STARS: SPATIAL TOOLS FOR THE ANALYSIS OF RIVER SYSTEMS VERSION 2.0.7 - A TUTORIAL

Dr. Erin E. Peterson
25 March, 2019

ARC Centre for Excellence in Mathematical and Statistical Frontiers (ACEMS) and the
Institute for Future Environments (IFE)
Queensland University of Technology, Gardens Point Campus
Brisbane, QLD, Australia
support@spatialstreamnetworks.com

Table of Contents

List of Figures

List of Tables

1. INTRODUCTION

Spatial autocorrelation is an intrinsic characteristic in freshwater stream environments where nested watersheds and flow connectivity may produce patterns that are not captured by Euclidean distance. Yet, many common autocovariance functions used in spatial models are statistically invalid when Euclidean distance is replaced with hydrologic distance. This issue made it necessary to develop new spatial statistical methodologies for stream networks (i.e. spatial stream-network models), which permit valid covariances to be generated based on a variety of hydrologic relationships (Peterson and Ver Hoef 2010; Ver Hoef and Peterson 2010).

Fitting spatial stream-network models requires multidisciplinary skills in aquatic ecology/biology, geographic information science, and spatial statistics. In addition, specialised geographic information systems (GIS) tools are needed to generate the data necessary to fit spatial stream-network models. The Spatial Tools for the Analysis of River Systems (STARS) custom toolset for ArcGIS has been provided to help users generate these spatial data (Peterson and Ver Hoef 2014). In previous versions, it was necessary to make use of the Functional Linkage of Waterbasins and Streams (FLoWS) toolbox (Theobald et al., 2006) during the pre-processing steps. However, the relevant FLoWS tools have been incorporated into STARS versions ≥ 2.0. The STARS toolbox provides all of the tools needed to generate the specific spatial data needed to fit spatial stream-network models.

This document has been written as a "recipe book" to help users use the STARS tools to generate the data required for spatial stream-network modelling (Peterson and Ver Hoef 2010; Ver Hoef and Peterson, 2010). An example dataset is available on the SSN & STARS website (http://www.fs.fed.us/rm/boise/AWAE/projects/SSN_STARS/software_data.html) and is referred to throughout this tutorial. There are also additional resources that may be useful, such as the FLoWS user manual (Theobald et al., 2006), the ArcGIS help pages, and the ESRI User Forums (http://forums.arcgis.com/). The last two references will not provide specific information about the STARS toolset, but may provide help with more general ArcGIS tasks.

2. SOFTWARE REQUIREMENTS

1. ArcGIS version ≥ 10.6
2. Advanced license and the Spatial Analyst extension
3. STARS version 2.0.7 geoprocessing toolbox for ArcGIS
4. Python version 2.7.14
5. PythonWin: **must be downloaded and installed separately from Python.** Go to this website: http://sourceforge.net/projects/pywin32/files/pywin32/, click on Build 221, and download this file: pywin32-221.win32-py2.7.exe.

3. STARS TOOLSET STRUCTURE

The STARS geoprocessing toolbox is written in Python version 2.7.14 for ArcGIS version 10.6. The STARS toolbox (Figure 1) contains three toolsets: 1) Pre-processing, Calculate, and Export. These tools are specifically designed to analyse, reformat, and export the spatial data as a .ssn ("dot s-s-n") object. The .ssn object can be directly imported to R statistical software (R

Core Team 2019) using the SSN package (Ver Hoef et al. 2014), where spatial stream-network models can be fit to streams data.
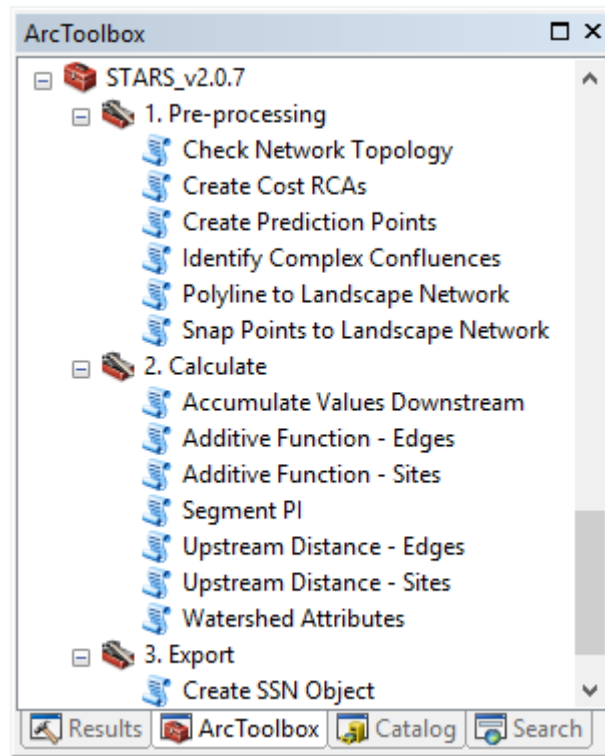
Figure 1. Spatial Tools for the Analysis of River Systems (STARS) geoprocessing toolbox for ArcGIS version 10.6.

4. DATA REQUIREMENTS

At a minimum, two datasets are required to generate the spatial data needed to fit a spatial stream-network model: streams and survey sites (Table 1). Additional data, such as reach contributing areas (RCAs), a digital elevation model (DEM), a waterbodies dataset and ancillary data are also useful if available. Details about how these data are used are provided in the following sections.

Table 1. Spatial data requirements for spatial stream-network modelling.

Data	Format	Description	Required
Survey sites	Vector	Survey data collected at sites	Yes
Reaches	Vector	Stream reaches	Yes
RCAs	Vector or raster	Reach contributing areas for each line segment	Optional
DEM	Raster	Digital elevation model	Optional
Waterbodies	Raster	Lakes, reservoirs, estuaries, etc.	Optional
Ancillary data	Raster	Descriptive data, such as land use, land cover, geology, climate, etc.	Optional

4.1 Example dataset description

An example dataset is provided on the SSN & STARS website (http://www.fs.fed.us/rm/boise/AWAE/projects/SSN_STARS/software_data.html), which is used in this tutorial. The intention is for users to work through the tutorial so that they understand the tools and successfully generate the spatial data necessary for spatial stream-network modelling. The *example* folder contains the input data for the STARS tools, which are used to create new tables, layers, and datasets in this tutorial (Table 2). The second folder, *example_final*, contains the data provided in the example folder, as well as the data derived in this tutorial (Table 3).

Table 2. A description of the spatial datasets contained in the example folder, which are used to derive new spatial data in this tutorial.

Name	Format	Description
sites.shp	Vector	Observed site locations and data
preds.shp	Vector	Prediction site locations
streams.shp	Vector	Stream reaches
waterbodies	Vector	Lakes, reservoirs, estuaries, etc.
dem	Raster	Digital elevation model
grazing	Raster	Grazing land use

Table 3. A description of the datasets contained in the example and example_final folders. These datasets are derived from the example dataset using the STARS tools and the instructions in this tutorial.

Name	Format	Description
dem	Raster	Digital elevation model
demburn5	Raster	Digital elevation model with streams "burned in" 5 metres
demfill	Raster	Filled digital elevation model
edgegrid	Raster	Raster of stream reaches
edgegrid5	Raster	edgegrid reclassed to values 0, 5
grazing	Raster	Grazing land use
rcagrid	Raster	Cost RCAs
wbgrid	Raster	Lakes, reservoirs, estuaries, etc.
costrcas.shp	Vector	Cost RCAs
graze_area.dbf	dbf table	Zonal statistics for RCAs and grazing land use
preds.shp	Point	Prediction site locations
rca_area.dbf	dbf table	Zonal statistics for RCAs
sites.shp	Point	Observed site locations and data
streams.shp	Vector	Stream reaches
waterbodies.shp	Vector	Lakes, reservoirs, estuaries, etc.
lsn	Folder	Contains the lsn.mdb and the lsn.ssn object folder
id_errors.txt	Text file	Contains the node point ID for complex confluences
lsn.mdb	Personal geodatabase	Landscape network (LSN) containing edges, sites, and prediction sites feature classes, as well as three dbf tables that store the topological relationships in the LSN
node_errors.shp	Vector	Potential topological errors
lsn.ssn	Folder	SSN object folder containing SSN associated files
edges.shp	Vector	Stream reaches; resides in lsn.ssn
sites.shp	Point	Observed sites; resides in lsn.ssn
preds.shp	Point	Prediction sites; resides in lsn.ssn
netID1.txt	Text file	Binary IDs for network 1; resides in lsn.ssn
⋮	⋮	⋮
⋮	⋮	⋮
netID16.txt	Text file	Binary IDs for network 16; resides in lsn.ssn

5. GETTING STARTED

1. Open ArcMap.
2. Add the STARS toolbox. In ArcToolbox, right click on 'ArcToolbox', and select 'Add Toolbox'. Navigate to the STARS toolbox and click OK. Then, right click on ArcToolbox, scroll down, select Save Settings, and click on To Default. Make sure that there are no spaces in the pathname where the STARS tools reside.
3. Change Environment Settings. Go to Menu, click on Geoprocessing, and select Environments. Expand M Values and set Output has M Values to Disabled. Repeat these steps for Z Values and click OK.
4. Overwrite outputs by default. In the menu, select Geoprocessing> Geoprocessing Options and check the box next to Overwrite the outputs of geoprocessing operations.

6. GENERAL DATA PRE-PROCESSING

1. Load the datasets into ArcMap

    1. Click on the Add data tool ⬇ and navigate to the directory that contains the streams, sites, and DEM data.
    2. Select the datasets and click Add. If a window pops up asking whether you would like to Create pyramids for dem..." click Yes.

    *Note: make sure there are no spaces in the pathnames where the datasets reside (e.g. c:\Users\MyName\My Documents\gisdata).

2. Project all of the input data to a common projection and datum

    ***Note**: the example dataset has already been projected to a common coordinate system.

    1. See the Data Management Tools> Projections and Transformations and the ArcGIS help if you need more information.

    *Note: You will be calculating and comparing distances and therefore, a geographic coordinate system (i.e. Latitude, Longitude) is not suitable. We recommend an Albers Equal Area or Universal Transverse Mercator projection instead.

3. Visually examine the digitized direction of the stream reaches for errors

    1. Click on the Table of Contents tab and then click on streams symbol ⬚ to open the Symbol Selector window. Scroll down and select the "Arrow at End" symbol →. Click OK.
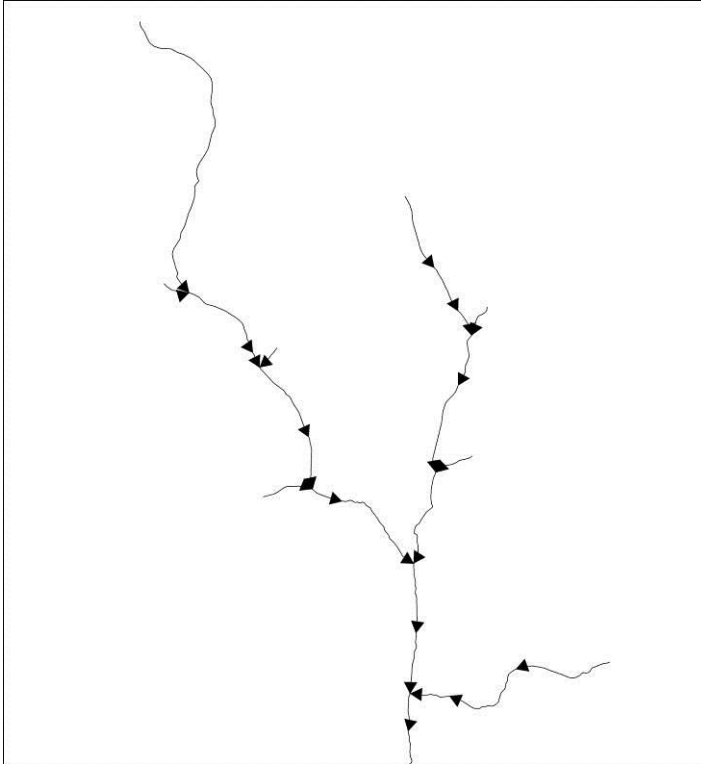
Figure 2. Stream segments correctly digitised in the downstream direction.

2. Undertake a brief visual examination of the segments to ensure that they are digitized in the downstream direction (Figure 2).

    If some segments are digitized in the upstream direction (i.e. from downstream to upstream), they will need to be edited manually.

    *Note: Any streams or waterbodies dataset can be used to build a Landscape Network. The datasets may be from a different sources or the streams could be derived from a DEM using the Spatial Analyst Tools>Hydrology toolset. However, the streams must be digitized in the downstream direction and free of topological errors. These issues will be described in detail below.

7. BUILD A LANDSCAPE NETWORK

A landscape Network (LSN) is a type of graph that is used to represent spatial context and relationships with additional geographic information (Theobald et al., 2006). A LSN is stored as an ESRI personal geodatabase and has four distinguishing features (Theobald et al., 2005):

- The LSN has the ability to store both the topology of a graph and the geometry of the nodes, reaches, and reach contributing areas (RCAs). Note that, the terms reach and segment are both used to refer to a single line segment in the polyline streams dataset.
- Nodes represent stream topologic breaks such as confluences, stream sources, or stream outlet points.
- Edges represent flow paths from node to node.

- RCAs represent the aerial extent that would theoretically contribute overland flow to a given edge in the absence of other hydrologic processes such as infiltration or evaporation.

For a detailed description of the LSN, please see Theobald et al. (2006) or Peterson and Ver Hoef (2014).

To begin, a simple LSN is generated from the streams data:

1. Create a new folder to hold the LSN. There should be no other files in this folder.

2. Open the Polyline to Landscape Network tool

    STARS>Pre-processing > Polyline to Landscape Network

3. Set the arguments (below) and click OK.

***Note**: **DO NOT use a shapefile named edges.shp.** An LSN will be produced, but there will be errors in the relationship tables.
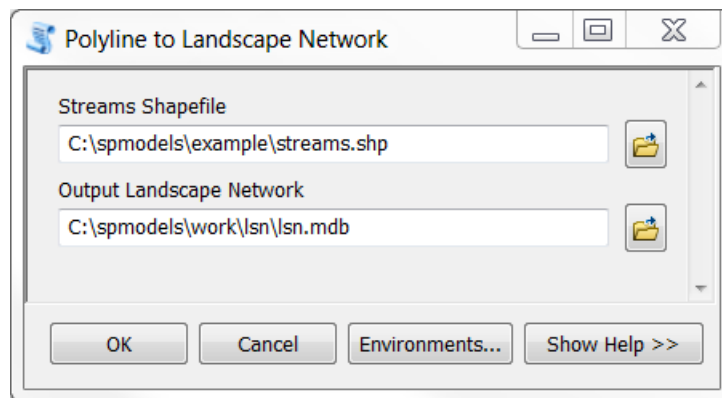
***Note**: Your path name may be different. **Be sure to include the file extensions .shp and .mdb.**

4. Load the LSN into ArcMap.

    *Note: An Unknown Spatial Reference warning box may appear indicating that the nodes feature class is missing spatial reference information. Click OK.

    If everything works correctly, the message FINISHED Polyline to Landscape Network Script will be shown in green. The tool produces a personal geodatabase with five components
    - **NODES (POINT FEATURE CLASS)**
    - **EDGES (POLYLINE FEATURE CLASS)**
    - **RELATIONSHIPS (TABLE)**
    - **NODERELATIONSHIPS (TABLE)**
    - **NODEXY (TABLE)**

    The program failed to produce a LSN if *any* of these components are missing.

*Note: It is not uncommon for the LSN to fail to build. If this occurs, please see the Trouble-shooting section for tips and potential work-arounds.

8. ELIMINATE TOPOLOGICAL ERRORS

It is ***extremely critical*** that the streams data are topologically correct to ensure that hydrologic distances and spatial relationships are calculated properly. Topological errors are common in GIS data and must be corrected before continuing. There are two tools provided in the STARS toolset to help identify these errors: the Check Network Topology tool and the Identify Complex Confluences.

Topological errors are identified in the edges feature class, but must be manually corrected in the streams.shp dataset that was used to build the LSN. Once the errors are corrected in the shapefile, a new LSN must be built. It is not possible to 'fix' the old LSN. It is also a good idea to run the Check Network Topology and the Identify Complex Confluences tools on each new LSN to ensure that it is free of topological errors.

Note, the development of a topologically corrected stream network can be the most time-consuming aspect of the modelling process.

8.1 Check network topology

The Check Network Topology tool is used to identify potential errors found in a LSN. *It does not work on data types other than a LSN.*

The Check Network Topology tool does two things to help identify topological errors. It adds and attributes a user-defined field in the node feature class (node_cat) with node categories or values, which include*: source, outlet, pseudo node, downstream divergence, converging stream,* and *confluence*.

Source nodes occur at the upstream node of an edge if there is not an adjacent edge upstream (i.e. headwater stream - does not have an upstream input).

Outlet nodes occur at the downstream node of an edge that does not flow into another downstream edge. These represent watershed outlets or pour points.

Converging stream nodes occur at the downstream node of two edges that converge, but do not flow into another downstream edge (Figure 3). **Converging stream nodes must be edited before an .ssn object can be produced.**

Pseudo nodes are locations where the upstream node of one edge is coincident with the downstream node of another edge. This occurs when two polyline segments make up a single reach. **Large numbers of pseudo nodes are undesirable in a LSN because they increase the storage space required for the .ssn object in R.**

Confluence nodes occur where the downstream nodes of two or more edges converge and flow into a single downstream edge.

Downstream divergence nodes are locations where one edge flows into two or more edges. They represent divergent flow associated with braided stream channels. **Downstream divergence nodes are not permitted in an .ssn object.**

These node categories can be displayed and visually examined in ArcMap to identify potential topological errors. For example, if an outlet node is located in the middle of a stream network it is likely a topological error. Although these categories are useful, it may still be difficult to identify errors visually when working with a large dataset.

The Check Network Topology tool also identifies potential topological errors, which are computed and stored in a shapefile called node_errors.shp. This file will be located in the same directory as the input LSN. Topological node and edge errors include outlet nodes that are within a user specified search tolerance distance of source nodes, and nodes that are snapped to an edge without a coincident node (i.e. an edge is connected to another edge at its middle).

Topological Error | Topologically Correct
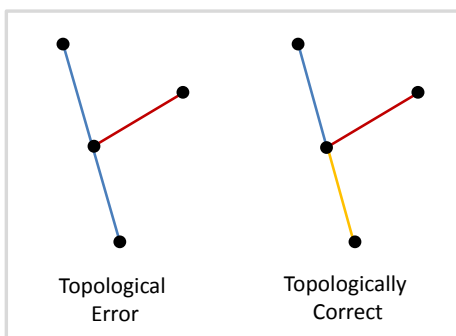
For example, the blue edge on the left does not have a coincident node at the end node for the red segment; this is a topological error. In contrast, there is a coincident node where the blue, red, and orange segment intersect on the right.

The points in node_errors.shp can be displayed visually in ArcMap, along with the edges feature class and the node categories, to help identify topological errors. You may need to zoom in on each of the points to look for topological errors in the edges feature class. Examples of topological errors and their effects are given in the FLoWS user manual (Theobald et al., 2006, pg. 23).

1. Open the Check Network Topology tool: STARS> Pre-processing> Check Network Topology
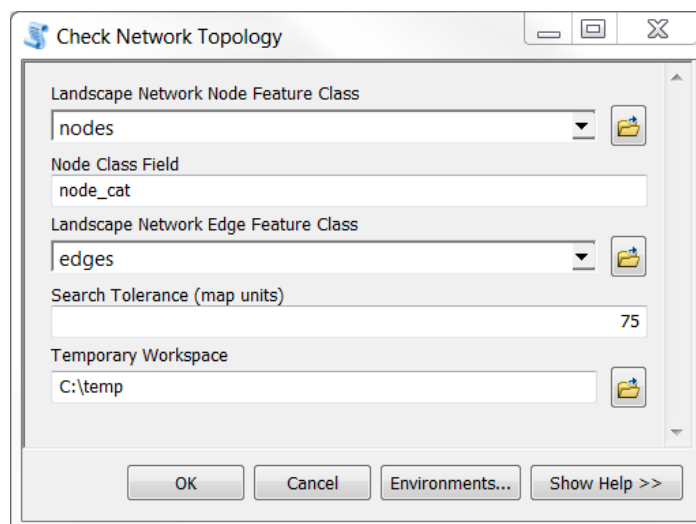
*Note: It may be necessary to browse for the data or to include the entire path name. Also, some computers do not have a C:/temp directory and so an alternative temporary workspace should be selected before running the Check Network Topology tool.

2. Set the arguments (above) and click OK.

A message, Finished Check Network Topology Script, will be shown in green if the tool completed successfully. If this message does not appear, delete any remnant files left in the c:/temp directory and try running the tool again.

3. *If necessary*, reset the source for the node feature class so that the node categories can be displayed.

    1. Double click on the nodes feature class to open the Layer Properties window.
    2. Select the Source tab.
    3. Click on the Set Data Source button, Set Data Source... , and navigate to the LSN geodatabase containing the nodes feature class. Select nodes, click Add, and click OK.

4. Add node_errors.shp to ArcMap and examine each potential error. The shapefile will be located in the same folder as lsn.mdb. The node_errors attribute table contains five fields: FID, Shape, pointid, xcoord, and ycoord. The xcoord and ycoord fields are populated with zeros. The pointid field in the node_errors shapefile corresponds to the pointid field in the nodes attribute table, which also contains the x, y coordinates.

    *Note: The node_errors shapefile contains *potential* errors; each point must be visually examined to determine whether or not it represents a true topological error.

5. Examine each node category for potential errors.

    1. Open the nodes attribute table, Click on the Table Options dropdown menu, and select Select By Attributes.
    2. Create the expression [node_cat] = 'Outlet' and click Apply to select all Outlet nodes in the LSN.
    3. Visually examine the selected Outlet nodes and look for Outlets that are not at known watershed outlets or along the edge of the map. If an Outlet is the result of a topological error, then it must be manually edited in the streams shapefile.
    4. Select [node_cat] = 'Converging stream'. Converging streams are not considered topological errors in the GIS; however, they must be manually removed before the .ssn object can be generated and exported.

    Note that, there are no converging streams in this LSN.
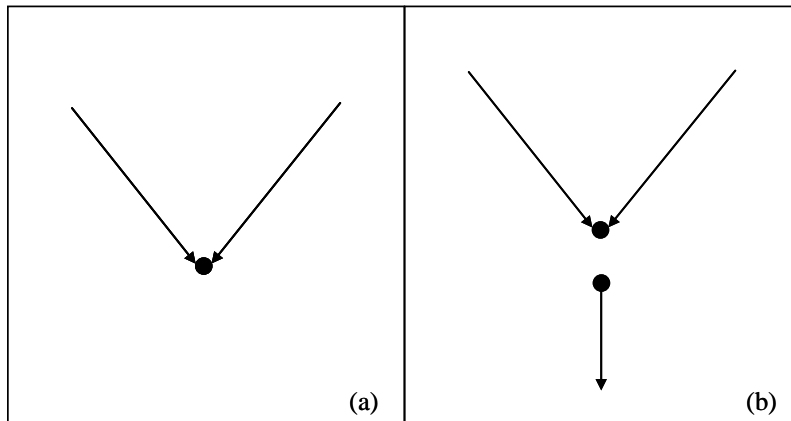
Figure 3. Converging stream nodes occur at the downstream node of two edges that converge (a,b), but do not flow into another downstream edge. This commonly occurs at the edge of the streams shapefile (a) or may be the result of topological errors within the stream network (b).

8.2 Braided channels

Braided stream reaches are a common occurrence in stream datasets that have been digitized from topographic map data. Although braiding is a natural process, braided streams are not permitted in the .ssn object. The Check Network Topology tool can be used to identify braided channels.

1. Open the Check Network Topology too, create the expression [node_cat] = 'Downstream Divergence' and click Apply to select all downstream diverging nodes in the LSN.

The streams dataset must be edited to remove braided channels before the LSN is constructed.

8.3 Identify complex confluences

A .ssn object cannot be created if the stream network contains confluences where the downstream nodes of three or more edges converge and flow into a single downstream edge (Figure 4a). Since this is not considered a topological error in a GIS or a LSN, the Identify Complex Confluences tool has been included in STARS to help users identify these features. This tool locates LSN nodes where the error occurs and produces a text file containing the pointIDs. These errors must be manually edited (Figure 4b) before a new LSN is generated.
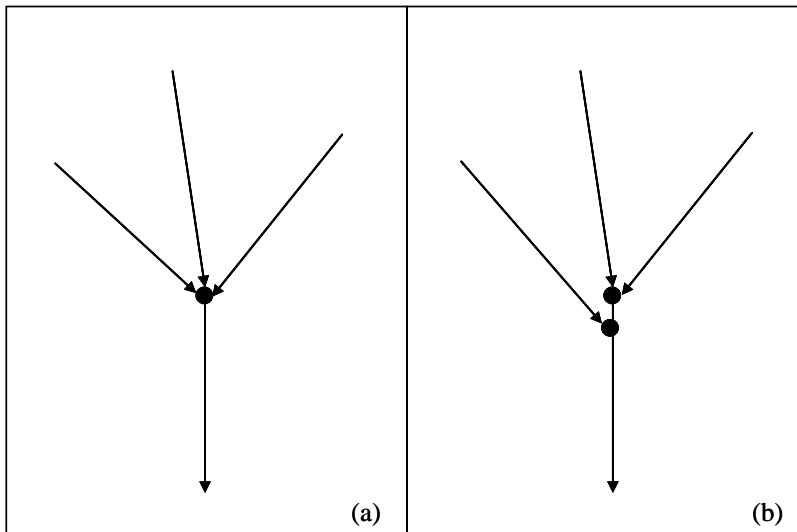
Figure 4. The stream network may contain confluences where three or more edges converge and flow into a single downstream edge (a). If this occurs, the nodes (black circles) must be identified using the Identify Complex Confluences tool and the error manually corrected (b) in the streams shapefile before the LSN is rebuilt.

1. Open the Identify Complex Confluences tool

    STARS>Pre-processing>Identify Complex Confluences

2. Set the arguments and click OK.
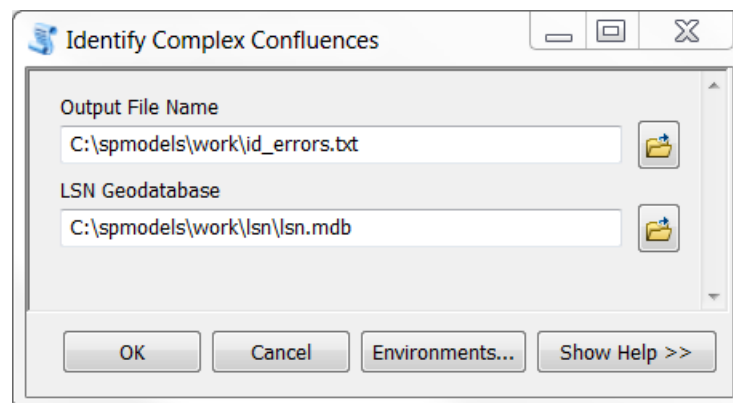
The Identify Complex Confluences tool produces a comma delimited text file (.txt), which contains the pointid values for nodes where more than two edges converge and flow into a single downstream edge.

A message, Program finished successfully, will be shown in green if the tool finished without errors.

*Note, if the output text file only contains a comma, then the tool has run successfully and no complex confluences were identified in the LSN. Also, remember to add the .txt extension to the filename.

9. CREATING REACH CONTRIBUTING AREAS

Reach contributing areas (RCAs) form a detailed tessellation of non-overlapping, edge-matching interbasins. The RCA boundaries are based on topography and include nearby areas that would theoretically contribute overland flow, were it to occur, to a given reach (Theobald et al., 2005). Oftentimes, national-scale streams datasets include RCAs, but are referred to using different terminology; for example, they are referred catchments and catchments in the US NHDPlus (Horizon Systems Corporation 2007) and subcatchments in the Australian Geofabric dataset (Bureau of Meterology 2012). If these data are available, it is unnecessary to generate them using the STARS toolset.

The Create Cost RCAs tool has been provided to generate the RCAs, which are delineated using edges, waterbodies, and a DEM as input. There is a one-to-one relationship between reaches (i.e., edges or segments) and RCAs, which allows the surrounding landscape to be explicitly linked to the reach.

1. Load the input data into ArcMap: LSN edges, the DEM, and a waterbodies dataset.

2. Set the Geoprocessing Environment. This ensures that new rasters have the appropriate cell size and extent. It also provides a way to 'snap', or align, new rasters with existing rasters.

   On the main menu, click on Geoprocessing, scroll down, and click Environments... This will open the Environment Settings window.

   > ***Note**: You may need to enable the Spatial Analyst Extension if it is not available. On the main menu, go to Customize>Extensions. This will open the Extensions window. Make sure that the box next to the Spatial Analyst is checked.

   1. Under Workspace, set the Current Workspace and Scratch Workspace.

   2. Under Processing Extent, set Extent to Same as layer dem.

      Note: You can also define a box using x,y coordinates if the extent of the DEM is much larger than the streams dataset. Be sure not to make the bounding box too small because the watershed for the stream network will be larger than the actual stream network. A visual analysis of the DEM and the streams dataset is usually sufficient to set the bounds.

   3. Under Raster Analysis, set the cell size to Same as layer dem. This will automatically set the cell size. No Mask is used.

   4. Click OK.

3. Format the edges and waterbodies datasets

An edges attribute should be created to relate the RCAs to the edges. It is better not to use the OBJECTID or rid fields because they may change if the LSN has to be rebuilt. Instead, create a new field for this purpose that will remain unchanged regardless of data transformations, conversions, and other manipulations.

1. Add a Long Integer field named reachid to the edges attribute table. This attribute will be used to relate the RCAs to the edges
    1. Calculate field: reachid = OBJECTID

        *Note: A message box may open warning you that you are about to calculate outside of an edit session. Click OK.

    2. Close the edges attribute table

2. Convert the edges to raster format
    1. In ArcToolbox, click on Conversion Tools, open the To Raster toolbox, and open the Polyline to Raster tool.
    2. Set the arguments (below) and click OK.
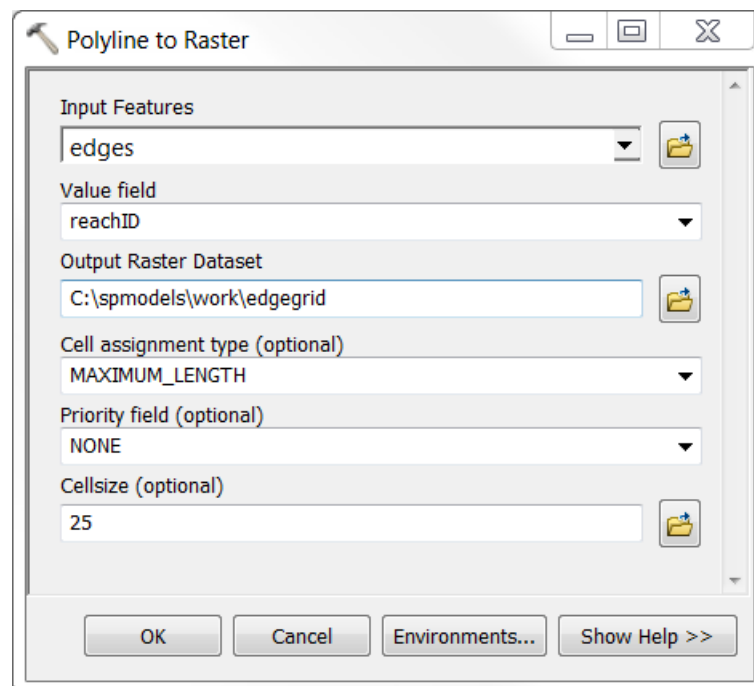
This will create a raster version of edges. Each edgegrid cell contains a Value attribute that is equal to the edge reachid value. Non-edge values are set to NoData. Overlay the new edge grid on top of the dem, zoom in, and check to ensure that the cells align perfectly.

3. Reclass the edges
    1. In ArcToolbox, click on the Spatial Analyst toolbox, select Reclass, and open the Reclassify tool.
    2. Set the first two arguments

            Input Raster = edgegrid
            Reclass Field = Value

3. Set the Reclassification Parameters
    1. Click on the Classify button to open the Classification window
    2. Set the Classification Method to Equal Interval and set the number of Classes to 1
    3. Click OK to close the window
    4. Under Reclassification, click on a New value cell to edit
    5. Set all New values = 5 and NoData values = 0
    6. Set the remaining arguments (below) and click OK.
    *Note: You may need to navigate to your working space to save the raster.
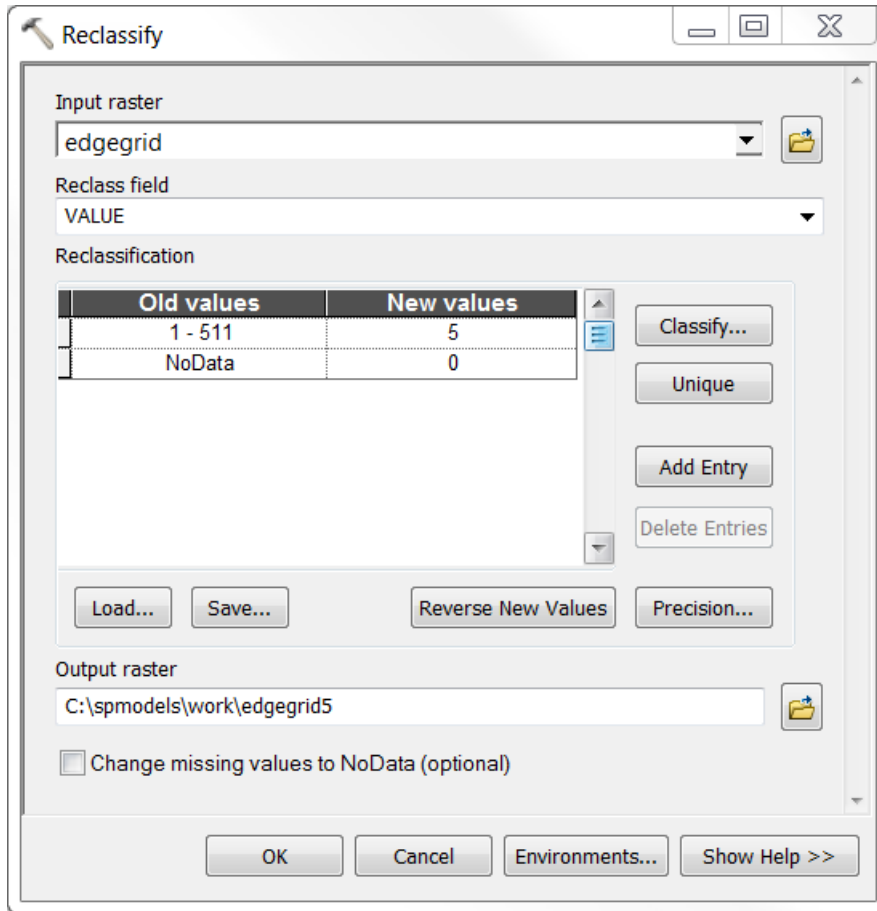
The result is a raster layer of the edges, with edge cells containing a value equal to 5 and all other cells a value equal to 0. The edgegrid5 will be used to 'burn in' the streams to the DEM.

4. In ArcToolbox, use the Conversion Tools > To Raster > Polygon to Raster tool to convert waterbodies dataset to raster format.
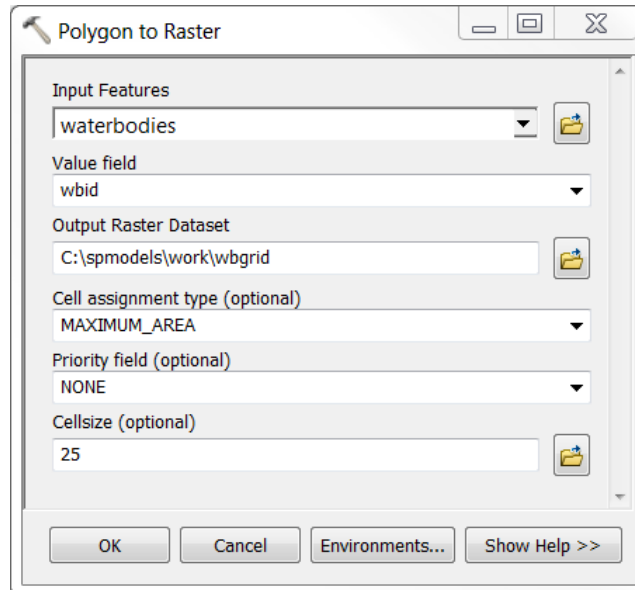
It does not matter which field is used for the Value field argument, as long as each waterbody is assigned a unique identifier. The actual value is overwritten when the RCAs are delineated. **However, the resulting wbgrid must be an integer grid.**

5. Remove waterbodies.shp from the Table of Contents.

4. Format the DEM

Sinks and peaks in the DEM may be errors resulting from the data resolution or the rounding of elevations. The sinks must be filled to ensure proper delineation of RCAs.

1. Burn the streams into the DEM. This step is necessary if the LSN edges feature class was derived from a vector of streams or another DEM. If it was derived from the same DEM that will be used to create the RCAs, skip these steps and go to step 5.2.

    1. Open the Spatial Analyst toolbox, scroll down and select the Map Algebra toolbox, and open the Raster Calculator tool.
    2. Type in the following expression (below), set the Output raster and name it **demburn5**, and click OK.
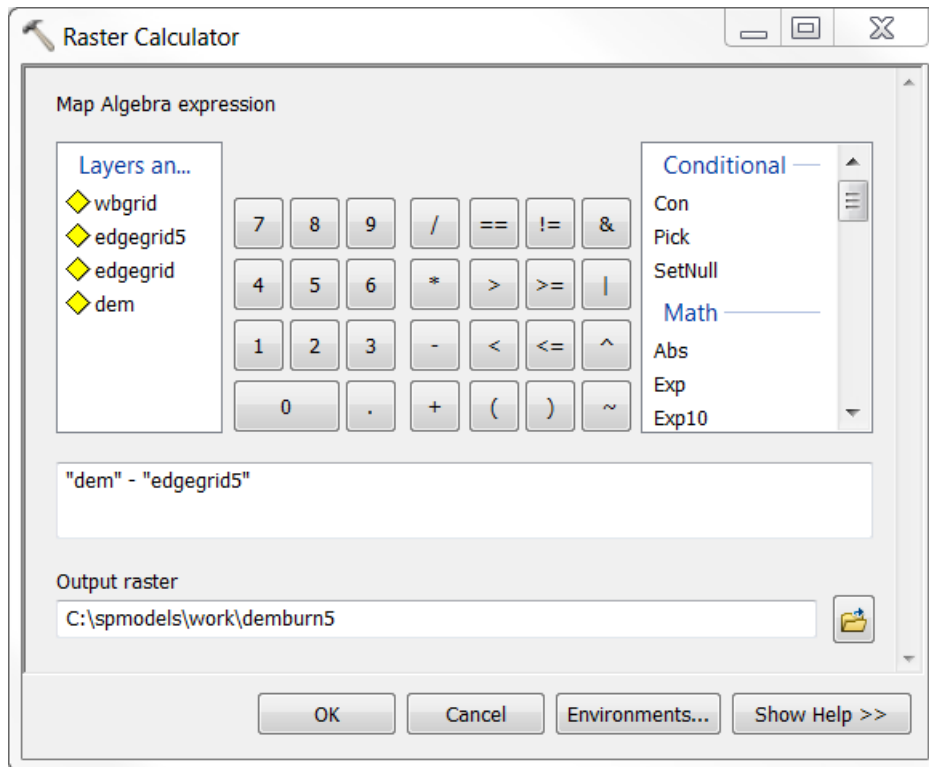
3. Remove the original DEM and edgegrid5 from the Table of Contents.

2. Fill demburn5 (or dem if you skipped step 5.1)
    1. Go to ArcToolbox>Spatial Analyst Tools>Hydrology>Fill to open the Fill tool.
    2. Set the arguments and click OK.
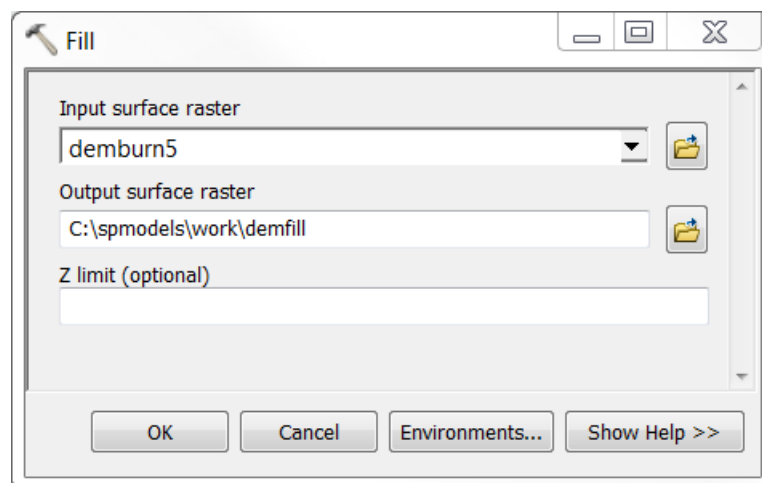
5. Create Cost RCAs

    1. Examine the DEM, edgegrid, and wbgrid rasters to ensure that they are properly aligned. All raster cells should have the same spatial resolution (Figure 5). If a dataset

is not properly aligned with the DEM, delete it, reset the Geoprocessing Environments (Section 9.2), and generate the dataset again.

demfill

edgegrid

Figure 5. The three rasters must have the same spatial resolution and must be aligned before the RCAs are created.

2. Go to ArcToolbox>STARS>Pre-processing>Create Cost RCAs. This will open the Create Cost RCAs tool.

3. Set the arguments and click OK. The program may take a few minutes to run.
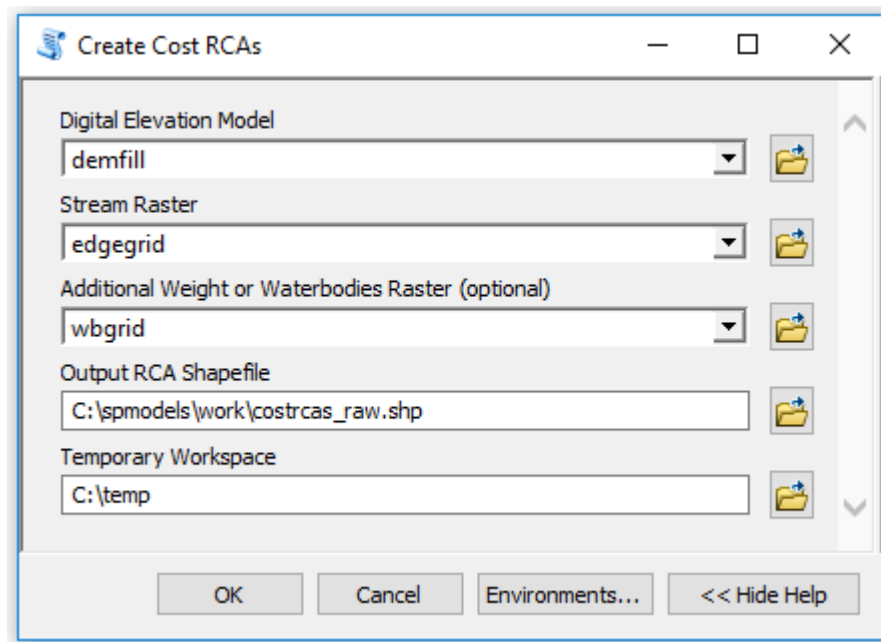
*Note: the .shp extension for the Output RCA Shapefile will be added automatically.**

The output of the Create Cost RCAs tool is a shapefile of RCAs, which is automatically added to the view. The gridcode attribute found in the RCA attribute table is equal to the reachid value assigned to the edges attribute table. This attribute enables each RCA to be directly linked to a single edge. Although there is generally a 1:1 relationship between edges and RCAs, this is not

always the case. For example, it is not uncommon to find that a relatively small number of edges have not been assigned an RCA. This may occur if the length of the stream reach is short in relation to the spatial resolution of the DEM (e.g. stream reach length = 10m and the DEM spatial resolution = 25m). When this situation occurs, the RCA is essentially too small to delineate. When a group of edges is part of a waterbody, such as a lake or reservoir, some of the edges will not be assigned RCAs. Also, edges outside the boundary of the DEM will not be assigned an RCA.

In a small number of cases, an edge may be assigned multiple RCA polygons (Figure 6a). In addition, an RCA may not be associated with an edge if waterbody does not intersect a stream reach (Figure 6b); these are generally assigned a gridcode value equal to 0.
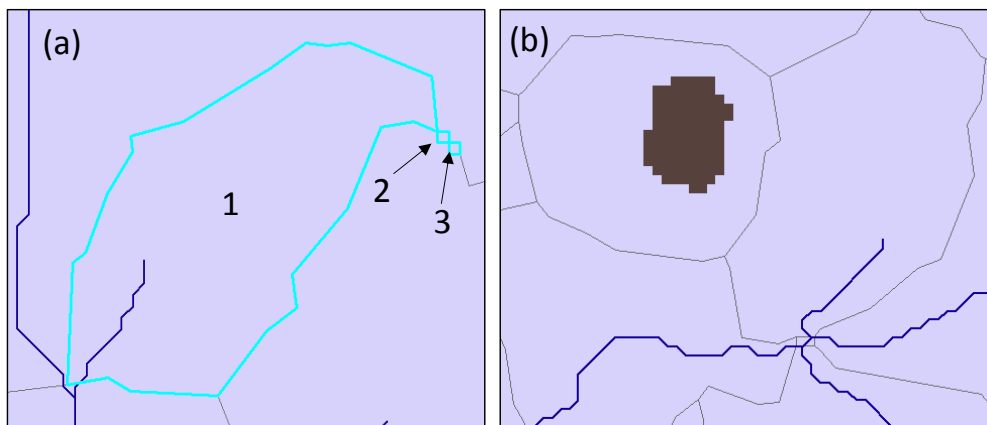
Figure 6. Multiple reach contributing areas (RCAs) may be associated with a single edge (a), while other RCAs may be associated with an isolated waterbody, but not an edge (b).

The edges feature class can be joined to the costrcas_raw shapefile using the edges 'reachid' field and the costrcas_raw GRIDCODE field. This provides a way to select edges that are not associated with an RCA ("costrcas_raw.FID" Is Null) so that they can be examined in ArcMap. In the example dataset, RCAs were not generated for six edges (Figure 7), which were located within a waterbody.
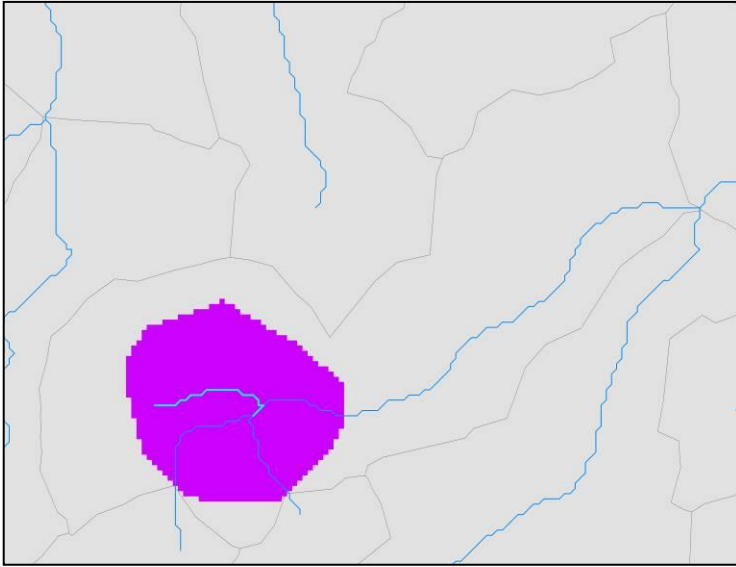
Figure 7. A relatively small number of edges will not be associated with an RCA. This may occur when the edge is found within a waterbody (purple area).

4. Open the edges attribute table and notice that the number of rows is larger than the number of rows in costrcas_raw. This is because there are duplicate polygons (Figure 6a). We will need to dissolve these polygons. Go to ArcToolbox > Data Management Tools > Generalization > Dissolve. This will open the dissolve tool. Set the arguments as shown below and click ok to create the final cost_rcas.shp.
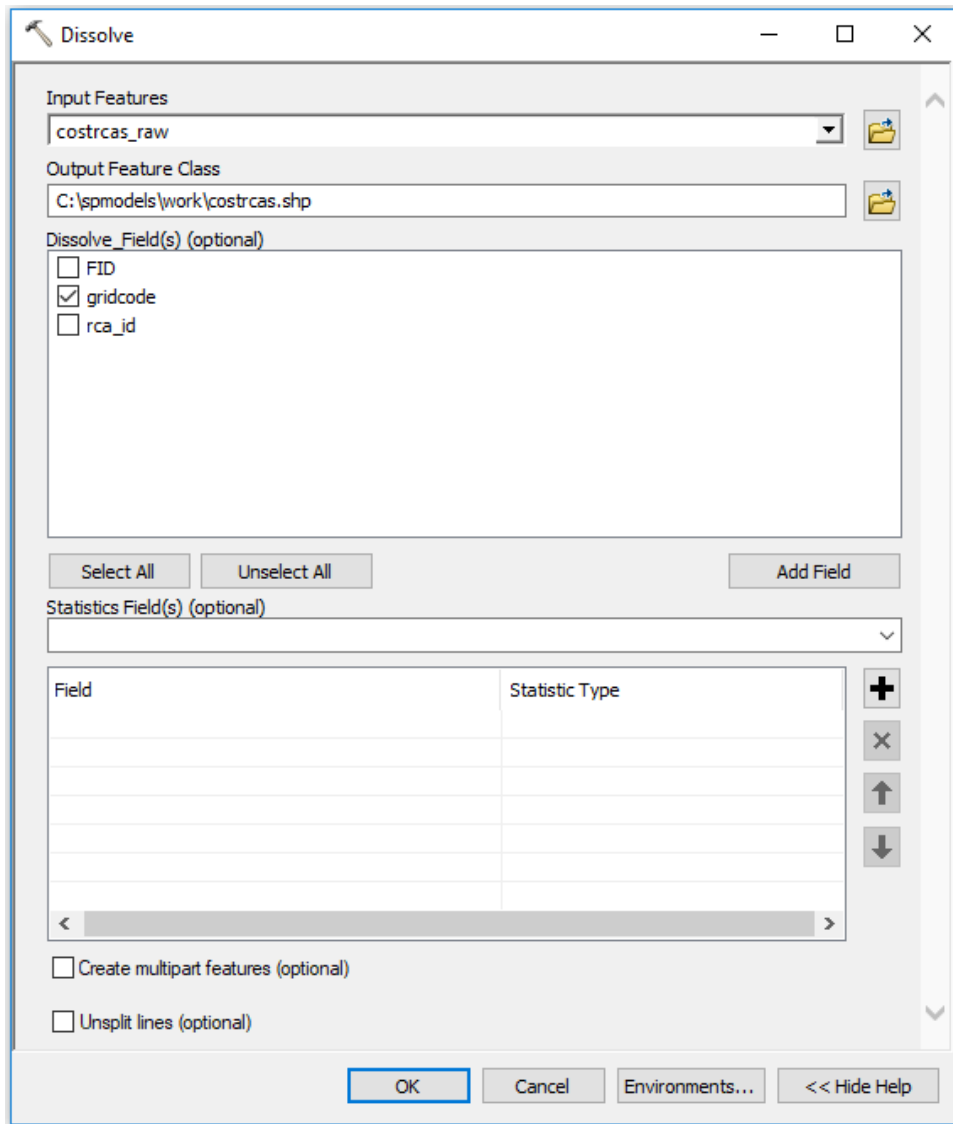
10. CALCULATING RCA ATTRIBUTES

The 1:1 relationship between RCAs and reaches can be used to provide information about the condition of the landscape immediately surrounding the reach. These landscape characteristics can include areas, means, and counts. For example, land use area, mean elevation, or the number of point source inputs in the RCA.

1. If you have not already done so, set the Geoprocessing Environment to match the filled DEM (demfill).

2. Convert the RCA shapefile to raster format using the Conversion Tools > To Raster > Polygon to Raster tool.

    Remember, the attribute GRIDCODE = reachid.
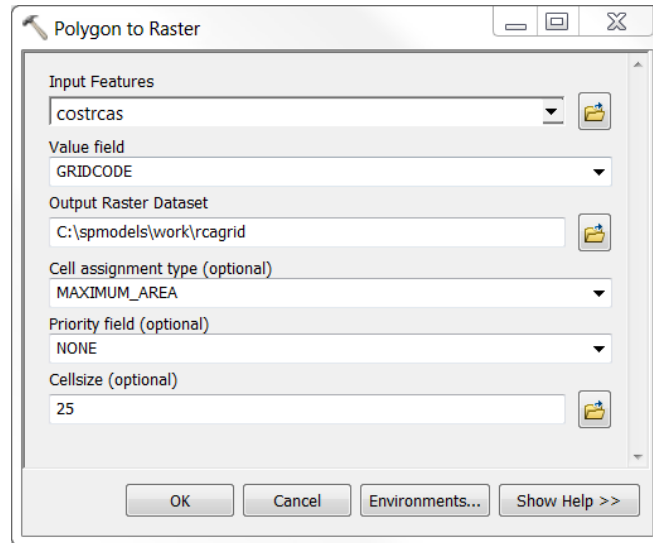
This will produce a raster of RCAs with a Value attribute that is equivalent to the reachid. Each cell contains the reachid that the cell is associated with.

3. Check to ensure that the new RCA raster is properly aligned with the other rasters (see Figure 5)

4. Reclass the landscape value raster

    This step is necessary when the landscape characteristic is made up of categorical values, such as land-use classes or presence/absence data. Skip this step if the characteristic is made up of continuous values, such as elevation or rainfall.

    Landscape cell value of interest = 1
    Other = NoData

    A reclassed land use raster, grazing, has been included in the example dataset. Add the grazing raster to ArcMap.

5. Calculate the RCA attribute

    1. Open the Spatial Analyst toolbox, scroll down to Zonal, and open the Zonal Statistics as Table tool.

    2. Set the argument values. In this example, the landscape value raster, grazing, represents grazed areas.

    3. Click OK.
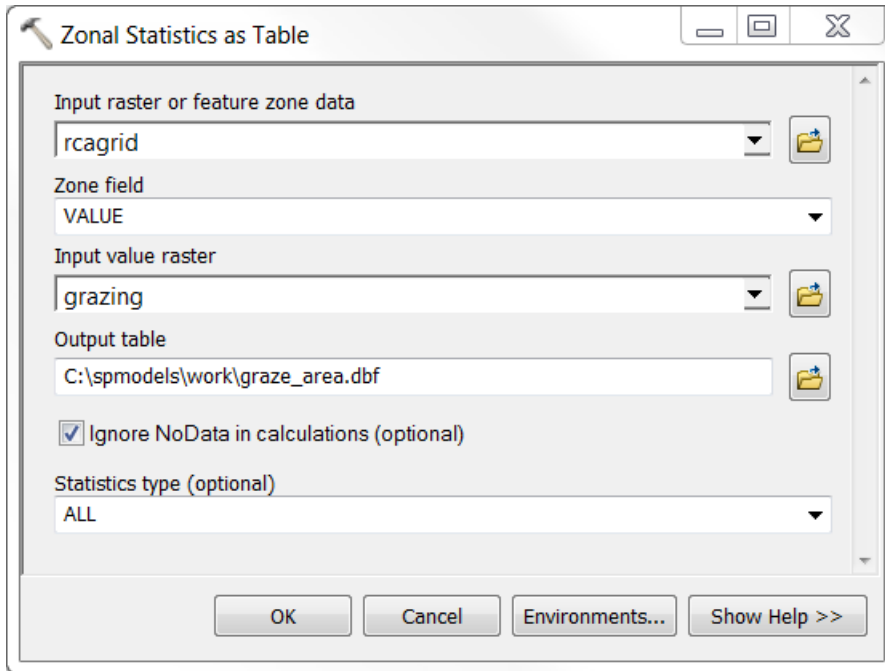
The Zonal Statistics as Table tool produces a dbf table with landscape value statistics, such as mean, count, or area for each RCA. The Value attribute is equivalent to the reachid, which enables these attributes to be associated with a single edge. The area shown in the dbf table will be in square projection units. In this case, they are in square meters because the projection is Albers.

6. Add the RCA attribute to the edges attribute table

   1. Add a field of type DOUBLE to the edges attribute table that represents the grazed land use in the RCA

      Name = rcaGrazKm2
      Type = Double

   2. Join the dbf table to the edges attribute table
      1. In the Table of Contents, right click on edges, scroll down, and select Joins and Relates>Join… This will open the Join Data window.

      2. Set the arguments (below) and click OK.
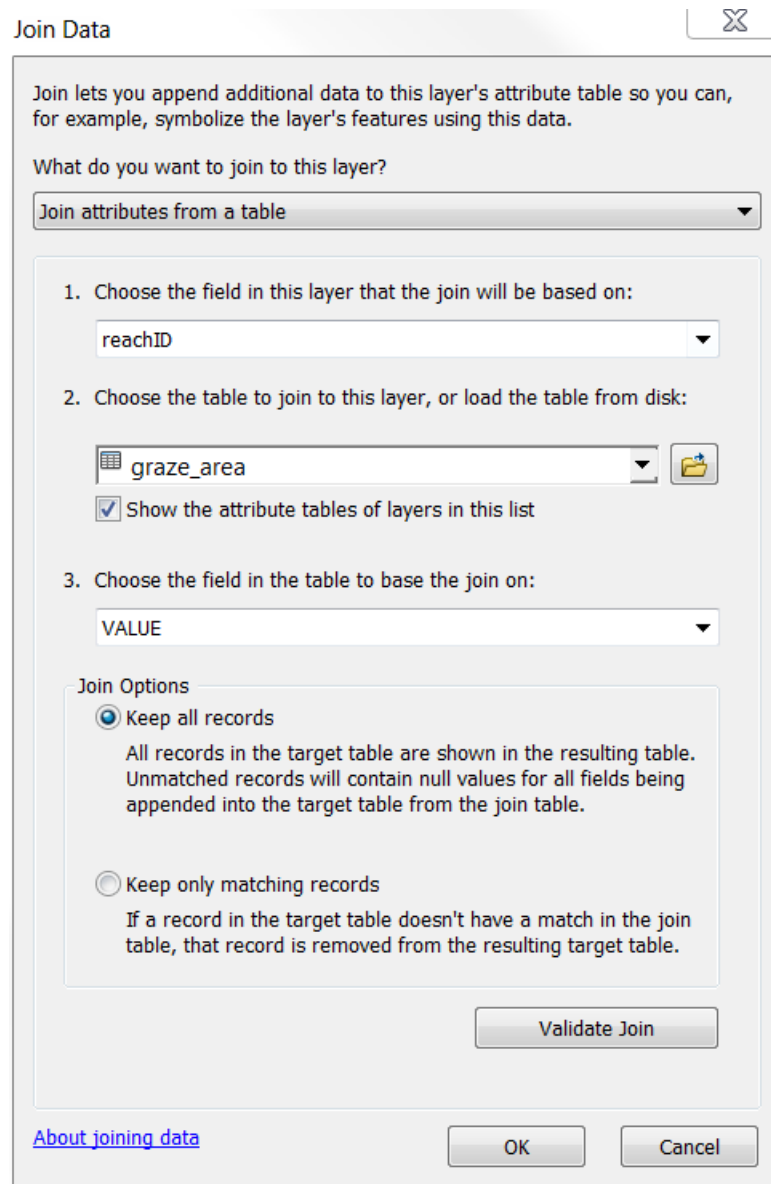
If a window opens asking whether you would like to index the field, click **No**.

3. Open the edges attribute table and scroll to the right until you see the empty RCA landscape value attribute. In this example, it will be edges.rcaGrazKm2.

4. Calculate the field. In this example, the area is also being converted from square meters to square kilometres.

edges.rcaGrazKm2 = [graze_area.Area] * 0.000001

When the edges attribute table and the zonal statistics table are joined, RCAs that do not contain grazed areas will have NULL statistics values.

*Note: In this example, we are interested in the area of developed land in each RCA. In other cases, we might be more interested in the mean or the count statistic.

7. Remove the Join
    1. In the Table of Contents, right click on edges, scroll down, select Joins and Relates>Remove Join(s)> Remove All Joins

8. Set NULL rcaGrazKm2 values = 0
    1. In the edges attribute table, open the Select by Attributes window (section 8.1.5.1).

    2. Select NULL values

       [rcaGrazeKm2]Is Null

    3. Use the Field Calculator to set the rcaGrazKm2 field = 0 for the selected records. Note, only the selected records will be calculated.

    4. Clear the selection and close the edges attribute table

       It is necessary to set the NULL values = 0 for areas and counts. However, this step should be skipped if means are being calculated for continuous values, such as elevation, because zero may be a meaningful value.

11. CALCULATING THE RCA AREA

The RCA area is calculated and used to derive the watershed area for the downstream node of each edge in the LSN. The watershed area is a useful attribute. It may be used as an explanatory variable in the spatial model or to convert other explanatory variables, such as grazed area, to percentages or proportions. It may also be used to calculate the Segment Proportional Influence (Section 17), which is needed to calculate the Additive Function (Section 18). Note that, national-scale datasets, such as the US NHD-Plus or the Australian Geofabric, often have attributes assigned to stream segments representing RCA and watershed attributes such as percent land use, RCA and watershed area, number of road crossings, or climate statistics. If these variables are available, there is no need to recalculate the RCA and watershed area attributes.

1. Calculating the RCA area is similar to calculating RCA attributes (Section 10) except that the RCA raster is used as both the zone dataset and the Value raster in step 10.5
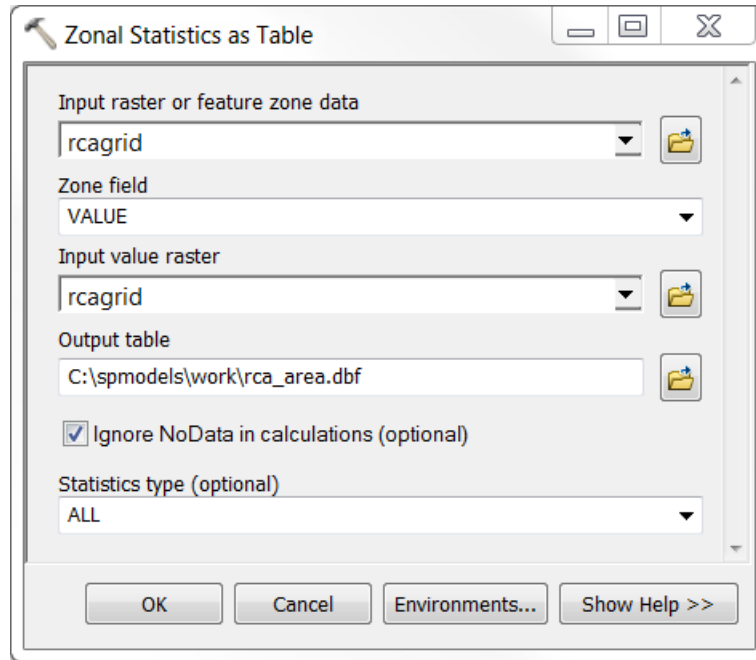
Using the RCA raster as both the zone and value raster is convenient when the dataset is relatively small, but the operation can be slow when the RCA raster is large. An alternative is to use another Value raster that provides continuous coverage, such as elevation, of the study area. When the zonal statistics are calculated using a continuous coverage, the COUNT attribute will represent the number of cells in each RCA and may be used to calculate the RCA area (COUNT * [cell resolution]^2). This is particularly convenient if an elevation attribute, such as the mean, is being calculated for each RCA.

Hint: It is better to use the Zonal Statistics as Table tool to calculate the RCA area rather than calculating it from the costrcas.shp polygon area. The areas of polygons in a shapefile will be different than the areas calculated using a raster dataset because of the stair-stepped edge pattern created by square raster cells. In addition, the RCA for some edges may be represented by multiple polygons. As a result, we prefer to calculate RCA area using data in a raster format.

Follow the steps in steps in Sections 10.5 to 10.8 to create a new field in the edges attribute table representing RCA area. Name the field rcaAreaKm2. Remember to set NULL values = 0.

12. ACCUMULATING WATERSHED ATTRIBUTES

The watershed includes the entire land area that contributes flow to a single stream outlet. Watershed attributes can be calculated by using the topological relationships in the LSN to accumulate the RCA attributes downstream. The STARS Accumulate Values Downstream tool produces watershed attributes for the downstream node of each edge, which are assigned as an attribute to the edges feature class. Again, there is no need to accumulate watershed attributes if your streams dataset already contains these attributes.

In this example, the rcaGrazKm2 attribute will be accumulated to calculate the total grazed area in the watershed.

1. Double click on the STARS>Calculate>Accumulate Values Downstream tool. This will open the Accumulate Values Downstream tool.

2. Set the argument values and click OK.
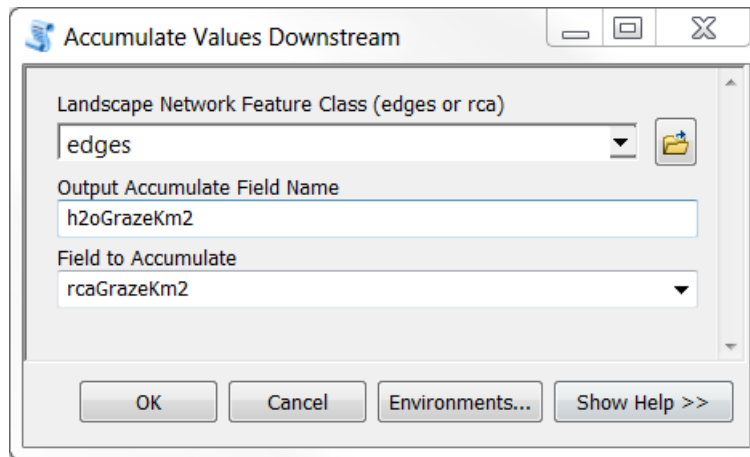
If the tool runs without errors, a green message will appear: Finished Accumulate Values Downstream Script.

*Note: You may need to load edges to the table of contents for this tool to run properly.

*Note: there is not a specific tool that can be used to accumulate means, such as mean elevation. A workaround is to create a new field and make it equal to the product of the mean and the RCA area. Then, accumulate this product. To get the area-weighted mean for each edge's watershed, create a new field and simply divide the accumulated product by the watershed area of the edge.

13. CREATE PREDICTION POINTS

The Create Prediction Points tool is included in STARS versions $\geq$ 2.0.7. This tool has been provided for those that do not have an existing set of prediction points to work with. If you already have a set of prediction points, move to Section 14.

1. Go to STARS v2.0.7 > Pre-processing > Create Prediction Points. This will open the Create Prediction Points window.
   a. Set the argument values (below). Note that the interval is in map units and in this example the units are meters.
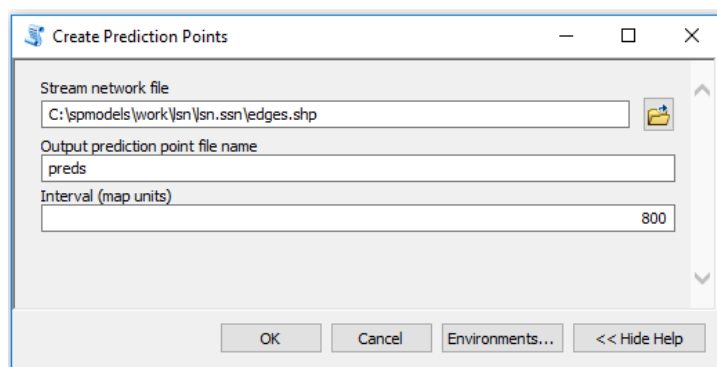
b. Click OK.

The tool creates a shapefile of prediction points that can be incorporated into the LSN using the Snap Points to Landscape Network tool (see Section 14).

14. INCORPORATING SITES INTO THE LSN

One challenge of working with GIS data is that sample sites collected within a stream are not always located directly on a stream segment, even though they should be. This is a common phenomenon that can result for a variety of reasons. GPS-based points generally have some error and do not always fall directly on a vertex or line segment representing a stream. Some stretches of river can move (e.g. meander) slightly from their mapped position. Streams are often represented on a map by lines and so samples collected on the banks of a large river may not fall directly on a line segment. When streams are represented at coarser scales the digital streams datasets may contain mapping errors and generalizations, such as the absence of small tributaries and the homogenization of form. Regardless of the error source, the sample sites must fall exactly on a stream line. Our solution is to 'snap' the sites to the nearest stream segment and manually examine each site to ensure that it is located on the correct stream segment.

The Snap Points to Landscape Network Edges tool allows point features (i.e., survey sites, dams, stream gauges) to be incorporated into the LSN using dynamic segmentation (Theobald et al., 2006). Dynamic segmentation involves intersecting each point with the closest edge segment, physically moving the point to that location, and calculating the distance ratio from the end of an edge to the point location. Once this is complete, distances between any two point locations in the LSN can be calculated using two pieces of information: the reach identifier (rid) and the distance ratio.

The Snap Points to Landscape Network tool will snap the sites to the nearest edge, but this may not be the correct location. We recommend visually inspecting each survey site to ensure that the closest edge is the correct edge. If the site is closer to another edge it must be moved manually using the editing tools provided in ArcGIS. If detailed information is available about the relative site location along an edge, all of the sites can be edited and moved to their appropriate locations.

Note, if the sites have been incorporated into an LSN previously, the OBJECTID, rid, and ratio fields must be deleted from the sites attribute table before snapping.

1. Go to STARS>Pre-processing>Snap Points to Landscape Network. This will open the Snap Points to Landscape Network window.

2. Set the argument values (below). If the sites have already been edited and snapped to the appropriate edges, set the Search Radius = 1. Otherwise, examine the data and choose an appropriate search radius based on the maximum distance between sites and segments.

   *Note: Your path name may be different. **Be sure that the file extensions .shp and .mdb are included.**
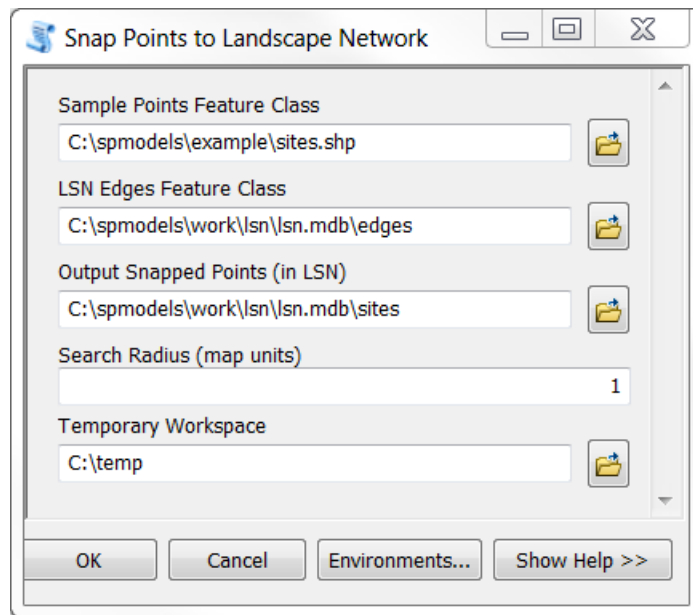
3. Click OK. The tool may take a few minutes to run.

   If the tool runs successfully, the message Finished Snap Points to Landscape Network Edges will appear.

A new sites feature class will be written to the LSN geodatabase if the survey locations are successfully incorporated. The new sites attribute table will contain some new fields, two of which are the *ratio* and *rid* fields. The rid field indicates which edge the site has been snapped to. The ratio for each site, r_i, provides the exact location along the edge:

$$r_i = \frac{d(l_j, s_i)}{L_j}$$

where $d(l_j, s_i)$ is the distance travelled along edge j (i.e. hydrologic distance) between the most-downstream location on an edge, l_j, and the site location, s_i, and L_j is the total length of the jth edge. Together, the rid and ratio values are used to identify a site's location within the LSN. This is extremely useful because it allows attributes to be estimated for site-specific locations along the edges. The site ratios range between 0 and 1 ($0 \leq$ ratio ≤ 1) because they are proportions; an error has occurred if values outside this range are present. Also, be sure to compare the total number of sites in the feature class to the total number of sites in the shapefile to ensure that all of the sites have been incorporated into the LSN.

Repeat steps 14.1 – 14.3 to incorporate preds.shp into the LSN. Name the new feature class preds.

*Note: The run time for the Snap Points to Landscape Network tool will increase depending on the number of sites and edges in the LSN.

15. CALCULATE WATERSHED ATTRIBUTES

The watershed attributes assigned to the edges represent the watershed attribute for the downstream node of each edge. However, survey sites can fall anywhere along an edge. The Calculate>Watershed Attributes tool enables watershed attributes to be estimated for any site that has been incorporated into the LSN.

$$W_i = \sum_{U[j]} \left(RCA_j \right) + \left[\left(1 - r_i \right) RCA_j \right]$$

where W_i is the watershed attribute for each survey site, s_i, RCA_j is an attribute summarized over the jth edge, $U[j]$ is the complete set of edges found upstream of the jth edge, and r_i is the ratio value for the site. A new field is added to the sites attribute table that contains the W_i for each site.

*Note: W_i is simply an *estimate* of the true watershed attribute because RCA_j does not contain spatial information about attribute (i.e. land use) variability at the sub-RCA level (Figure 8).
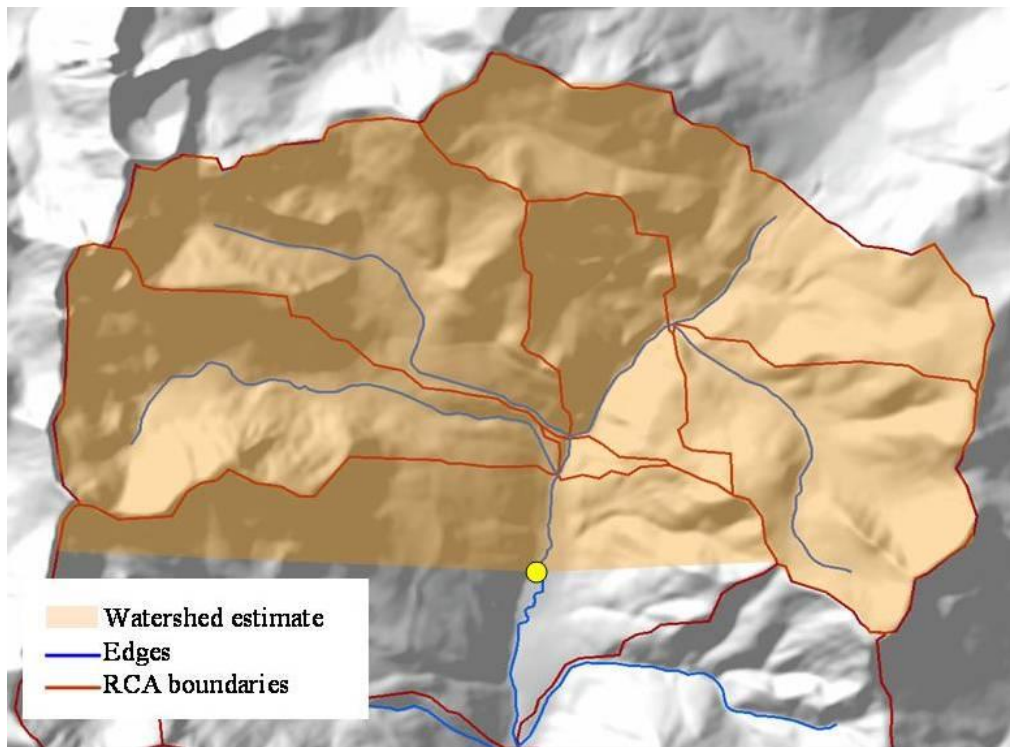
Figure 8. The Watershed Attributes tool provides estimates of the true watershed attribute. The watershed area, shown in orange, follows hydrologic boundaries in the upper RCAs, but appears as a straight line in the lowest RCA. This is because we simply estimate the percent coverage in the lowest RCA based on the ratio value rather than delineating the true boundary.

1. Go to STARS>Calculate>Watershed Attributes. This will open the Watershed Attributes window.

2. Set the argument values
   1. Watershed attributes may be calculated for multiple site feature classes simultaneously (i.e. observed and prediction sites)
   2. The Edge Watershed Attribute Name must be an accumulated field found in the edges attribute table
   3. The Edge RCA Attribute Name should be the RCA field that was accumulated to produce the Edge Watershed Attribute Name field
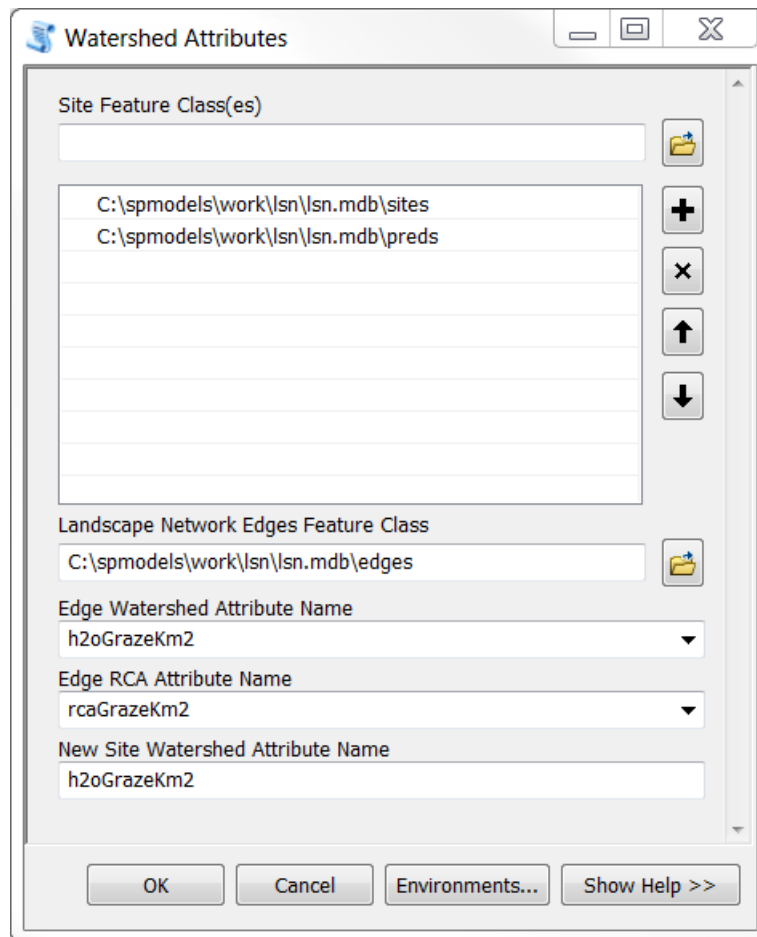
3. Click OK.

If the program finishes without errors, a green message will appear: Program finished successfully. A new field will be added to the sites attribute table that contains the watershed attribute for each site. If the tool is run separately for multiple sites feature classes (i.e. observed and then later prediction sites), ensure that the New Site Watershed Attribute Name is identical in all of the feature classes.

This field can be easily converted to % grazing in the watershed. Simply divide it by the total watershed area at the site and then multiply it by 100 to get the percentage.

*Note: Try running the tool again if an error message appears that says "The program DID NOT finish successfully" the first time you run it. Otherwise, go to the Troubleshooting section.

16. CALCULATE UPSTREAM DISTANCE

There are two tools provided in the STARS toolset to calculate the upstream distance between the stream outlet (i.e., the most downstream location in the stream network) and each of the edges and sites: Upstream Distance – Edges and Upstream Distance – Sites. The Calculate>Upstream Distance - Edges tool calculates the total distance from the uppermost location on each line segment (upstream node) to the stream outlet (i.e., the most downstream location in the stream network) and records it in the edges attribute table. The Calculate>Upstream Distance - Sites tool calculates the total distance from each site location to the stream outlet and records it in the sites attribute table. The new attributes in both the edges and sites attribute tables have the same units as the edges Shape_Length attribute. These attribute values provide part of the information needed to calculate flow-connected and flow-unconnected hydrologic distance measures in R.

1. Calculate Upstream Distance - Edges
    1. Double click on STARS>Calculate>Upstream Distance - Edges. This will open the Upstream Distance – Edges window.

2. Set the Edges Feature Class argument (below), select the Shape_Length attribute from the dropdown list, and click OK.
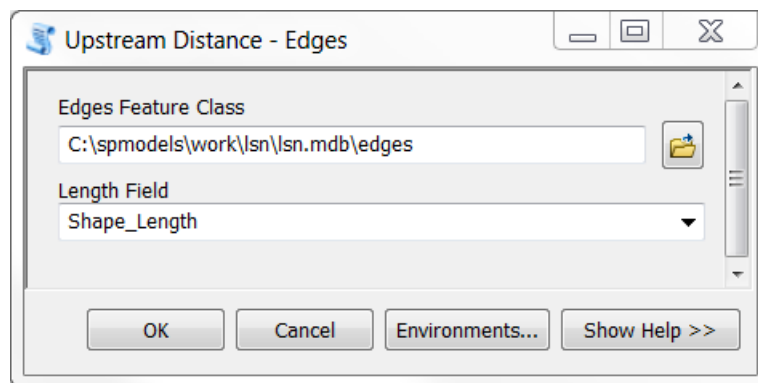
You should see a green message Program finished successfully. In addition, a new field should appear in the edges attribute table named *upDist*.
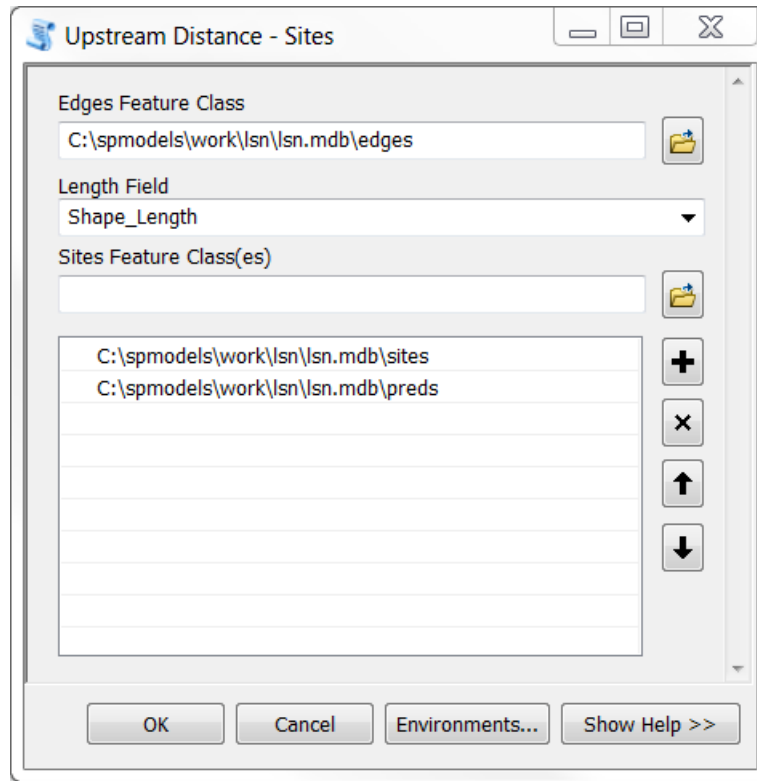
3. Calculate Upstream Distance – Sites
    1. Double click on STARS>Calculate>Upstream Distance - Sites. This will open the Upstream Distance – Sites window.

    2. Set the arguments (below) and click OK.

If no errors occurred, a green message, Program finished successfully, will appear.

*Note: The upDist attribute must be present in the edges attribute table before the Upstream Distance – Sites tool can run.

17. CALCULATE SEGMENT PROPORTIONAL INFLUENCE

Calculating the spatial weights needed to fit a spatial statistical model to streams data is a three step process: 1) calculating the segment proportional influence (PI), 2) calculating the additive function values, and 3) calculating the spatial weights (Peterson and Ver Hoef 2010). Steps 1 and 2 are performed using the STARS toolset, while step 3 is undertaken in R using the SSN (Spatial Stream Networks) package. The segment PI is defined as the relative influence that a stream segment has on the segment directly downstream (Figure 9). In this example, the segment PI is based on watershed area, but other measures (i.e. slope or Shreve's stream order) could also be used.

To begin, watershed area is calculated for the downstream node of each edge segment in the network, W_j (refer to Section 12). The cumulative watershed area at each confluence, pseudo, or outlet node (Figure 1a), $\sum_{k=1}^{n} W_k$, is calculated by summing the watershed area for the $n \leq 2$ edges that flow into it. The PI for each edge that flows into the node, ω_j, is then

$$\omega_j = \frac{W_j}{\sum_{k=1}^{n} W_k}$$

The segment PIs directly upstream from a confluence always sum to 1 because they are proportions.

$$\omega_A = \frac{W_A}{W_A + W_B}$$
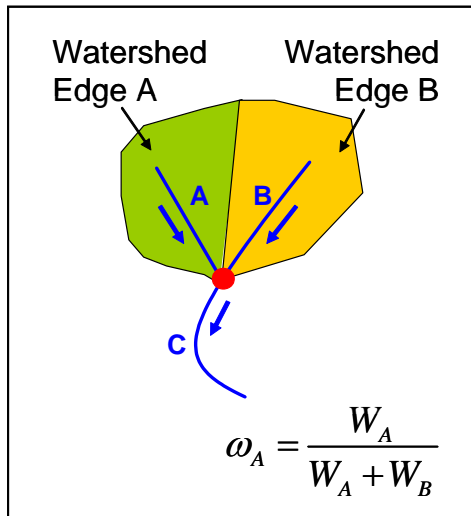
Figure 9. Calculating the segment proportional influence (PI).

1. Calculate the watershed area for each edge in the LSN using the STARS> Calculate>Accumulate Values Downstream tool (see section 12 for instructions). Set 'Field to Accumulate' to rcaAreaKm2 and name the new field 'h2oAreaKm2'.

2. Double click on the STARS>Calculate>Segment PI

3. Set the arguments (below) and click OK.
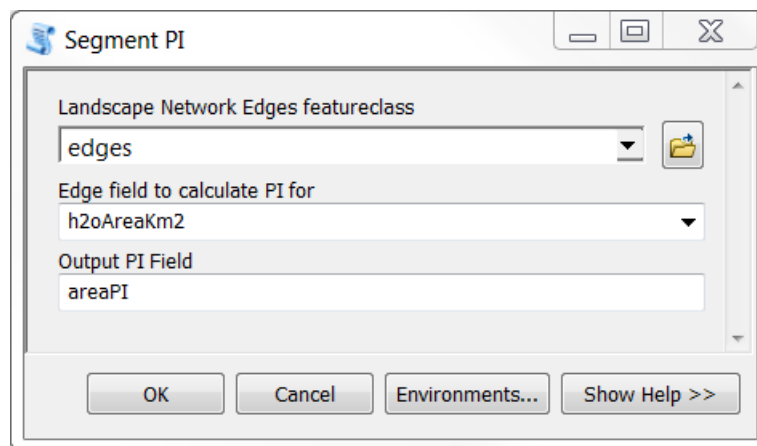
A green message will appear, Program finished successfully, if it runs without errors.

This tool adds a new field to the edges attribute table that contains the segment PI values. In this example, it is named areaPI.

The segment PI values are proportions ($0 \leq$ segment PI ≤ 1). Though, PI values equal to zero will occur if the segment attribute used to calculate the PI is equal to 0. If any PI values are greater than 1, then an error has occurred and the segment PIs should be recalculated.

4. Check the areaPI field to ensure that values range between 0 and 1.

    1. Open the edges attribute table, right click on the areaPI field, scroll down and select Statistics. This will open the Statistics of edges window, which contains summary statistics for the areaPI field.

    2. Examine the minimum and maximum values to ensure that they range between 0 and 1.

18. CALCULATE ADDITIVE FUNCTION

Two tools have been provided in the STARS toolset, which are used to calculate the additive function value (AFV) for every edge and site in the LSN (Figure 10): Additive Function – Edges and Additive Function – Sites. The AFV for a given edge, j, is equal to the product of the segment PIs found in the path downstream to the stream outlet.

$$AFV_j = \prod_{m=1}^{n} \omega_{D_{j,m}}$$

Note that, the set of edges in the downstream path to the stream outlet, $D_{[j]}$, also includes the jth edge. The AFV for the outlet segment (Figure 10, edge G) is equal to 1. The AFV for a site, AFV_i, is simply equal to the AFV_j of the edge it lies on. For additional details, please see (Peterson and Ver Hoef 2010, Appendix A).

$$D_{[B]} = [B, E, G] = [D_{B,1}, D_{B,2}, D_{B,3}]$$

$$AFV_B = \prod_{m=1}^{n} \omega_{D_{B,m}} = \omega_B \times \omega_E \times \omega_G$$

$$AFV_1 = AFV_B$$
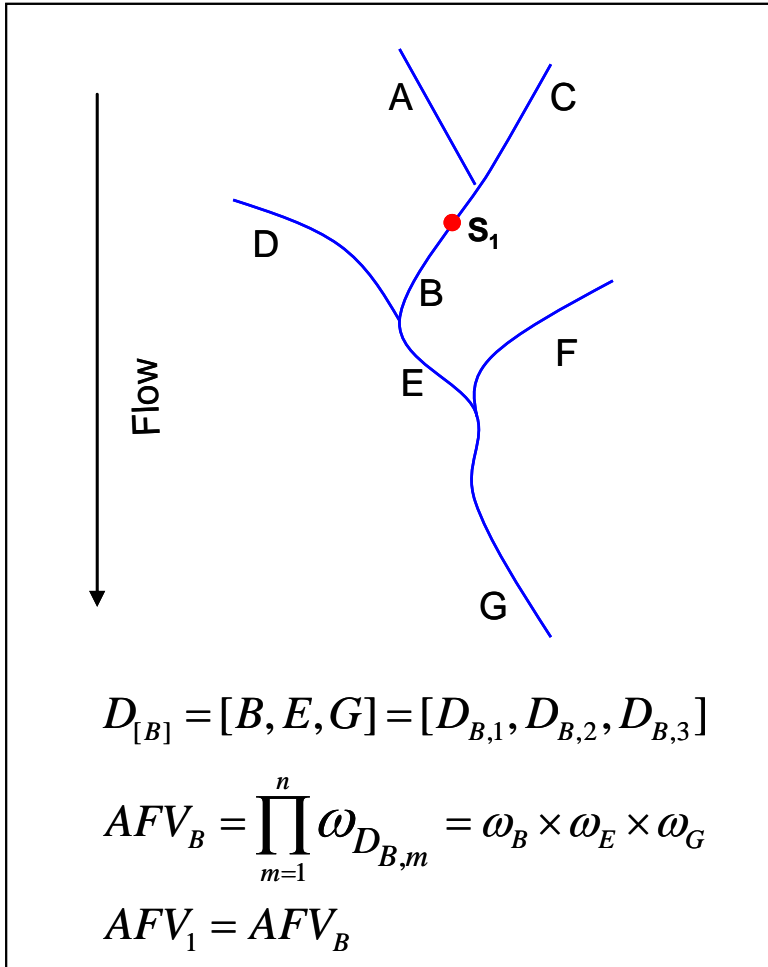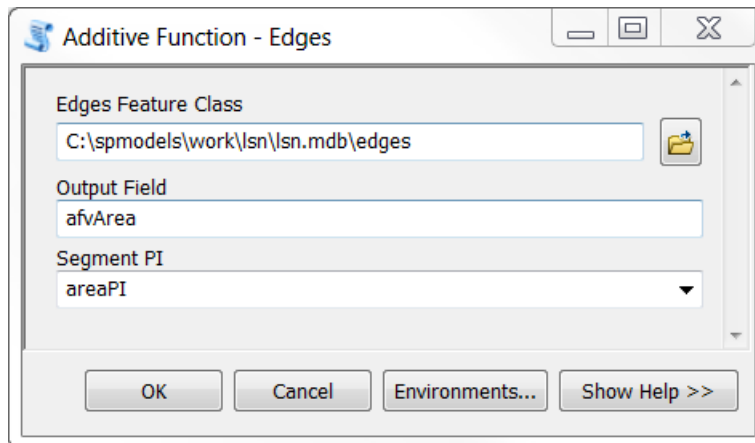
Figure 10. Calculating the additive function values (AFV) for every site and edge in the landscape network. Note that the ith site, in the example above, can be located anywhere on segment B.

1. Calculate Additive Function – Edges
   1. Go to STARS>Calculate>Additive Function - Edges. This will open the Calculate Additive Function - Edges tool.
   2. Set the arguments and click OK.

If the script runs without errors a green message will appear: Finished Get Additive Function Script.

2. Calculate Additive Function - Sites
   1. Go to STARS>Calculate>Additive Function - Sites. This will open the Calculate Additive Function - Sites tool.

   2. Set the arguments and click OK.
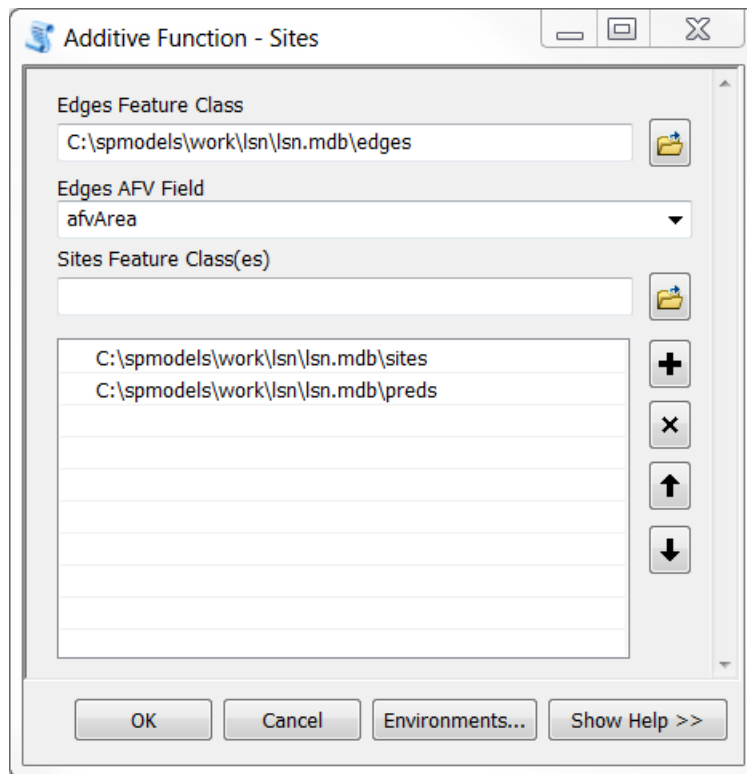
When the tool finishes running, a green message should appear: Finished Additive Function Script.

The Additive Function tools create new fields in both the sites and edges attribute table representing the AFV values. The AFV is a product of proportions (segment PI values) and so the AFV should always range between 0 and 1. Check to ensure that this is the case.

19. CREATE SSN OBJECT

The purpose of the Create SSN Object tool is to reformat the LSN as a Spatial Stream Network (.ssn) object. The .ssn object represents the spatial data and the topology of the network in a format that can be easily accessed and efficiently stored and analysed in R statistical software using the SSN package.

In a LSN, the relationships, nodexy, and noderelationships tables provide an efficient way to store and quickly analyse topological relationships between edges, sites, and nodes in the network. This information is powerful because it allows the spatial relationship between any two locations to be analysed, accounting for connectivity, flow direction, and distance (see Theobald

et al. (2006) or Peterson and Ver Hoef (2014) for a detailed description of the relationships tables). The LSN is currently stored as a personal geodatabase in ArcGIS, which is a Microsoft Access file (Figure 11). Unfortunately, it is difficult to access the feature geometry of feature classes in R when they are stored in this format. As an alternative, we have chosen to create the .ssn object, which is used to store the feature geometry, attribute data, and topological relationships of each spatial dataset contained in the LSN in a format that can be efficiently stored, accessed, and analysed in R.
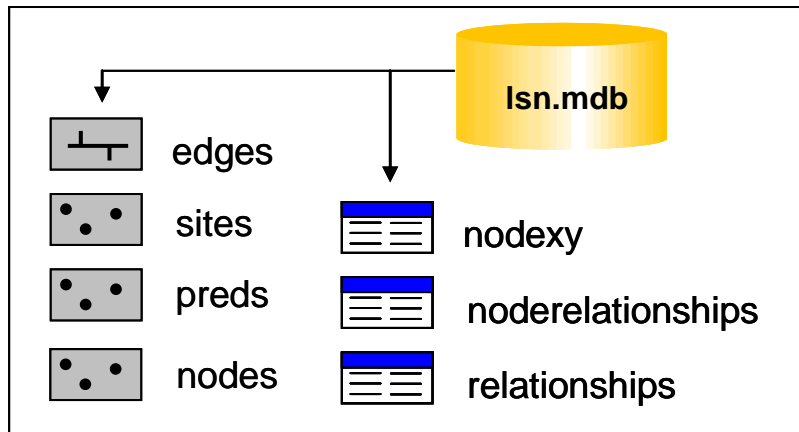

Figure 11. A landscape network (LSN) must contain six datasets before it can be used to calculate the data needed to fit the spatial statistical models: three feature classes: edges, nodes, and sites, as well as, three Access tables: nodexy, noderelationships, and relationships. Additional feature classes representing prediction locations may also be included.

The spatial datasets are stored as shapefiles in the .ssn object, since shapefiles can be easily imported into R, with all of the associated shape geometry and attributes. However, shapefiles of the edges and sites in their present form cannot be used to represent the topological relationships in the LSN. Our solution was to use network and binary identifiers (IDs).

Before the binary IDs are calculated, a new folder is created to hold the files that make up the .ssn object. The folder is located in the same directory as the lsn.mdb and the naming convention is *lsn name*.ssn (e.g. lsn.ssn).

The process of assigning binary IDs is relatively straightforward. The outlet edge is identified for a network in the LSN and assigned a binary ID = 1 (Figure 12). The upstream node of one outlet edge represents the downstream node of the two edges directly upstream (Figure 12, black circles). Binary IDs are assigned to the two upstream edges by appending a 0 or a 1 to the downstream binary ID (i.e. 1 ⟵ 10 and 11, Figure 12). This process of moving upstream and assigning binary IDs continues until every edge in the stream network has been assigned a binary ID.

It is common for LSNs to contain multiple stream networks, with unique stream outlets. For instance, there are 16 individual stream networks in the edges feature class provided in the example. Two edges may have the same binary ID if they reside on different stream networks (Figure 12). Therefore, a network identifier (netID) is also assigned to the edges, sites, and prediction sites attribute tables to differentiate between two edges with the same binary ID. In addition, the observed and prediction sites are assigned a location ID (locID) and a point ID (pid). The locID for a record will only be unique if the dataset does not contain repeated

measurements at a single location (e.g. measurements taken over time at a single location). However, the pid for each record is always unique.
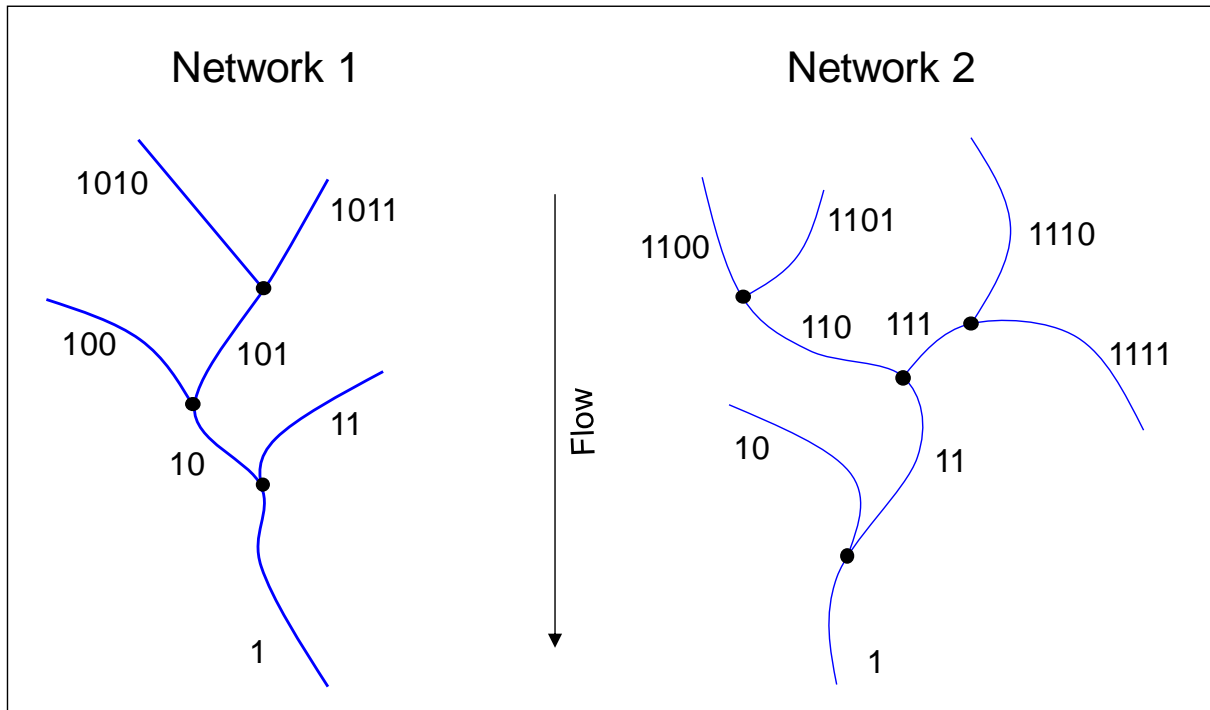

Figure 12. Binary IDs are assigned to each edge in the LSN. Edges are represented by blue lines and nodes (black circles).

The rid and binary ID for each edge are stored in a comma delimited text file (Figure 13), with a separate binary ID file for each network. The naming convention for these files corresponds to the network ID (e.g., net1.dat, net2.dat, etc.). All files are stored in the lsn.ssn folder.

```
"rid", "binaryID"
 418,  1
 409,  10
 410,  11
 407,  100
 408,  101
 176,  110
 405,  111
 347,  1010
```

Figure 12. Binary ID text file format.

Once the binary, network, location, and point IDs have been calculated and assigned, the edges, sites, and prediction sites feature classes are converted to shapefiles, which are stored in the lsn.ssn folder. When the Create SSN Object tool is complete, the .ssn object contains the spatial, attribute, and topological information of the LSN (Figure 14).
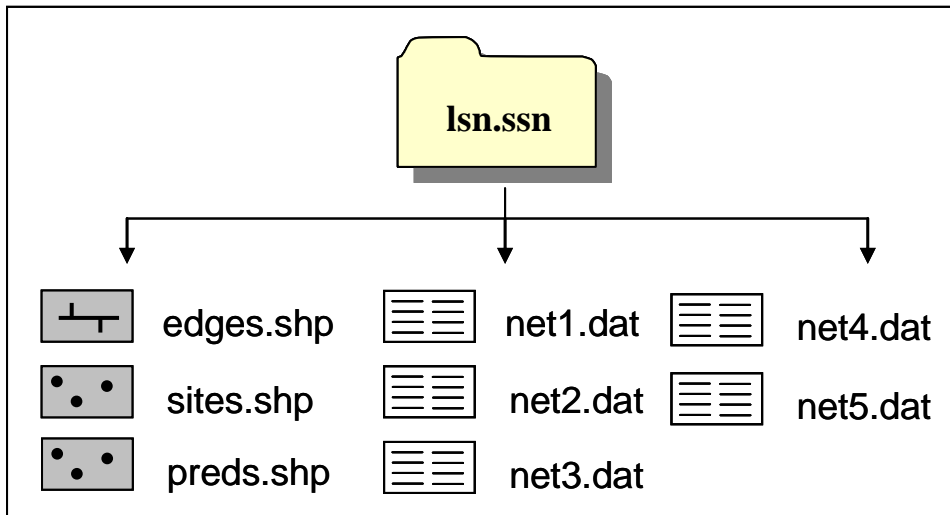
Figure 14. The .ssn object contains the spatial, attribute, and topological information of the LSN. It always contains at least two shapefiles edges and sites, as well as multiple text files containing the edge binary IDs.

To create an .ssn object:

1. Double click on STARS>Export>Create SSN Object. This will open the Create SSN Object tool.

2. Set the parameters (below) and click OK.
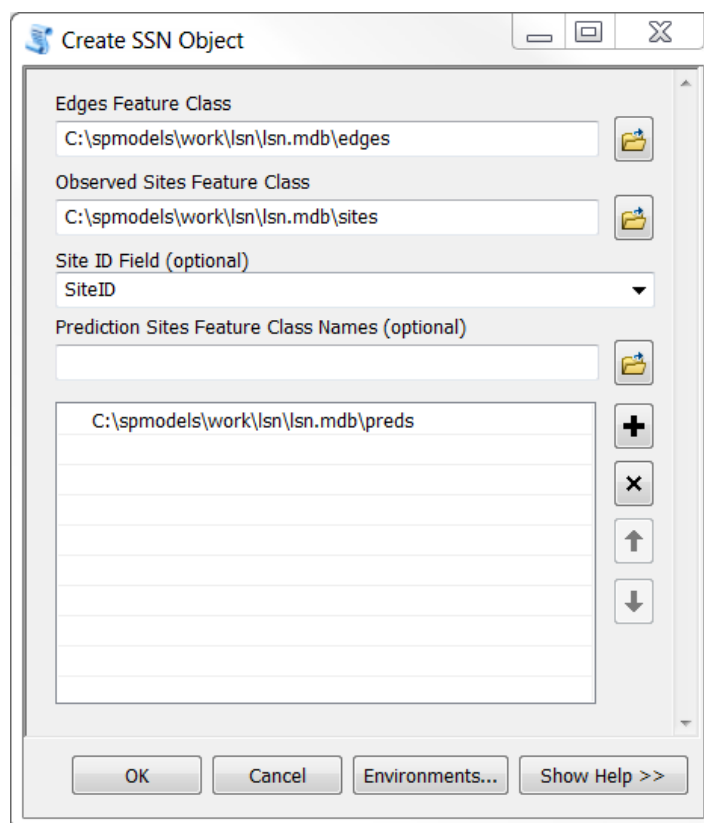
*Note that, the Site ID Field is optional, but must be specified if the Observed Sites Feature Class contains repeated measurements (multiple measurements at a single location). The Site ID Field should represent a site identifier, such as a site code, and **must be present and populated** in both the observed and prediction sites (if specified) feature classes. Site identifiers should be unique for each unique location.

The Create SSN Object tool may take a while to run depending on the number of edges and sites in the LSN. When the tool has finished successfully, a green message, Successfully Finished Create SSN Object Script, will appear. An .ssn object will also be created in the same directory as the LSN used to create it. It will have the same file structure as shown in Figure 14.

20. REFERENCES

Bureau of Meteorology (2012). \Australian Hydrological Geospatial Fabric (Geofabric) Product Guide." URL http://www.bom.gov.au/water/geofabric/documentation.shtml.

Horizon Systems Corporation (2007). \National Hydrography Dataset Plus: Documentation." URL http://www.horizon-systems.com/NHDPlus/NHDPlusV2_documentation.php.

Peterson E.E. and Ver Hoef J.M (2010) Capturing multiple patterns in stream ecosystems: A moving-average approach to geostatistical modelling. Ecology, 91(3): 644-651.

Peterson E.E. and Ver Hoef J.M. (2014) STARS: An ArcGIS toolset used to calculate the spatial data needed to fit spatial statistical models to stream network data. Journal of Statistical Software, 56(2), 1-17.

R Core Team (2017) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 30900051-07-0, URL http://www.R-project.org

Theobald D.M., Norman J., Peterson E.E., and Ferraz S. (2005) Functional Linkage of Watersheds and Streams (FLoWS): network-based ArcGIS tools to analyze freshwater ecosystems. Proceedings of the ESRI User Conference 2005, San Diego, CA, July 26.

Theobald D.M., Norman J.B., Peterson E., Ferraz S., Wade A., and Sherburne M.R. 2006. Functional Linkage of Water basins and Streams (FLoWS) v1 Users Guide: ArcGIS tools for Network-based analysis of freshwater ecosystems. Natural Resource Ecology Lab, Colorado State University, Fort Collins, CO. 43 pg.

Ver Hoef J.M. and Peterson E.E. (2010) A moving average approach for spatial statistical models of stream networks. Journal of the American Statistical Association, 489:22-24.

Ver Hoef J.M., Peterson, E.E., Clifford D., and Shah R. (2014) The SSN Package: An R package used to fit spatial statistical models to stream network data. Journal of Statistical Software, 56(3), 1-45.

21. TROUBLESHOOTING

21.1 General solutions for all tools:

1. Do not use file paths with spaces in them (e.g. "C:\Documents and Settings\pet221\My Documents". This includes the path to your data, as well as the path where the STARS tools are stored.

2. Check to make sure that you have a C:\temp directory because the tool writes temporary files to this folder. If you do not have a C drive, you must change this folder name in the python scripts that are associated with this tool.

   *Note: this will also be an issue with other tools.

3. Check the C:\temp directory. Data may have been written to C:\temp, but not deleted in a previous model run. Go to C:\temp and manually delete the output.

   1. Close and reopen all ArcGIS products.

   2. Restart the computer.

3. When other ArcGIS products, such as Arc Catalog, are open it can create a schema lock on individual datasets. This may prevent the tools from executing successfully. Close other ArcGIS products and rerun the tool.

21.2 Build a landscape network (LSN)

1. When they are incorporated into the LSN, the data are converted from a shapefile to a polyline feature class and given the name 'edges'. Errors in the LSN will occur if the original streams shapefile is named edges.shp. Rename edges.shp (i.e., streams.shp) and build the LSN.

2. It is not uncommon that the LSN fails to build and an error similar to this one is produced
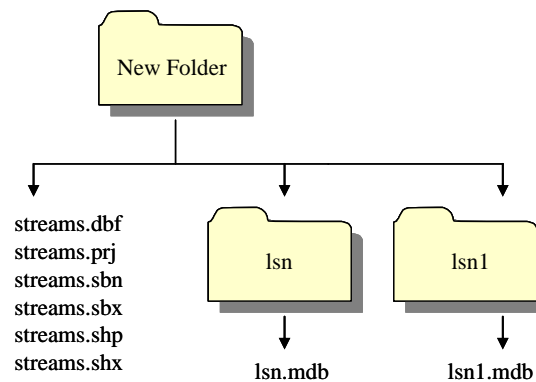
   ```
   Error in script PolylineToLandscapeNetwork.
   Error in executing: cmd.exe /C
   C:\PROGRA~1\GISTools\FLoWS\RCA200~2\POLYLI~1.PY
   "E:\projects\shortcourse\teachmat\data\nhd_alb3.shp"
   "E:\projects\shortcourse\lsn7.mdb"

   Failed to execute (PolylineToLandscapeNetwork_3).
   ```

   Potential work-arounds:

   1. Make sure that the shapefile does not have an OBJECTID field or an rid field. If it does, then delete those fields and try again.
      1. Create a new streams shapefile by exporting the data to a new folder.
      2. Try writing the LSN to a new directory folder.
      3. See General Solutions.

2. Create a new folder and export the streams shapefile to this folder. Make sure to create a new dataset (export) rather than simply copying the dataset. Create another folder within this folder (i.e., lsn) and generate the new LSN in this folder. Whenever a new LSN must be generated, create a new folder (i.e., lsn1) and generate the new LSN in this folder.

New Folder

streams.dbf
streams.prj
streams.sbn
streams.sbx
streams.shp
streams.shx

lsn

lsn1

lsn.mdb

lsn1.mdb

3. PolylineM and PolylineZ geometries pose difficulties for some of the STARS tools. NHDPlus hydrography layers, such as nhdflowlines commonly used in the United States, contain both M and Z coordinates. These may need to be removed from the shapefile prior to building the LSN.

4. The Polyline to Landscape Network tool fails to run and produces the error: Dataset nodes already exists

   Potential workarounds
   1. Sometimes the Polyline to Landscape Network tool will fail to run properly if there is a nodes featureclass from another LSN is present the table of contents. Remove the other nodes featureclass from the TOC and delete the relationship tables that were created as part of the failed tool run.

5. The Polyline to Landscape Network tool may fail to run properly if you are attempting to overwrite an existing LSN. Either delete the existing LSN or create a new directory and write the new LSN to the new location. Note that, if the tool failed to run previously it may have produced relationship tables, but not the full LSN. These relationship tables must also be deleted.

21.3 Check network topology

1. Program fails to execute
   1. Make sure that the edges feature class has a rid field. If not, recreate the LSN and rerun the Check Network Topology tool.
   2. The Check Network Topology tool stores temporary files in C:/temp. It is important to clear this folder before you re-run the Check Network Topology tool.

21.4 Creating RCAs

1. Cannot convert feature to raster
   1. Check the projection of your shapefile. Is it in the same projection as the DEM? If not, reproject your data so that it is in the same projection and try again.

2. Try a new name. Sometimes the error occurs if you try to use the same name for a raster that was previously created and deleted. If you want to use the same name, you can navigate to the working directory in Windows Explorer and delete the <rastername>.aux file. You will not be able to do this in ArcCatalog.

2. The Create Cost RCAs tool seems to have finished and successfully generated cost RCAS, but they are not in the working directory.

    1. The tool appears to have completed in an unrealistically short period of time and nothing was created. Check to make sure that the waterbody grid is an integer grid. If not, convert it to an integer, delete the files that were left in the C:\temp folder and try again.

3. Create Cost RCAs tool does not complete successfully.

    1. In the Create Cost RCAs tools window, the user does not need to include the .shp extension Output RCA Shapefile argument. If the extension is included, (i.e., costrcas.shp instead of costrcas), the program will add an additional extension to the filename (i.e. costrcas.shp.shp), which is not recognized as a shapefile format.

21.5 Export SSN

1. All of the locIDs are equal to zero.

    1. If a Site ID Field was specified, check to ensure that each site has a unique site identifier in both the observed and prediction feature classes.