
- The "network_sediment_transporter_NHDPlus_HR_network.ipynb" file in the "Associated_Files" folder

---

The Major Steps

1. Download NHDPlus HR Data
2. Install QGIS Plugins

3. Convert Shapefile Format
4. Join Geodatabase Fields to Shapefile
5. Create Temporary Node Shapefile
6. Clip the Basin of Interest
7. Find and Remove Disconnected Links
8. Find and Fix Multiple Downstream Links
9. Create The Node Shapefile
10. Add a Pour Point Node
11. Export the Final Shapefiles
12. Import Shapefiles into Landlab

Example Scenario

This example describes how to import a section of the Soque River (a tributary of the Chattahoochee River in northern Georgia) into Landlab and model its sediment flux.

In the "NHDPlus_HR_Adapter" folder, go to "Associated_Files/Soque_Basin_Polygon/". The files in this folder can generate an outline of the basin of interest. Open Google Earth, drag and drop the "Soque_Basin.shp" file onto the screen, and click "no" when asked "Do you want to apply a style template to the features you ingested?". An outline of the basin should appear in northern Georgia (Figure 1). Pan around to get a feel for the basin's terrain, land cover, and stream topology—having a good sense of a basin's characteristics can help you recognize and solve any unforeseen problems or errors.

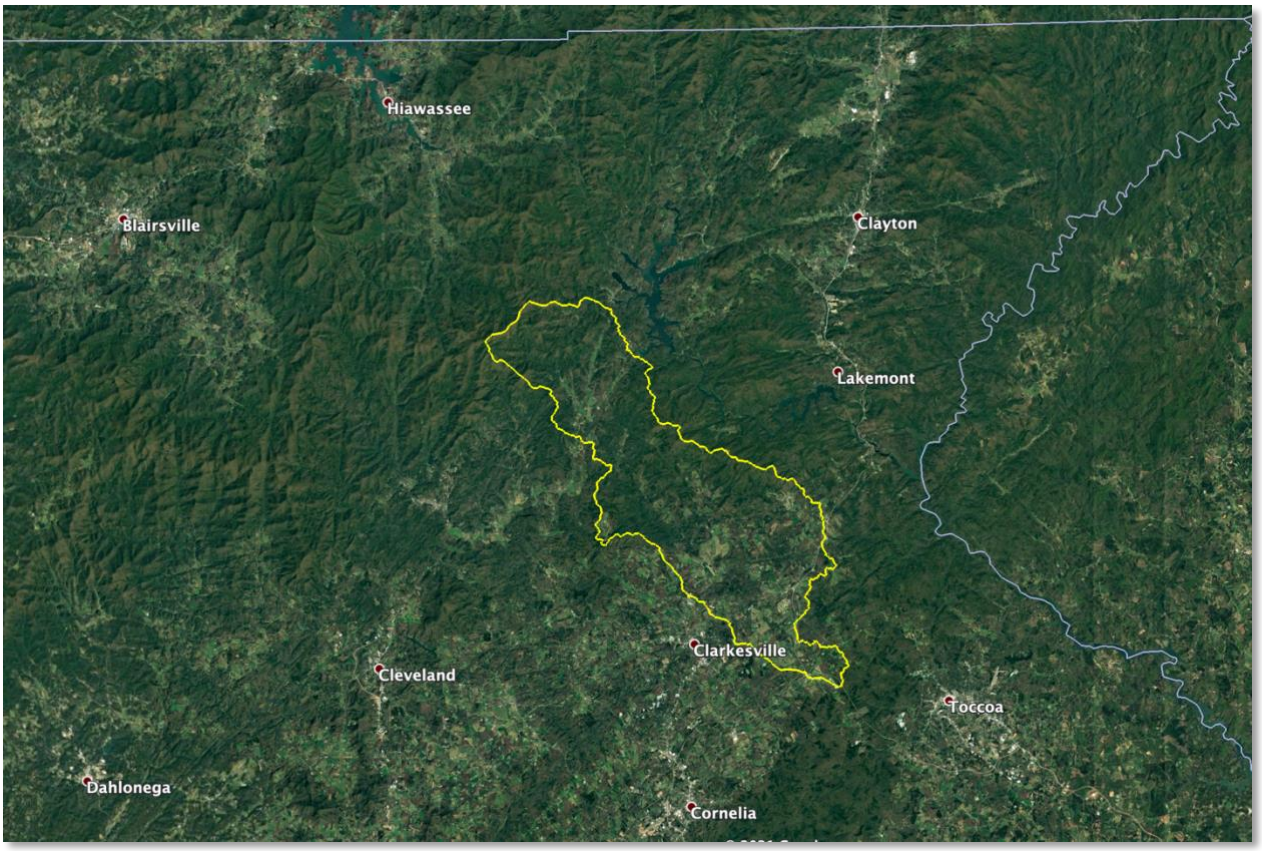

Figure 1: View of the Soque River basin outline (yellow) in Google Earth. The outline is generated by the "Soque_Basin.shp" file.

Using the "Workflow" folder

File organization will be extremely important as you follow these instructions. You will create several, preliminary shapefiles before the final products are generated in Section 11. The philosophy applied here is to save and preserve each of these preliminary files. This is beneficial for two reasons: A) It creates a record of all the processing steps you applied, and B) If at any point you make a mistake, you can return to a previous section of this procedure and reload the files that were being used at that time.

Inside the "NHDFlowline_HR_Adapter" folder, there is another folder called "Workflow". The "Workflow" folder contains several subfolders; you can use these to save shapefiles generated in each step (they also contain compressed versions of the example shapefiles).

---

1. Download NHDPlus HR Data

The first step is to download NHDPlus HR data for the larger HU4 basin containing the Soque River (read more about Hydrologic Units (HU) on the USGS Watershed Boundary Dataset webpage). The NHDPlus HR datasets are part of a collection of National Hydrography products available on the USGS National Hydrography website.

A. Go to the USGS National Hydrography website (https://www.usgs.gov/core-science-systems/ngp/national-hydrography). To download NHDPlus HR data, click on the Access National Hydrography Products tab, scroll down to the section titled "NHDPlus High Resolution (NHDPlus HR) Direct Download Links", and click "Download the NHDPlus HR by 4-digit Hydrologic Unit (HU4) for CONUS or by 8-digit Hydrologic Unit (HU8) for Alaska".

B. The Soque River is located within a large basin with the HU4 code 0313. Scroll down and click on the "NHDPlus_H_0313_HU4_GDB.zip" link to download a compressed folder containing the HU4 file database. Once downloaded, unzip the database folder somewhere on your computer.

---

2. Install QGIS Plugins

Before using QGIS, you will need to install two plugins.

A. Open QGIS and start a new project. Go to Plugins → Manage and Install Plugins.

B. Install the "Disconnected Islands" plugin. In the Manage and Install Plugins window, click the "All" tab and scroll down until you find the "Disconnected Islands" plugin. Then, click install (Figure 2).

C. Repeat the previous step for the "Locate points along lines" plugin.

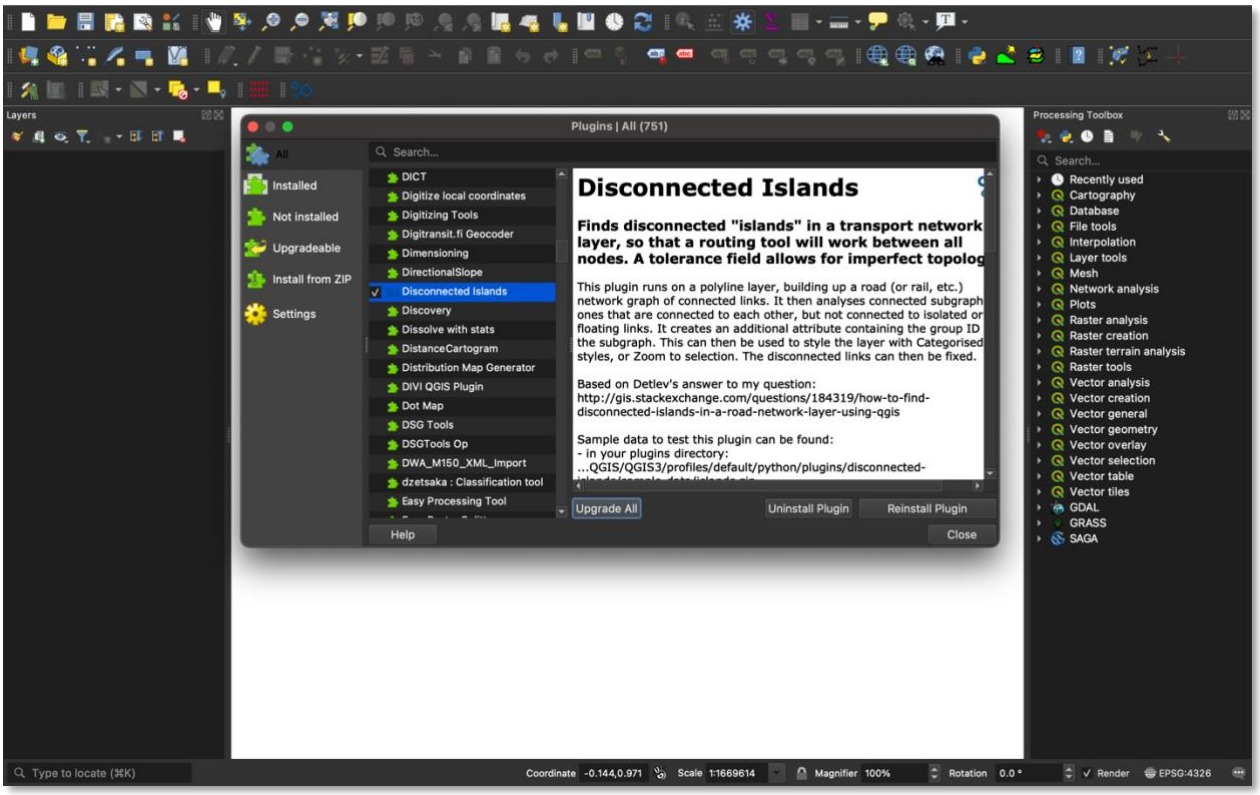

Figure 2: The Manage and Install Plugins window open in QGIS. You can use the search bar at the top of the window to find and install a desired plugin.

---

3. Convert Shapefile Format

The Landlab NetworkModelGrid component requires two shapefile inputs: a polyline structure of links and a multi-point structure of nodes. You will use files from the HU4 database folder to generate the link structure, and later, you will use the link structure to generate the node structure. Here, you will generate the link structure.

A. Go to the "NHDPLUS_H_0313_HU4_GDB.gdb" folder, and you should see several hundred files. Select all the files and drag-drop them into the QGIS workspace. If prompted with a "Select Transformation" dialogue box, click "OK".

B. Many layers will appear in the workspace. There will also likely be an "Invalid Data Source" warning displayed (Figure 3); disregard and close this warning.

C. In the "Layers" panel, find the polyline layer called "a00000058 NHDFlowline" (the first part may vary depending on which HU4 basin you are using, but the second part should still be "NHDFlowline"). This is the layer that you will use to create the link structure.

D. Right-click this layer in the "Layers" panel and go to Export → Save Features As.... In the "Save Vector Layer as..." window, click the "..." button to the right of "File name". In the "Save Layer As" window, navigate to the "Workflow" folder and select the "1_Converted_NHDFlowline" folder. For "Save As", type "1_Converted_NHDFlowline". Click "Save" to close the "Save Layer As" window.

E. For “Geometry type” in the “Save Vector Layer as...” window, select “Linestring” (The default setting is probably “Automatic”, so it is very important that you change this), make sure that the “Force multi-type” and “Include z-dimension” boxes are unchecked (Again, their default setting may be checked, so it is very important that you uncheck them), and click “OK”.

F. A new layer called “1_Converted_NHDFlowline” will appear in the workspace.

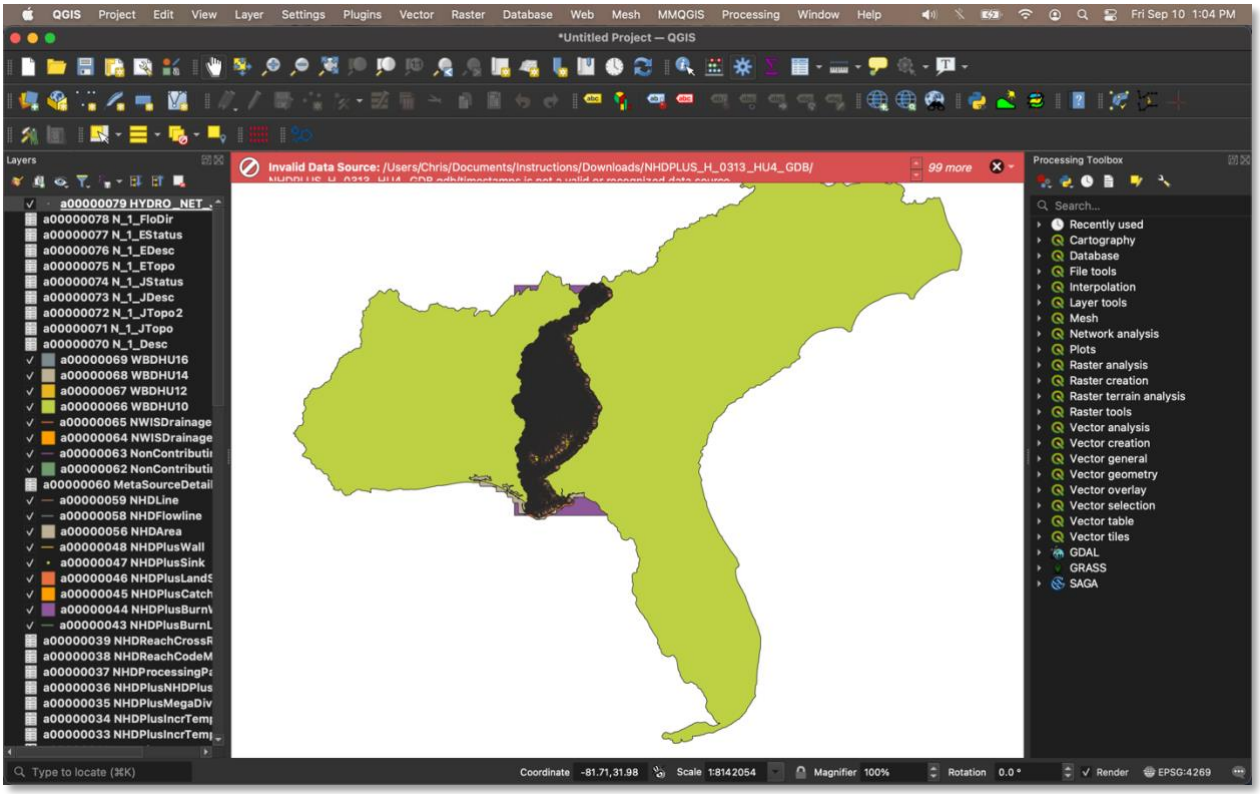

Figure 3: Files from the “NHDPLUS_H_0313_HU4_GDB.gdb” folder displayed in the QGIS workspace. Note the warning (red bar) at the top of the workspace; this can be disregarded and closed.

---

4. Join Geodatabase Fields to Shapefile

The next step is to join information from one of the data tables to the shapefile. Importantly, you want the shapefile to contain information concerning topographic elevation, slope, and drainage area. These are contained in the data table named “a00000018 NHDFlowlineVAA” (the first part may vary depending on which HU4 basin you are using, but the second part should always be “NHDFlowlineVAA”).

A. In the “Processing Toolbox”, search for and open the “Join Attributes by Field Value” window.

B. For “Input layer”, select “1_Converted_NHDFlowline”. For “Table field”, select “NHDPlusID”. For “Input layer 2”, select “a00000018 NHDFlowlineVAA”. For “Table field 2”, select “NHDPlusID”.

C. Click the “...” button to the right of “Joined layer [optional]” and click “Save to File”. In the “Save File” window, navigate to the “Workflow” folder and select the “2_Join_attributes_by_field_value” folder. For “Save As”, type “2_Join_attributes_by_field_value”. Click “Save” to close the “Save Layer As” window.

D. Back in the “Join Attributes by Field Value” window, click “Run”. A new layer called “2_Join_attributes_by_field_value” should appear in your workspace. Remove all layers from the workspace except for this one (Figure 4).

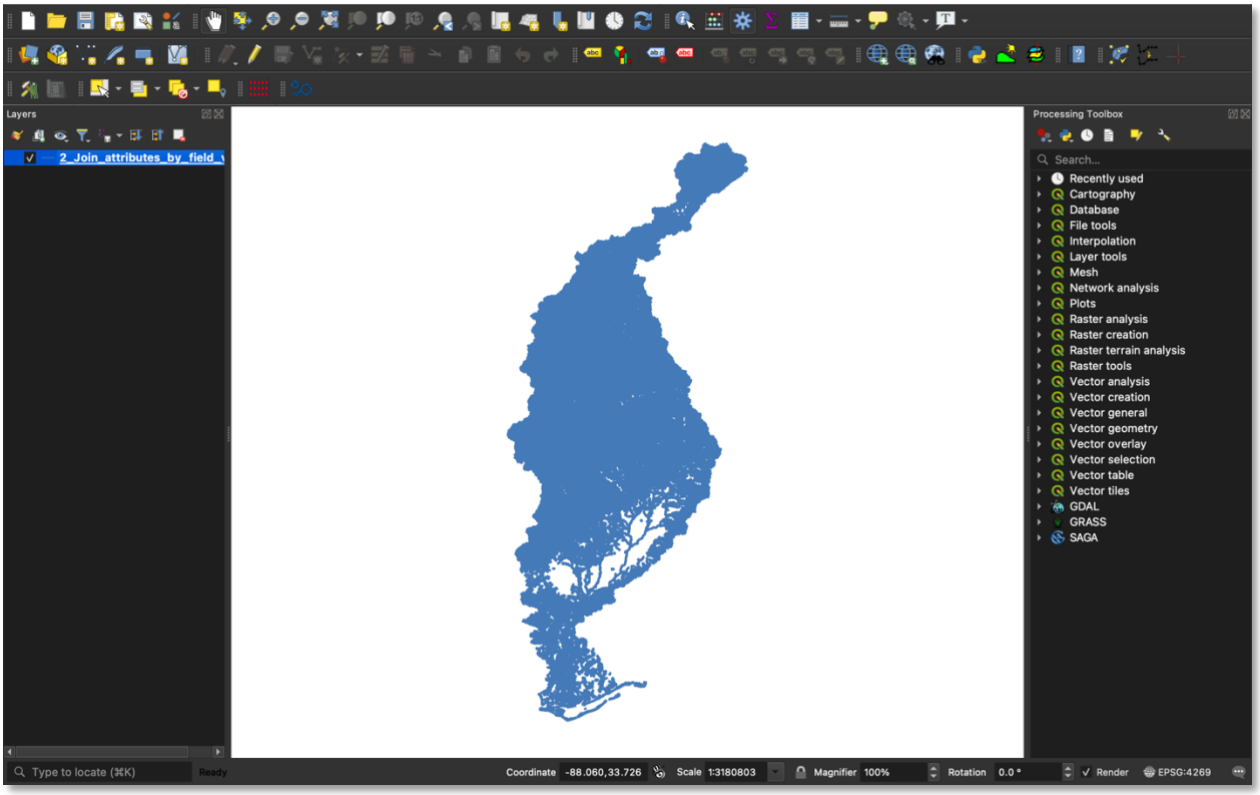

Figure 4: The “2_Join_attributes_by_field_value” layer in the workspace. The layer color may be different than this example (in this step and all the following steps).

---

5. Create Temporary Node Shapefile

Now, use the “2_Join_attributes_by_field_value.shp” file to create a temporary node structure. You will only use these nodes for visualization purposes; the final node structure will be generated later.

A. Open the “Locate points along lines” plugin (An icon should have appeared in one of the top panels when you installed it; it looks like a 5 x 3 grid of red dots).

B. For “Input polyline layer”, select “2_Join_attributes_by_field_value”. For “Output layer name”, type “Temporary_Nodes”. Check the “Keep attributes” box. Then, click “Run”. A layer called “Temporary_Nodes” should appear in the workspace (Figure 5).

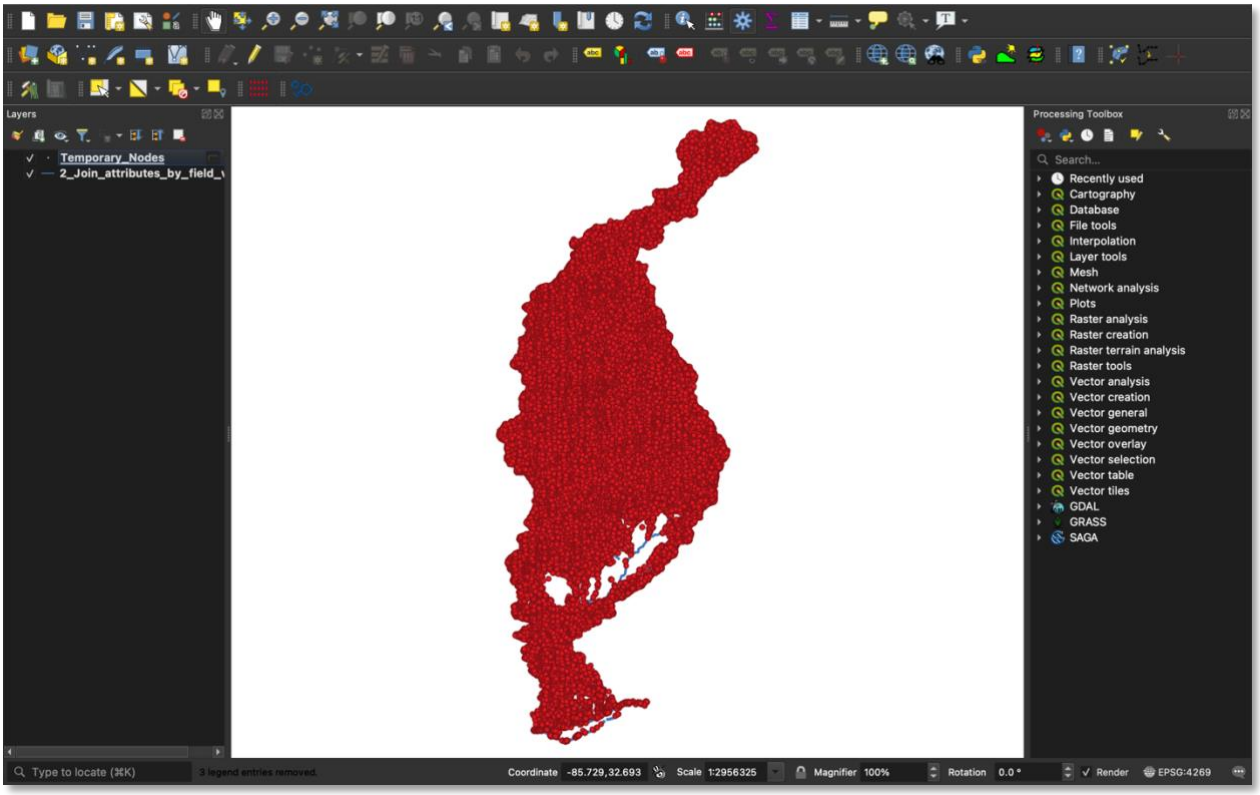

Figure 5: The “Temporary_Nodes” layer in the wokspace.

---

6. Clip the Basin of Interest

The “2_Join_attributes_by_field_value” layer covers a much larger area than the basin of interest. You need to clip the layer down to the Soque River basin. Note that if you wanted to import an entire HU4 basin into Landlab, you could skip this step.

A. Drag-drop the “Soque_Basin.shp” file into the workspace. A polygon of the basin should appear in the workspace.

B. Go to Vector → Geoprocessing Tools → Clip... to open the “Clip” window. For “Input layer”, select “2_Join_attributes_by_field_value”. For “Overlay layer”, select “Soque_Basin”.

G. Click the “...” button to the right of “Clipped” and click “Save to File”. In the “Save File” window, navigate to the “Workflow” folder and select the “3_Clipped” folder. For “Save As”, type “3_Clipped”. Click “Save” to close the “Save Layer As” window.

C. Back in the “3_Clip” window, click “Run”. A new layer called “3_Clip” should appear in your workspace. Remove the “2_Join_attributes_by_field_value” and “Soque_Basin” layers

from the workspace, right-click on the “3_Clip” layer in the layers panel and select “Zoom to Layer(s)” (Figure 6).

D. You do not want any extra polylines downstream of the outlet node. Zoom in to the outlet of the basin. At a scale of ~250:1, you can see that when you clipped the streams, a small segment just downstream of the pour point node was not deleted (Figure 7).

E. Left click on “3_Clip” in the “Layers” panel so that it is highlighted. Then, click on the “Select Features by Area or Single Click” button (it looks like a white curser clicking on a yellow rectangle inside a larger grey rectangle). Next, click on the “Toggle Editing” button (it looks like a single yellow pencil). Click on the unwanted stream segment, and it will be highlighted yellow with red crosshairs at its vertices. Press “Delete” on your keyboard to delete this segment.

F. Click the “Toggle editing” button again, and when promoted whether “you want to save the changes to layer Converted_NHDFlowline”, click “Save”.

G. Remove the “Temporary_Nodes” layer from the workspace and click the “Zoom Fill” button (it looks like a magnifying glass with three arrows pointing outward from the center) (Figure 8).

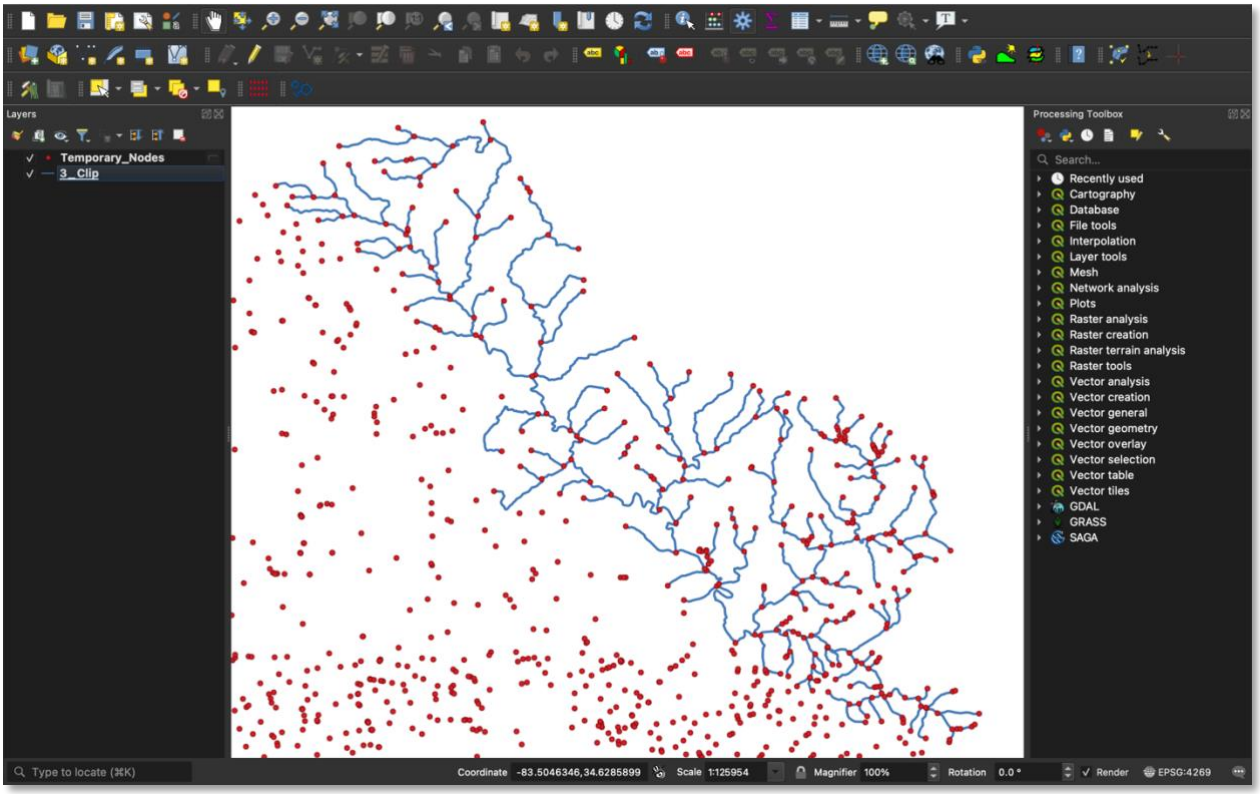

Figure 6: The “3_Clip” and “Temporary_Nodes” layers in the workspace. Note how the “Temporary_Nodes” layer covers the former extent of the “2_Join_attributes_by_field_value” layer.

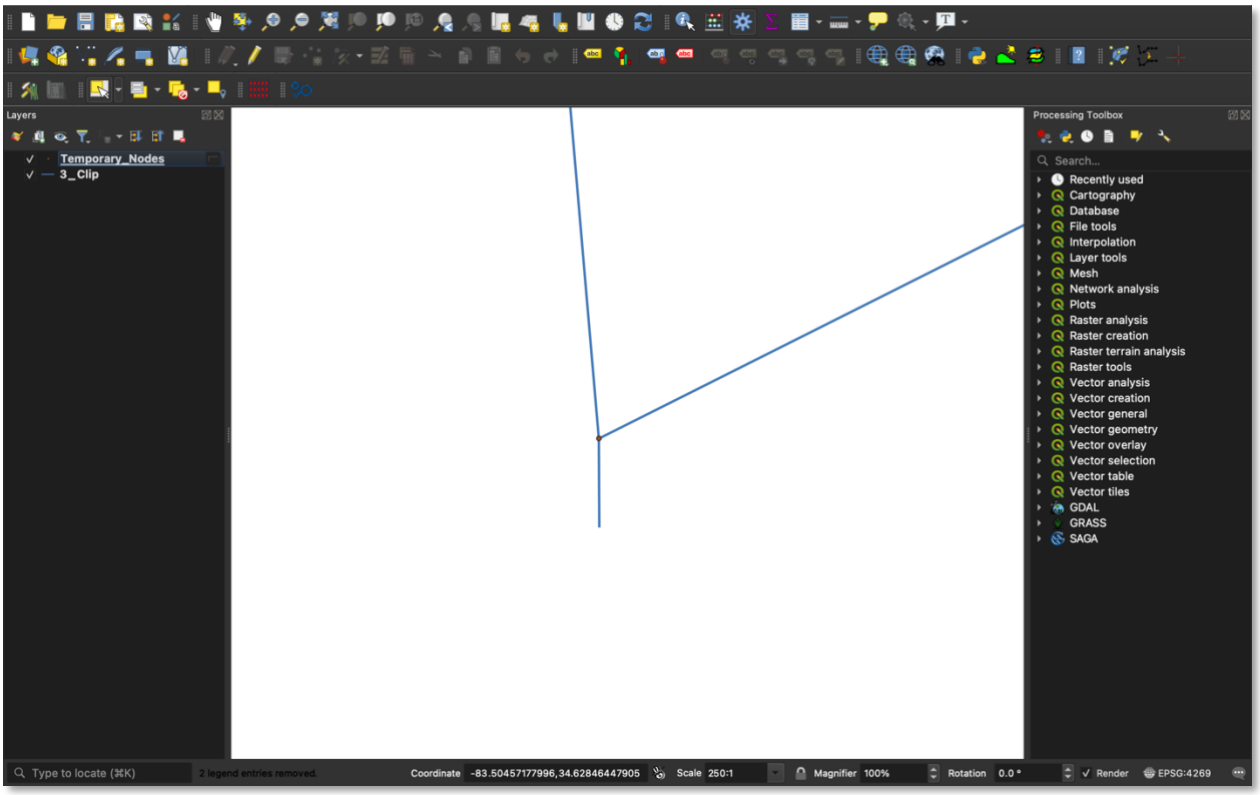

Figure 7: View of the "3_Clip" layer zoomed in to the basin outlet at a scale of 250:1 (the scale is displayed at the bottom center of the figure, just below the workspace). The small segment extending south of the node needs to be deleted.

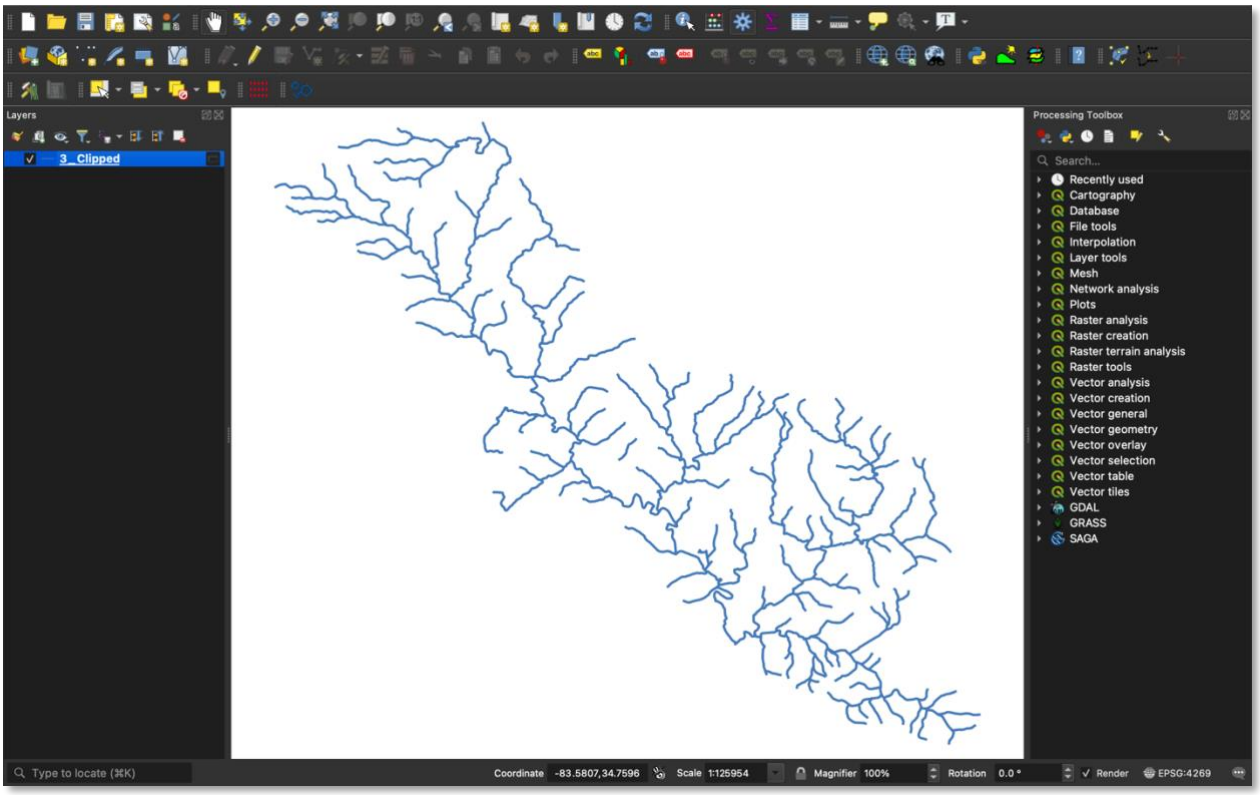

Figure 8: Full extent of the "3_Clip" layer after Step 6.

---

7. Find and Remove Disconnected Links

The "3_Clip.shp" file is now in a format that can be handled by Landlab. However, there are some issues with the network topology that need to addressed before you can model sediment transport. The first issue is the connectivity of the stream network. The Landlab NetworkModelGrid component requires that the links comprise one continuous network, meaning that there cannot be any disconnected links. To find and remove disconnected links follow these steps:

A. First, you need to create a copy of the "3_Clip" layer that you can edit. Right-click "3_Clip" in the "Layers" panel and go to Export → Save Features As.... In the "Save Vector Layer as..." window, click the "..." button to the right of "File name". In the "Save Layer As" window, navigate to the "Workflow" folder and select the "4_Deleted_Islands" folder. For "Save As", type "4_Deleted_Islands". Click "Save" to close the "Save Layer As" window. Back in the "Save Vector Layer as..." window, click "OK".

B. A new layer called "4_Deleted_Islands" should appear in the workspace. Remove the "3_Clip" layer from the workspace.

C. To check for any disconnected links, open the "Check for Disconnected Islands" plugin (An icon should have appeared in one of the top panels when you installed it- it looks like three blue circles with two of them connected by a line). Under "Select layer", select "4_Deleted_Islands", and click OK.

D. When the plugin is finished running, a new attribute field will be added to the "4_Deleted_Islands" layer, and it should be displayed in the Layers panel (Figure 9). The new field has values ranging from 0 to 1; this means that there is one link segment that is disconnected from the main network (the main network has a value of 0).

E. You need to manually delete the disconnected segments. Uncheck the "0" attribute within the "4_Deleted_Islands" layer. This will leave only the single disconnected island visible.

H. Click the "Toggle Editing" button. Click the "Select Features by Area or Single Click" button, select the disconnected segment, press "delete" on your keyboard to delete the segment, and click the "Toggle Editing" button again. When promoted whether "you want to save the changes to layer 4_Deleted_Islands", click "Save".

I. To verify that the disconnected segment was successfully deleted, open the "Check for Disconnected Islands" plugin again. Under "Select layer", select "4_Deleted_Islands". Check the "Overwrite existing values" box and click OK. The layer should now only have a single attribute named "0" (Figure 10).

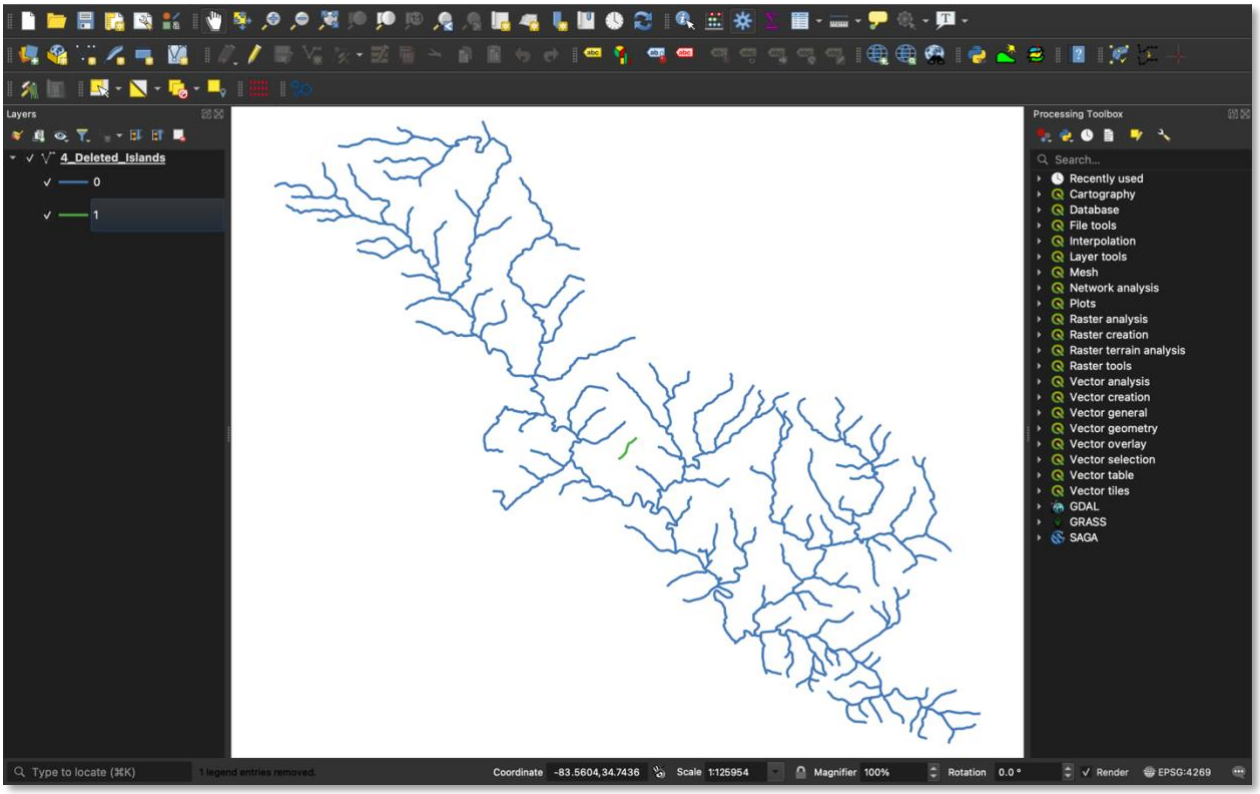

Figure 9: View of the "4_Deleted_Islands" layer after running the "Check for Disconnected Islands" plugin. There is a single disconnected link in the layer (near the center of the workspace, colored green).

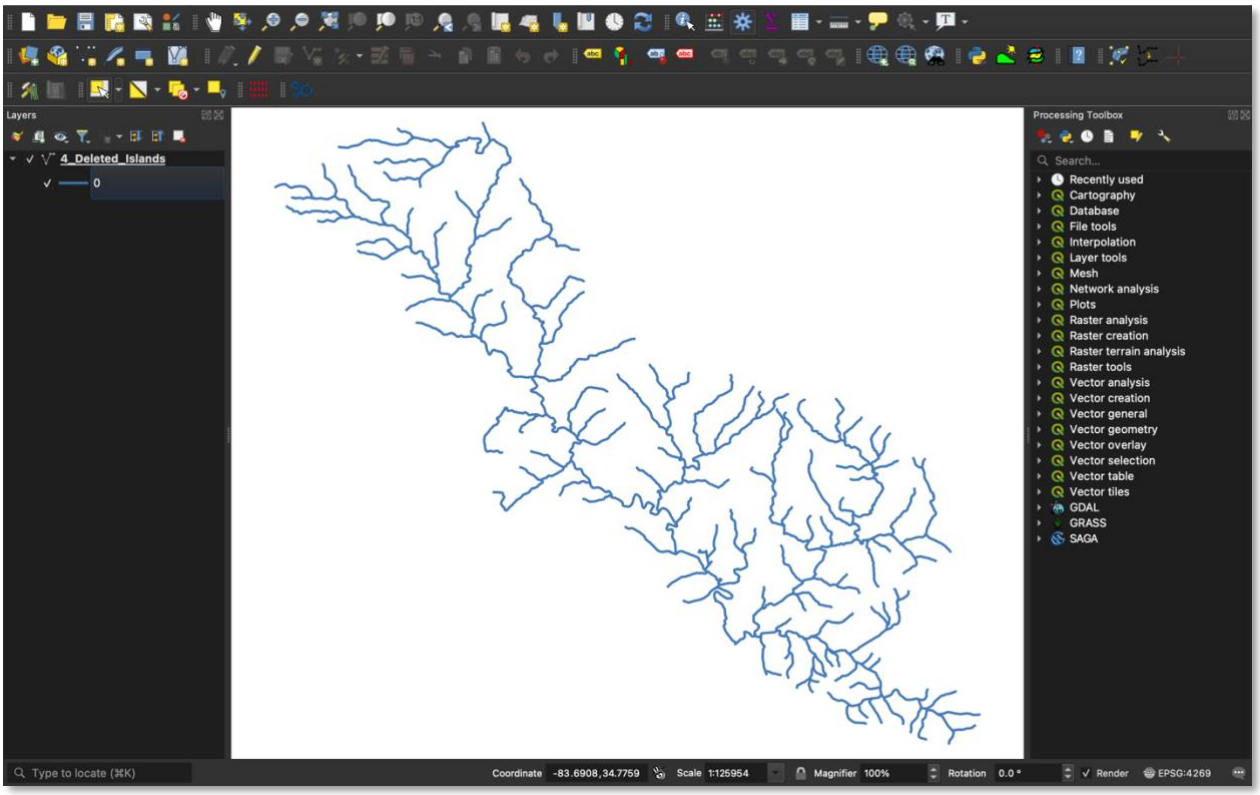

Figure 10: View of the "4_Deleted_Islands" layer after deleting the disconnected link.

---

8. Find and Fix Multiple Downstream Links

The next issue involves the number of links associated with each node. The Landlab NetworkModelGrid component requires that each node have only one link connected to its downstream end. However, the USGS NHDFlowline files sometimes have two links downstream of a node to represent anabraching channels or large islands. Within the clipped region of the Soque River, this does not occur anywhere. However, we will discuss how to resolve the issue in case it occurs in your own basin of interest.

There are several possible ways that you could resolve this issue, and whichever one you choose depends on how you wish to think about the physical system. For many purposes, the best solution is to delete the smaller of the two channels at each location, and so that is what you will do in this example.

Fortunately, there is a simple way to do this. Wherever two stream segments branch off from the same node, the "Divergence" attribute of each segment will be one of two values: 1 (the larger branch) or 2 (the smaller branch). To solve the issue, you can delete all line features that have a "Divergence" value of 2.

A. First, you need to create a copy of the "4_Deleted_Islands" layer that you can edit. Right-click "4_Deleted_Islands" in the "Layers" panel and go to Export → Save Features As.... In the "Save Vector Layer as..." window, click the "..." button to the right of "File name". In the "Save Layer As" window, navigate to the "Workflow" folder and select the "5_Deleted_Divergence" folder. For "Save As", type "5_Deleted_Divergence". Click "Save" to close the "Save Layer As" window. Back in the "Save Vector Layer as..." window, click "OK".

B. A new layer called “5_Deleted_Divergence” should appear in the workspace. Remove the “4_Deleted_Islands” layer from the workspace.

C. Go to Edit → Select → Select Features by Value.... In the “5_Deleted_Divergence – Select Features” window, enter “2” for “Divergence”, and click “Select Features” (Since this layer does not have any multiple downstream links, no features will be selected, but otherwise they would be!).

D. Click the “Toggle Editing” button, press “Delete” on your keyboard, and click the “Toggle Editing” button again. When promoted whether “you want to save the changes to layer Converted_NHDFlowline”, click “Save” (Figure 11).

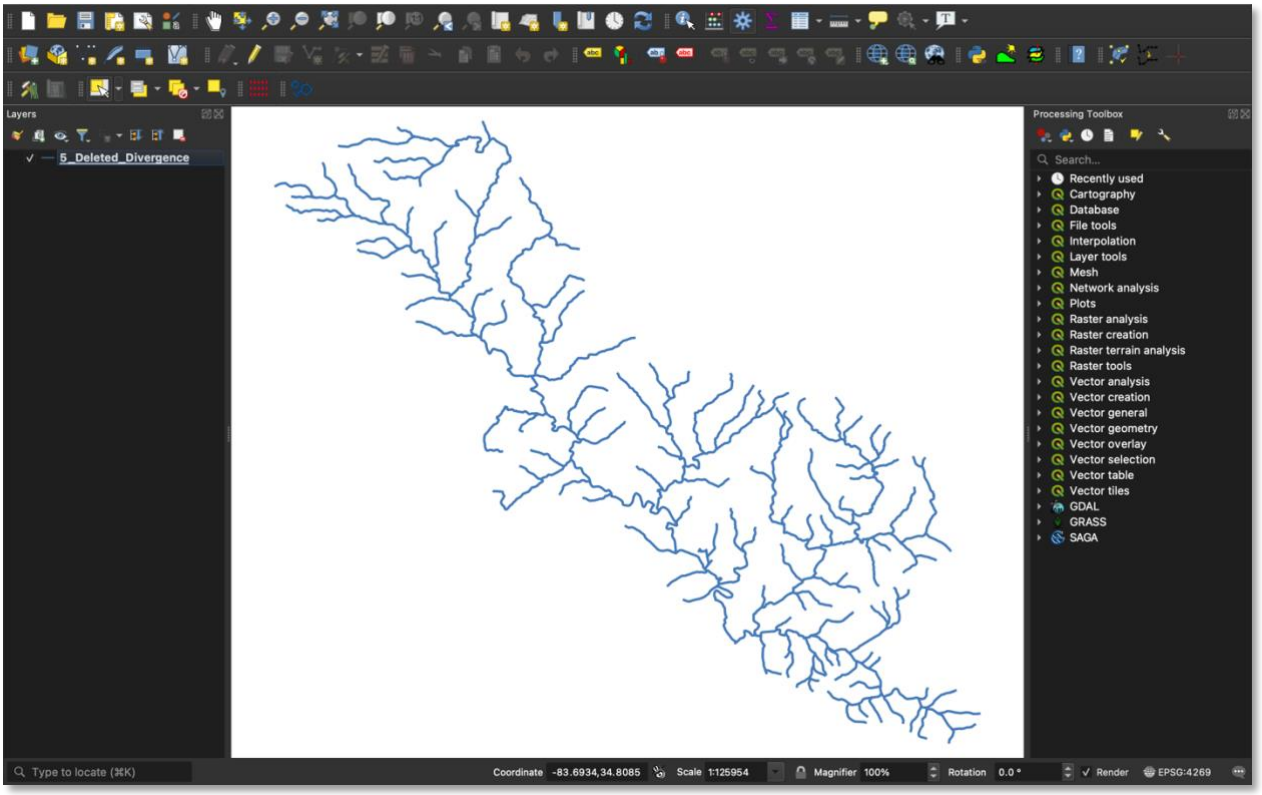

Figure 11: View of the “5_Deleted_Divergence” layer. Note that the layer appears the same as the “4_Deleted_Islands” layer in Figure 10, because there were no divergences in this reach of the Soque River. Divergences are common in the NHDPlus HR though, so if you are following this procedure for a different river, you will likely need to complete this step.

---

9. Create The Node Shapefile

Now, you will use the “5_Deleted_Divergence.shp” file to begin creating a node structure.

A. Open the “Locate points along lines” plugin. For “Input polyline layer”, select “5_Deleted_Divergence”. For “Output layer name”, type “Temporary_Nodes”. Check the “Keep attributes” box. Then, click “Run”. A layer called “Temporary_Nodes” should appear in the workspace.

B. The “Temporary_Nodes” layer is currently a virtual layer, meaning that it has not yet been saved on your machine. Right-click the layer’s name in the “Layers” panel and go to Export → Save feature as.

E. Right-click “Temporary_Nodes” in the “Layers” panel and go to Export → Save Features As.... In the “Save Vector Layer as...” window, click the “...” button to the right of “File name”. In the “Save Layer As” window, navigate to the “Workflow” folder and select the “6_Deleted_Divergence_Nodes” folder. For “Save As”, type “6_Deleted_Divergence_Nodes”. Click “Save” to close the “Save Layer As” window. Back in the “Save Vector Layer as...” window, click “OK”. A new layer called “6_Deleted_Divergence_Nodes” should appear in the workspace.

C. Remove the “Temporary_Nodes” layer from the workspace (Figure 12).

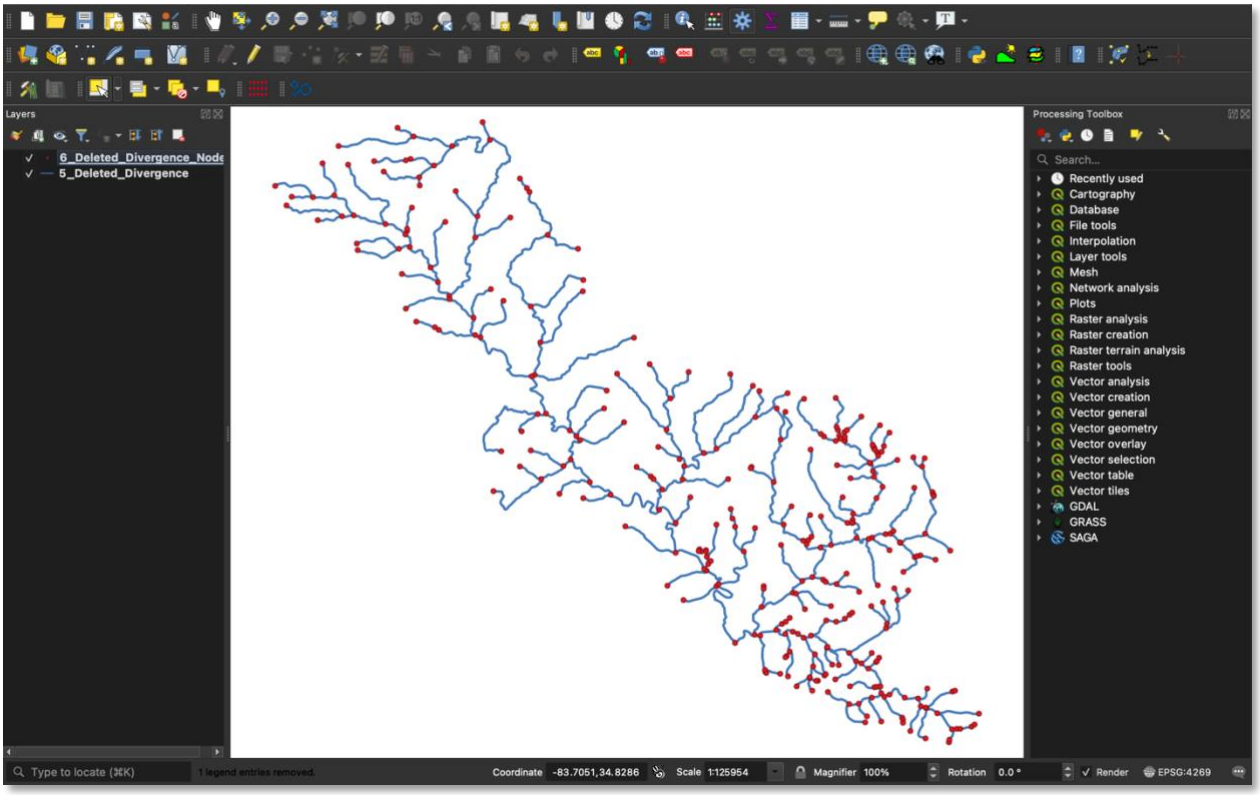

Figure 12: View of the “5_Deleted_Divergence” and “6_Deleted_Divergence_Nodes” layers.

---

10. Add a Pour Point Node

At this time, the “5_Deleted_Divergence” layer should contain 317 links, and the “6_Deleted_Divergence_Nodes” layer should contain 318 nodes (you can verify this by double-clicking on the layer in the “Layers” panel, navigating to the “Information” tab, and checking the value of “Feature count”). Each link has a corresponding node on its upstream end that shares its attribute values.

With one exception: the node at the outlet of the basin (i.e., the pour point node) does not have a corresponding link downstream (hence the number of nodes is equal to the number of links plus one). In order for Landlab to identify this node as the pour point, you must delete all of its attribute values. You can do that by deleting the node and creating a new one.

A. First, you need to create a copy of the "6_Deleted_Divergence_Nodes" layer that you can edit. Right-click "6_Deleted_Divergence_Nodes" in the "Layers" panel and go to Export → Save Features As.... In the "Save Vector Layer as..." window, click the "..." button to the right of "File name". In the "Save Layer As" window, navigate to the "Workflow" folder and select the "7_Create_Outlet_Node" folder. For "Save As", type "7_Create_Outlet_Node". Click "Save" to close the "Save Layer As" window. Back in the "Save Vector Layer as..." window, click "OK".

B. A new layer called "7_Create_Outlet_Node" should appear in the workspace. Remove the "6_Deleted_Divergence_Nodes" layer from the workspace.

C. Zoom in to the outlet of the basin (just as you did in Section 6).

D. Left-click on "7_Create_Outlet_Node" in the "Layers" panel so that it is highlighted. Click on the "Select Features by Area or Single Click" button. Click on the "Toggle Editing" button. Click on the outlet node and press "Delete" on your keyboard. The node will be deleted.

E. Click the "Add Point Feature" button (it looks three green dots and a yellow rectangle containing a gear). Then, zoom in as far as you can to the intersection of the two downstream-most line segments (Figure 13).

F. Left-click once at the intersection of the two lines. This will open the "7_Create_Outlet_Node – Feature Attributes" window. You want the pour point node to have no attribute values so that Landlab can distinguish it, so click "Ok" without entering any additional values.

G. A new point feature has now been created, but it might not render on your screen right away. If you can't see it, try clicking the "Zoom Full" button and then zooming back into the outlet (Figure 14).

D. Click the "Toggle Editing" button again. When promoted whether "you want to save the changes to layer 7_Create_Outlet_Node", click "Save".

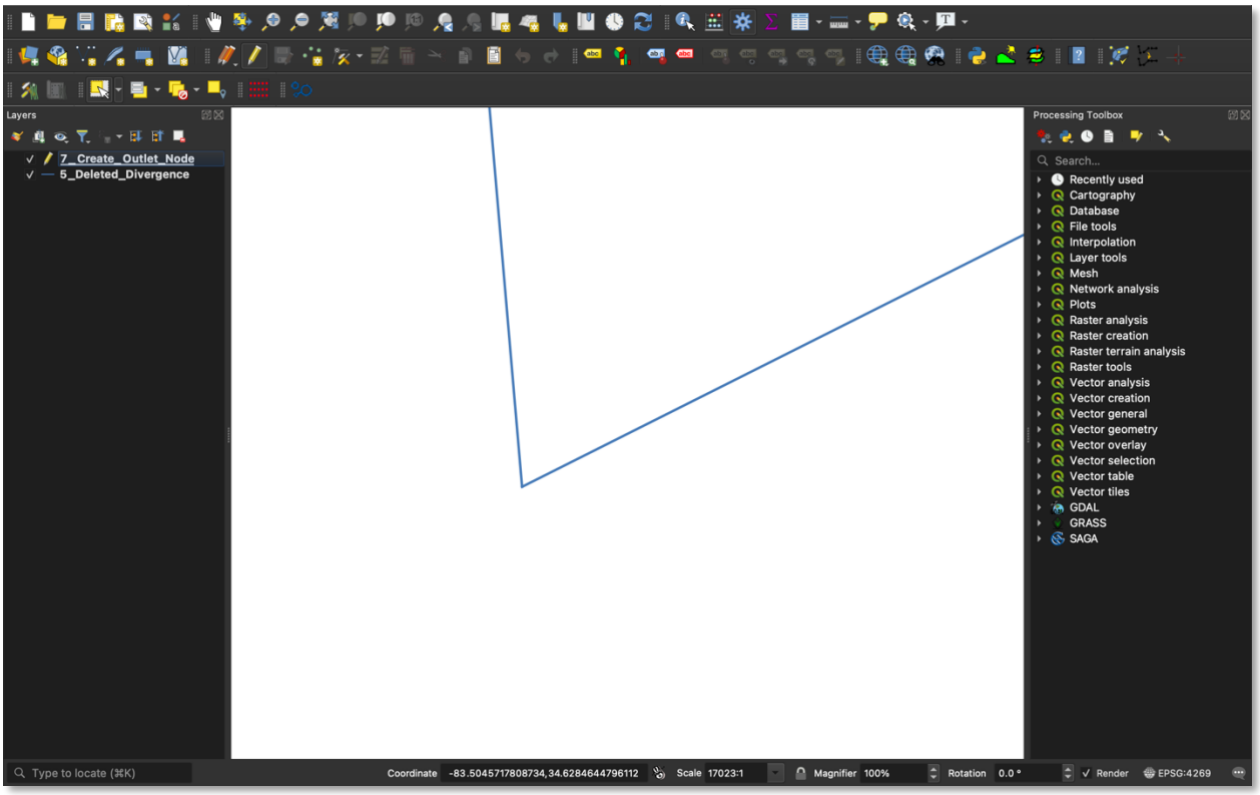

Figure 13: View of the basin outlet zoomed in to the maximum scale (17023:1, this might vary on your own machine).

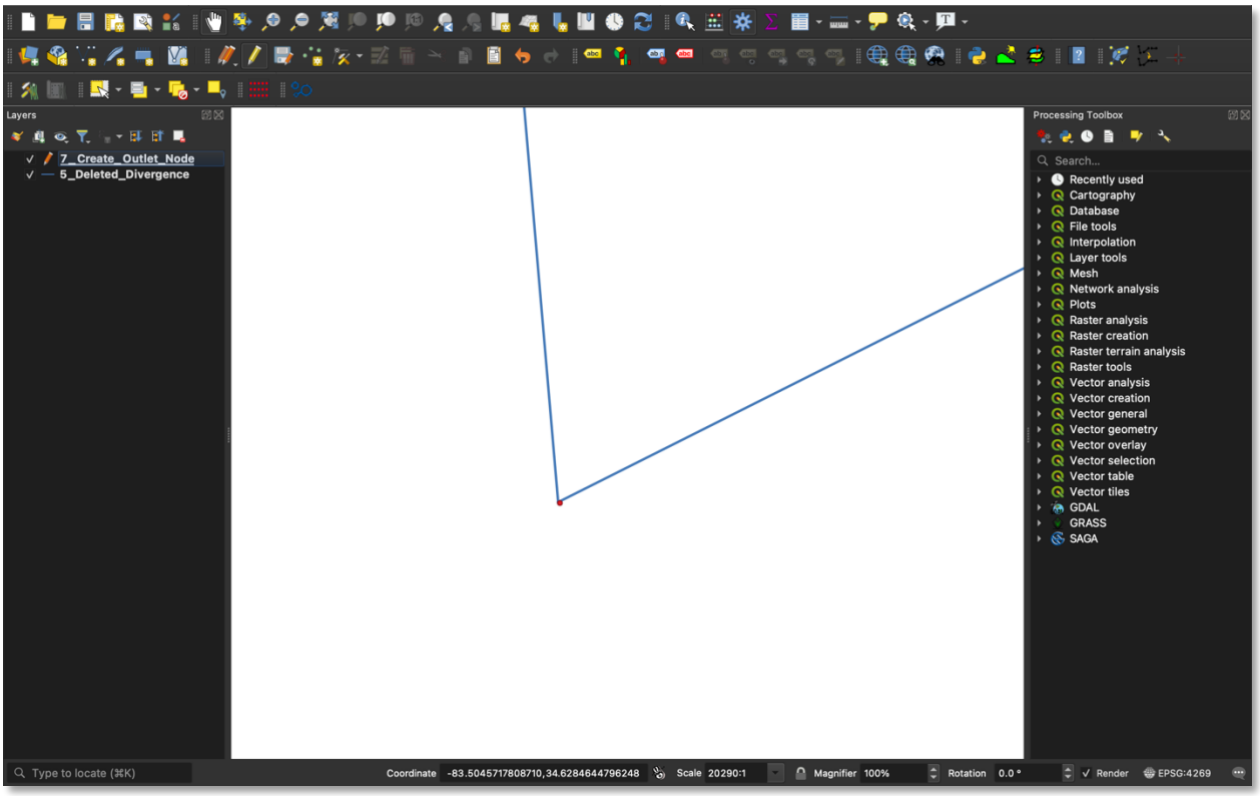

Figure 14: Same view as Figure 13 after adding a new node.

11. Export the Final Shapefiles

The “5_Deleted_Divergence.shp” and “7_Create_Outlet_Node.shp” files are now suitable to be imported into Landlab. However, let’s save them under new names in a new location so that you can distinguish them as the final products.

A. Right-click the “5_Deleted_Divergence” layer in the “Layers” panel and go to Export → Save Features As.... In the “Save Vector Layer as...” window, click the “...” button to the right of “File name”. In the “Save Layer As” window, navigate to the “Workflow” folder and select the “8_Final_Products” folder. For “Save As”, type “8_Final_Links”. Click “Save” to close the “Save Layer As” window, and then click “OK” in the “Save Vector Layer as...” window. A new layer called “8_Final_Products” will appear in the workspace.

B. Repeat the previous step but with the “7_Create_Outlet_Node” layer. Save it to the “8_Final_Products” folder under the name “8_Final_Nodes”.

C. Remove the “5_Deleted_Divergence” and 7_Create_Outlet_Node” layers from the workspace (Figure 15).

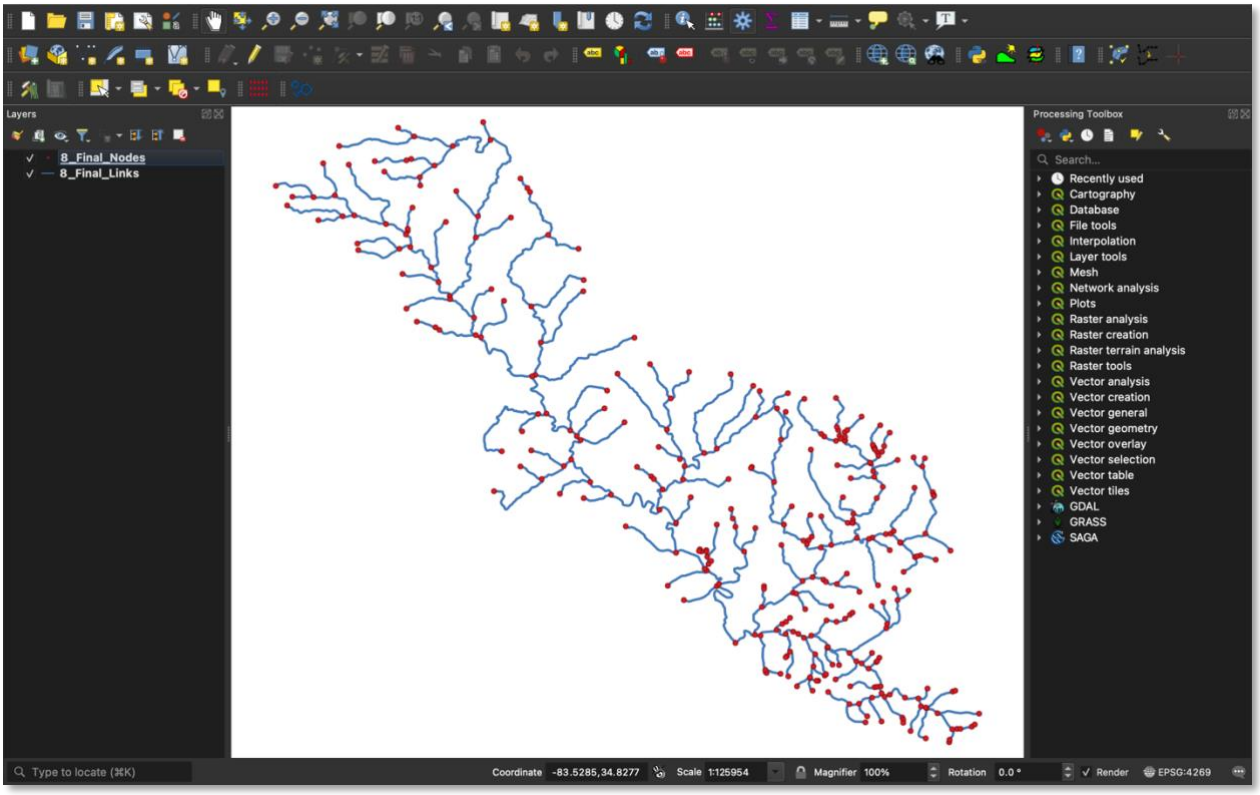

Figure 15: View of the final two shapefiles that will be imported into Landlab.

12. Import Shapefiles Into Landlab

The “8_Final_Links.shp” and “8_Final_Nodes.shp” files are ready to be imported into Landlab and used with the NetworkSedimentTransporter component. To verify that the files are in the correct format and learn how to use them in Landlab, you can load the files into the Jupyter Notebook in the “Associated_Files” folder:

A. Open the “network_sediment_transporter_NHDPlus_HR_network.ipynb” Jupyter Notebook.

B. In the second code block, enter the paths to the “8_Final_Links” and “8_Final_Nodes” files (both the .shp and .dbf files).

C. Run the notebook.