Supplementary Information 3: Upstream Shade Buffers

for

Spatial and Temporal Variability in Stream Thermal Regime Drivers for Three River Networks During the Summer Growing Season

Matthew R. Fuller, Naomi E. Detenbeck, Peter Leinenbach, Rochelle Labiosa, and Daniel Isaak

Oak Ridge Institute for Science and Education Postdoc at the Atlantic Coastal Environmental Sciences Division (Fuller), U.S. Environmental Protection Agency, Narragansett, Rhode Island, USA [Currently: Northern Research Station, U.S. Forest Service, Amherst, Massachusetts, USA]; Atlantic Coastal Environmental Sciences Division (Detenbeck), U.S. Environmental Protection Agency, Narragansett, Rhode Island, USA; Region 10 (Leinenbach, Labiosa), U.S. Environmental Protection Agency, Seattle, Washington, USA; Rocky Mountain Research Station (Isaak), U.S. Forest Service, Boise, Idaho, USA (Correspondence to Fuller: matthew.fuller@usda.gov)

ABSTRACT: Many cold-water dependent aquatic organisms are experiencing habitat and population declines from increasing water temperatures. Identifying mechanisms which drive local and regional stream thermal regimes facilitates restoration at ecologically relevant scales. Stream temperatures vary spatially and temporally both within and among river basins. We developed a modeling process to identify statistical relationships between drivers of stream temperature and covariates representing landscape, climate, and management-related processes. The modeling process was tested in 3 study areas of the Pacific Northwest USA during the growing season (May [start], August [warmest], September [end]). Across all months and study systems, covariates with the highest relative importance represented the physical landscape (elevation [1st], catchment area [3rd], main channel slope [5th]) and climate covariates (mean monthly air temperature [2nd] and discharge [4th]). Two management covariates (ground water use [6th] and riparian shade [7th]) also had high relative importance. Across the growing season (for all basins) local reach slope had high relative importance in May, but transitioned to a regional main channel slope covariate in August and September. This modeling process identified regionally similar and locally unique relationships among drivers of stream temperature. High relative importance of management-related covariates suggested potential restoration actions for each system.

KEYWORDS: covariate relative importance; information criterion; model selection; Pacific Northwest, USA; restoration; river networks; spatial stream network model; temperature; water resources management; water quality

Upstream Shade Buffers: As-the-fish-swims

By Matthew Fuller

Last Updated: 15 April 2020

The following protocol uses the ArcGIS Network Analysis tools to generate upstream buffers using either stream distance or travel time to delineate the buffer from a set of identified points along the river network upstream. The primary data required for this analysis is a completed SSN object or stream network with flowlines that have pre-established connectivity (e.g., National Stream Internet database for CONUS extent), a set of locations (points) along the flow network that need upstream buffers to be calculated, and the flowline attributes from NHDPlus data used to join to the flowlines in your network.

The output from the buffer analysis/calculation are polygons that surround the stream flowlines by ~50m on either side. These buffers can be used to average values from overlapping raster datasets (e.g., shade estimates, slope, and elevation) to estimate upstream statistics for different parameters. Because the buffers are nested and overlap, GME tools are used to make these polygon-raster calculations, but other means are possible using ArcGIS or R.

The example data used in the below protocol for display purposes uses an established SSN object from the NorWeST project for the Middle Columbia River Processing Unit and the observation and prediction points for temperature in that SSN object for the Middle Fork John Day River (nested within the Middle Columbia Processing Unit). Upstream buffers are established for several stream length distances (1, 2, 3, 4, 5, and 10km) upstream as well as various travel times (1, 2, 3, 4, 5, 6, 12, and 24 hours) upstream of each observation and prediction point. A shade raster is used to demonstrate how to use the buffer polygons to extract mean upstream shade conditions for each buffer.

PART 1: Delineating buffers upstream of points “as the salmon swims” using distance or travel time upstream as a buffer delimiter.

1. Begin by loading data into ArcMap and ensuring the layers have similar projections.
    a. Locate the NorWeST SSN object for the area of interest. Within the processing unit directory is an “edges.shp” file. This file will be used to generate the Network Dataset in ArcMap.
    b. Add the edges.shp file to the map layers, then project (Data Management > Projections and Transform > Project) it into the NAD_1983_Albers projection. This projection matches the observation and prediction site shapefiles.
    c. Add the observation and prediction points to the map. Confirm these are in the same projection as the edges.

2. Prepare the edges data for generating a network dataset. Three fields need to be added/created in the edges attribute table before creating a network dataset. A (1) “Length” field for determining upstream buffers as a function of stream distance, (2) “Oneway” field to restrict the direction of flow from the upstream to downstream

direction, and (3) “TravelTime” field to use as a measure for upstream buffers via water travel time.

a. Confirm there is a “Length” field in the edges shapefile. This field will be automatically recognized by ArcGIS as a field to use for identifying impedance breaks for a service area analysis. The units should be either in meters of kilometers. Either one can be set accordingly from the network dataset analysis properties settings, but noted which units they are for future network dataset analysis layer property settings.
b. When establishing flow directionality in the network, we need to set up one-way edges in the stream network. To do this, you need to create a “Oneway” text field (Length: 2) in the edge attribute table. This allows us to restrict movement in the upstream direction when desired. In this text field, place “FT” in the cells for all flowline features using the Calculate Field tool. This designates a one-way flowline in the direction the flowline is digitized (“From” node to the “To” node associated with each line feature).
c. Several steps are required to establish a “TravelTime” field that include, (1) obtaining water velocity data for the travel time estimates, (2) joining them to the edges, and (3) using the length and mean annual velocity estimates to calculate an approximate travel time for each flowline.
    i. The NHDPlus V1 attributes (NHDV1 used because this was the flowline data used to generate the SSN object for the Middle Columbia River Processing Unit) contain estimates of mean annual flow velocity for each NHD flowline. This flowline attribute data can be joined to the edges data from the SSN object via the common “COMID” field.
        1. Load the attribute database file (“flowlineattributesflow.dbf”) to the map.
        2. Right click the edges shapefile and select Joins and Relates > Join…
        3. Within the Join Data window, choose “Join attributes from a table”
            a. Choose “COMID” for a field
            b. Select the “flowlineattributesflow.dbf” file to join to
            c. Choose “COMID” for a field
            d. Select the radio button for “Keep all records” as a Join Option
            e. Click OK to start the join.
    ii. After the join, locate the “MAVelU” field. This field contains the “Mean Annual Velocity (fps) at bottom of flowline as computed by Jobson Method (1996) using the flow in MAFlowU.” This field measures velocity in “feet per second”, so we will also need to convert it to metric units of “meters per second”.
    iii. Add a new field to the attribute table and call it “MAVelU_m”
        1. Type: double
        2. Precision: 10

3. Scale: 8

iv. Use the “Calculate Field” tool to convert feet per second to meters per second using the following conversion: 1 foot = 0.3048 meters

v. Next we need to estimate the time it requires to travel along each flowline using the mean annual velocity and length of each flowline.

1. Add a new field to the attribute table and call it “TravelTime”

a. Type: Double
b. Precision: 15
c. Scale: 9

2. Use the “Calculate Field” tool to estimate the travel time using the velocity data (MAVelU_m) and the flowline length data (Length) via the following equation:

TravelTime = ((LENGTHKM * 1000) / MAVelU_m) / 3600

3. The above equation for TravelTime has units of Hours. Adjust the “3600” denominator to get minutes (60) or seconds (1) if these units are necessary or preferable.

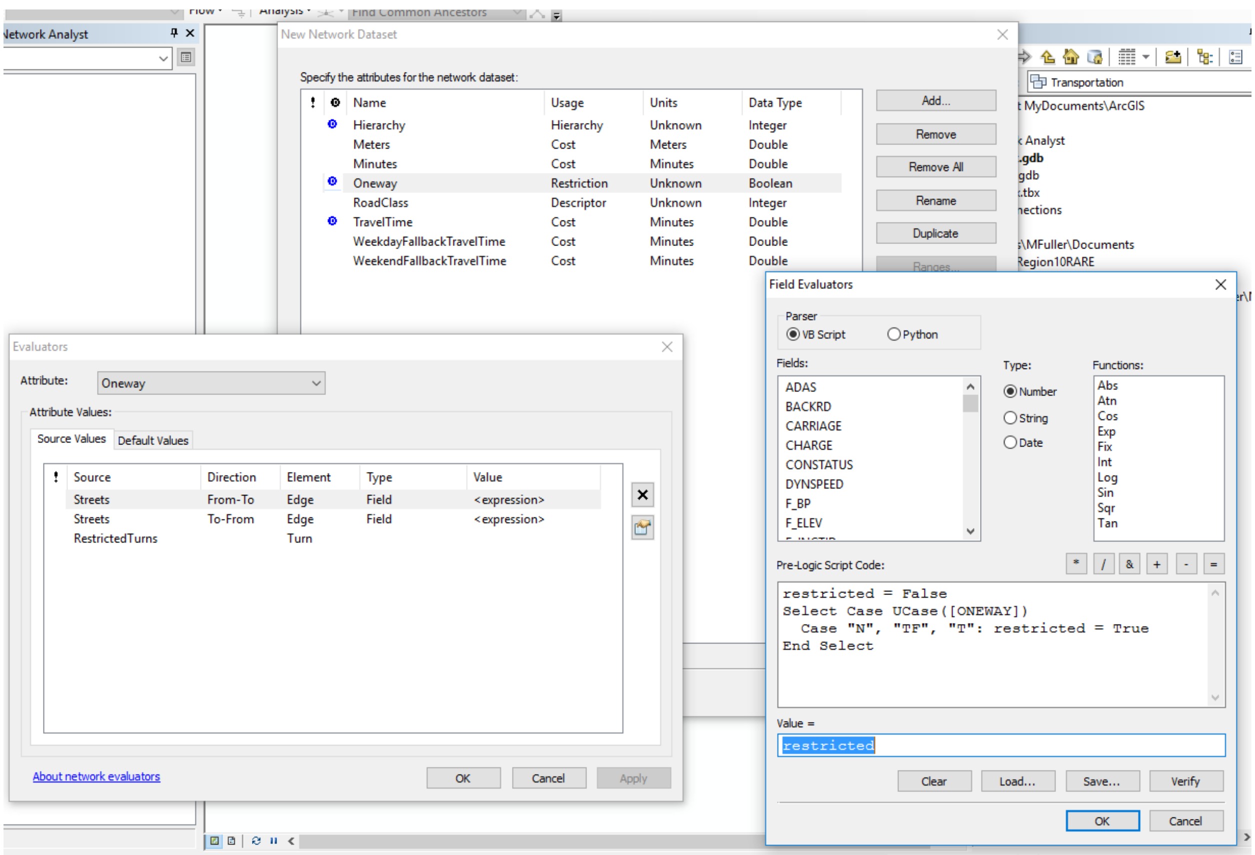
3. Prepare databases to hold the network dataset

a. First create a new file geodatabase.
   i. In ArcCatalog, right click the directory you want to hold the file geodatabase and select New > File Geodatabase. Name it accordingly.
b. Create a new Feature Dataset within the new geodatabase
   i. Right click on the geodatabase and select New > Feature Dataset. Name accordingly.
   ii. Choose the projection of the rest of the data in your map (NAD_1983_Albers). You can find it in the “Layers” projections directory.
   iii. Select no coordinate system for the vertical coordinate system.
   iv. You can leave the defaults for X, Y, Z, and M tolerance and leave the box checked for “Accept default resolution and domain extent”.
   v. Click Finish.
c. Next, import the features to your feature dataset.
   i. Right click the new feature dataset and select Import > Feature Class (Multiple) …
      1. Add the edges alone as Input Features
      2. leave the output geodatabase as the newly created feature dataset
      3. Click OK

4. Build the network dataset

a. Right click on the new Feature Dataset and select New > Network Dataset …
b. Name the network dataset, choose ArcGIS version 10.1, and click Next
c. Select only the edge feature class and click Next

d. Set turns in the model to “No”, click Next
e. Leave default connectivity for the line features. Click Next
f. We won’t be using elevation in the network features so we select “None” for how to address elevation in the model.
g. There should be three attributes for the network dataset
    i. “Length”
        1. Usage: Cost
        2. Units: Meters
        3. Data Type: Double
    ii. “Oneway”
        1. Usage: Restriction
        2. Units: unknown
        3. Data Type: Boolean
        4. The evaluator for the Oneway attribute should be as in the below screenshot. The only difference in the script code between the “From To” and “To From” Field Evaluators script code is in line three where the “TF” is reversed to “FT”.
        5. Usually when ArcGIS reads an attribute with the name “Oneway” it assumes this information and populates the script codes itself. It is good to check though.



    iii. “TravelTime”
        1. Usage: cost
        2. Units: hours

3. Data Type: Double
4. Usually when ArcGIS reads an attribute with the name "TravelTime" it assumes the units are in "Minutes", so if you calculated the values in hours (Section 2.c above), be sure to update this unit information here.

iv. If you had more than one travel time field and consequently they are named something other than "TravelTime", then ArcGIS wont automatically recognize these travel time fields. Instead, you need to add them manually here.
1. Click "Add…" button
2. Type in the name of the field holding your travel time data.
3. Select the appropriate unit for travel time from the drop down list

v. Click Next

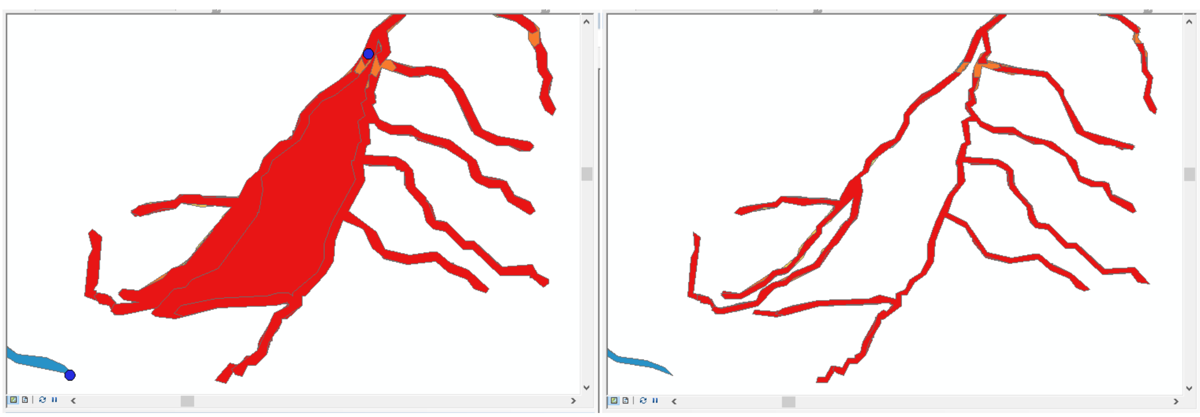
h. Leave travel mode fields blank
i. We will not establish any driving routes, so we can select "No" for directions.
j. We will be conducting service area analyses, so we will select the box to build a service area index.
k. Then finish. When it is complete, you should choose to build the network from the pop-up message/window. Building the network took a long time (~10 min) for the MFJD because it uses the entire MidColumbia Processing Unit SSN object edges file. This will likely take several minutes for any large SSN object edges file.

5. From the Network Analyst toolbar use the drop down list to start a "New Service Area" analysis. You may need to open this tool bar by going to the top menu item "Customize > toolbars > Network Analyst" and selecting it to use in ArcMap.

6. Load the Facilities locations for a Service Area analysis. The facilities, in this case, are designated by the point files for observations and predictions across the network. These locations are where we need to map upstream buffers.

a. From the Network Analyst toolbar, select the drop down menu from Network Analyst and select a New Service Area analysis
b. First load facilities for the analysis. These are the observation and prediction points.
i. Right click "Facilities (0)" and select "Load locations…"
ii. Load the point features that you want to use as the facilities to search upstream from in the "Load From:" field
iii. Choose a field for the Location Analysis Properties Name (often a point ID or some other feature ID to track each feature. I've created a new field for each points file called "FacilityID" – Long integer). This step is important so that the polygons can be joined via this name field back to the point file for the facilities. Make a note of which field you use for naming the polygons for future reference.
iv. Set the search radius for 30 meters to snap facilities to the network edges since they may be slightly offset or not associated via a "Snap" procedure to the line edges. This distance can be adjusted to the data set and how far offset the observation points may be from the network edges/flowlines.

7. Set analysis properties.
    a. Polygon Generation Tab
        i. Generate Polygons: box checked
        ii. “Generalized” radio button selected (not “Detailed”)
        iii. Polygon Type: Generalized
            1. Trim Polygons: box checked; 50 meters for length input
                a. ***Note for end of analysis*** If your buffer polygons, at the end of the service area analysis, are merging because streams are very close to each other and causing the stream tributaries to merge and generate polygons covering upland area or other tributaries that should not be included, adjust this buffer trim down. For the Wind River This required reducing the trim distance to 10 meters. See the examples below with 50m trim on the left and 10m trim on the right



        iv. Excluded sources: boxes unchecked
        v. Multiple Facilities Options: Overlapping
        vi. Overlap Type: Disks
    b. Line Generation Tab
        i. Generate Lines: box not checked
    c. Accumulation Tab
        i. Accumulation Attributes: neither Length (Meters) or Minutes (Hours) box checked
    d. Network Locations Tab
        i. Network Location Field Mapping
            1. Location Type: Facilities
            2. Do not need to change any of the default values.
        ii. Finding Network Locations:
            1. Search Tolerance: 30 meters
            2. Point Barrier Location Type set “BarrierType” to “Added Cost”
    e. Analysis Settings Tab ***CHANGE TO TOGGLE DISTANCE (“i” and “ii”) OR TRAVEL TIME (“iii” and “iv”) BUFFERS***
        i. Impedance: Length (Meters)
        ii. Default Breaks: 1000,2000,3000,4000,5000,10000

            1. These values state the number of meters to look upstream.
    iii. Impedance: TravelTime (Hours)
    iv. Default Breaks: 1,2,3,4,5,6,12,24,1000 (1000 hours is the full watershed assuming it takes less than 1000 hours to reach all the tips of the headwater tributaries.)
            1. These values state the travel time in hours to look upstream.
    v. Use Time: unselected
    vi. Direction: Towards Facility
    vii. U-Turns at Junctions: Allowed
    viii. Ignore Invalid Locations: box checked
    ix. Restrictions: “Oneway” box checked

8. Solve the Service Area analysis

    a. Click the Network Analyst Toolbar Solve Button  to run the service area analysis.

    b. When the analysis has completed, polygons should be generated that reach upstream from the observation/prediction points according to the breaks set for impedance (either distance or travel time) in the analysis properties (Section 6.e.i-iv).

PART 2: Once buffer polygons are generated for the observation and prediction sites, average values within these buffers can be calculated using the GME isectpolyrst tool. This GME tool requires that the data you are averaging along the upstream buffer be in raster format of the same projection as the polygons. (Note: spatial joins between shade points and the polygons will not work because these spatial joins are one-to-one and not the one-to-many joins which are required for nested polygons).

We will be calculating average shade values for the nested upstream buffer polygons.

9. Use GME to extract summarized raster values for each polygon.

    a. Inside GME, select the “isectpolyrst” tool. There are several parameters to set to run this tool for overlapping polygons that need to extract raster data.

        i. Select a shade buffer polygon layer as the “in” file.
        ii. Select the desired shade raster (either current or potential/future” for the “raster” input.
        iii. Select a six character or less prefix for the output field. For current shade, I used “rwCSFD” as a representation of the “raw Current Shade as the Fish swims by Distance upstream” and “rwCSFT” for “raw Current Shade as the Fish swims by travel Time upstream”. For the future/potential shade, I replaced the “C” in the prefix with an “F” for “rwFSFD” and “rwFSFT”.
        iv. Optional: Metrics: entered “MN” to get the mean shade value within each polygon.
        v. Optional: Thematic, Proportion, and Medquant were both set to False

vi. Optional: allowpartialoverlap was set to True
vii. Optional: Where was left blank.

b. Click the "Copy to clipboard" button and keep a text file record of the commands that are run along with their approximate start and stop times.

c. Click the "Copy to Command Text window" button

i. This will bring you to a new tab within the Commands tab window.
ii. Click the "Run" button to run the analysis.

d. Clicking the Run from the "Command Text" tab opens up the "Output" tab immediately. Here you will find a progress bar in the lower left corner of the window to track your analysis.

e. The output from the analysis exists within the buffer polygon file as a new field titled "rwCSFDMN" and should be renamed. This can be done by adding a new field of the same time and then transferring the values to the new field using Field Calculator.

10. Convert raster values to the correct format. This is necessary because the shade raster has values that range from 0 to 1000 where 0 indicates 0.0% shade and 1000 indicates 100.0% shade. Therefore, it is necessary to divide each value in the "rwCSFDMN" column by 1000 to generate values that range from 0 to 1 and identify the average proportion of shade cover within each buffer

f. Add a field to the polygon file and format/name it accordingly.

i. Name: "mnCSFD" (substitute the prefix form from the input polygon file from section 8.a.iii)
ii. Type: Double
iii. Precision: 10; Scale: 9

g. Use the field calculator to divide the "rwCSFDMN" column by 1000.

h. This field is the final output for average shade along each buffer distance/travel time upstream of each observation/prediction point.