Supplementary Information

IDW-Plus: An ArcGIS toolset for calculating spatially explicit watershed attributes

Erin E. Peterson[1] and Alan R. Pearse[2]

[1] Corresponding Author: Erin.Peterson@qut.edu.au; Principal Research Fellow, ARC Centre for Excellence in Mathematical & Statistical Frontiers and the Institute for Future Environments, Queensland University of Technology, Brisbane, QLD, Australia, 4001.
[2] Research Assistant, ARC Centre for Excellence in Mathematical & Statistical Frontiers (ACEMS), Queensland University of Technology, Brisbane, QLD, Australia, 4001.

Table of Contents

1 Introduction 3

2. Software & Hardware Requirements 3

3. The IDW-Plus Toolset Structure 4

4. Data Requirements 4

4.1 Example Dataset Description 5

5. Getting Started 7

5.1 Tutorial Directory Structure 9

6. Pre-processing Steps 9

6.1 General Pre-processing Steps 10

6.2 Creating the Stream Raster 11

6.3 Formatting the DEM 14

6.3.1 Burning in the streams 14

6.3.2 Filling the DEM 16

6.3.3 Creating the Flow Direction and Flow Accumulation Grids 17

7. Using the IDW-Plus tools 19

7.1 Tool Inputs and Arguments 21

7.1.1 Field name 22

7.1.2 Inverse distance weighting power 23

7.1.3 Snap tolerance 23

7.1.4 WHERE clause 24

7.2 The iFLO Tool 25

7.3 The HAiFLO Tool 26

7.4 The iEDO Tool 27

7.5 The iFLS Tool 28

7.6 The HAiFLS Tool 29

7.7 The iEDS Tool 30

8. Output 31

9. References 32

1 Introduction

This tutorial has been written to help users of the Inverse Distance Weighted Percent Land Use for Streams (IDW-Plus) tools recreate the example in Peterson and Pearse (In Review). Additional resources may be helpful such as the ArcGIS version 10.3x help pages and the ESRI User Forums (http://forums.arcgis.com/). The last two references will not provide specific information about the IDW-Plus toolset, but may provide help with more general ArcGIS tasks.

2. Software & Hardware Requirements

The following software is required for the toolbox:

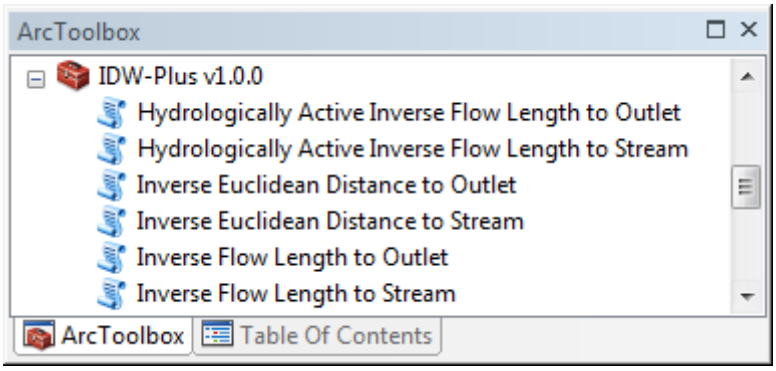
1. An Advanced licence for ArcGIS 10.3x.
2. A licence for the Spatial Analyst extension.
3. Python 2.7.8 (automatically installed with ArcGIS 10.3x).

The tools generate intermediate data when calculating the metrics. These data are stored in memory, which considerably improves execution speed. However, this can present problems if the input datasets are large. To avoid complications with in-memory processing, it is recommended that at least 4GB of RAM are installed on your computer.

3. The IDW-Plus Toolset Structure

IDW-Plus is an ArcGIS custom toolset for calculating a suite of spatially explicit land use metrics described in Peterson *et al.* (2011). It contains six tools:

1. Hydrologically Active Inverse Flow-Length to Outlet (HAiFLO)
2. Hydrologically Active Inverse Flow-Length to Stream (HAiFLS)
3. Inverse Euclidean-Distance to Outlet (iEDO)
4. Inverse Euclidean-Distance to Stream (iEDS)
5. Inverse Flow-Length to Outlet (iFLO)
6. Inverse Flow-Length to Stream (iFLS)



4. Data Requirements

At a minimum, five datasets are required to generate spatially explicit watershed attributes: survey sites, a digital elevation model (DEM), a flow direction raster, a flow accumulation raster, and a land use raster (Table 1). A streams raster is an optional input to the tools. Details about how these data are used are provided in the following sections.

Table 1. Spatial data requirements for the IDW-Plus tools.

Data	Format	Description	Required
Survey sites	Vector	Survey data collected at sites	Yes
DEM	Raster	Digital elevation model	Yes
Flow direction	Raster	Flow direction from the DEM	Yes
Flow accumulation	Raster	Flow accumulation from the DEM	Yes
Land use	Raster	Land use or landscape attribute	Yes
Streams	Raster	Streams	Optional

4.1 Example Dataset Description

The example dataset used in this tutorial can be downloaded at http://www.fs.fed.us/rm/boise/AWAE/projects/SSN_STARS/other_software.html. The data are from the South-East Queensland Ecosystem Health Monitoring Program (EHMP; Bunn *et al.* 2010). It includes five EHMP survey sites, as well as raster data for streams, elevation, and agricultural pasture land use at a 25 metre spatial resolution (BRS, 2002; Moreton Bay Waterways and Catchments Partnership, 2005; QNRW, 2000) (Figure 1). The intention is for users to work through the tutorial with the example data so they understand the tools and can successfully generate spatially explicit watershed attributes.

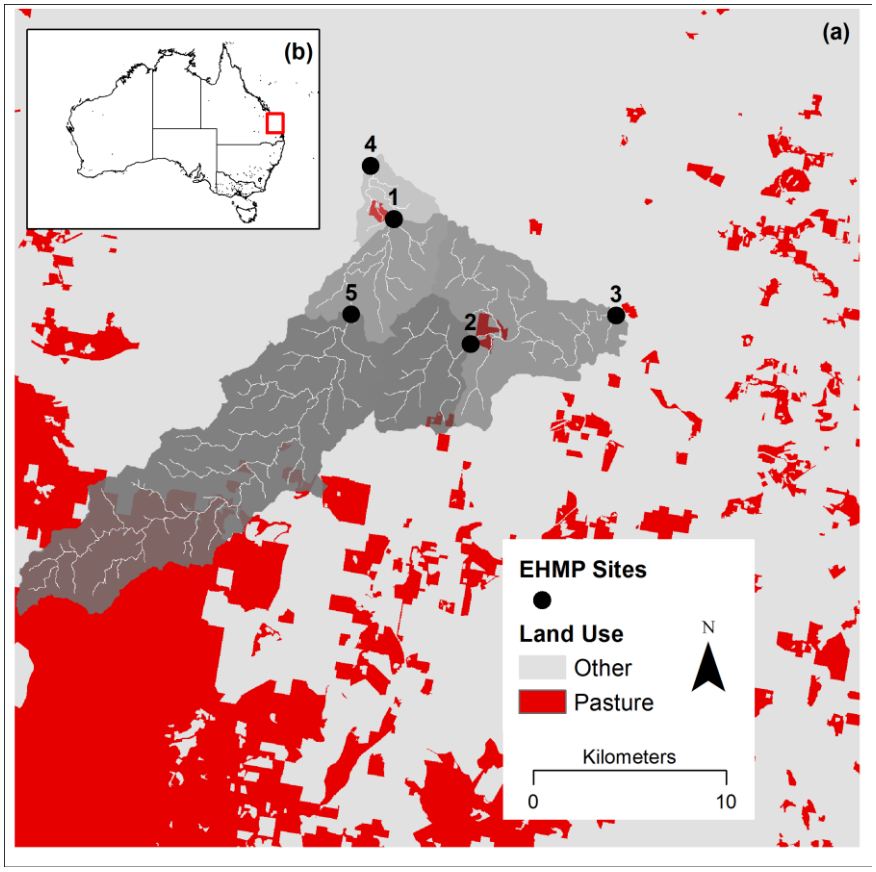


Figure 1. Map of the study area (a), which is located in Southeast Queensland, Australia. Black circles represent the 5 Ecosystem Health and Monitoring Program (EHMP) survey sites, which are labelled using the corresponding SiteID. Watersheds for each EHMP site are displayed in varying shades of grey. Lighter shades of grey represent larger watersheds within which the darker watersheds are nested. The streams running through the watersheds are shown in white.

The *example* folder contains the input data for the IDW-Plus tools. The *example_final* folder contains the data provided in the example folder, as well as the data derived in this tutorial (Table 2).

Table 2: Datasets provided in the *example* and *example_final* folders.

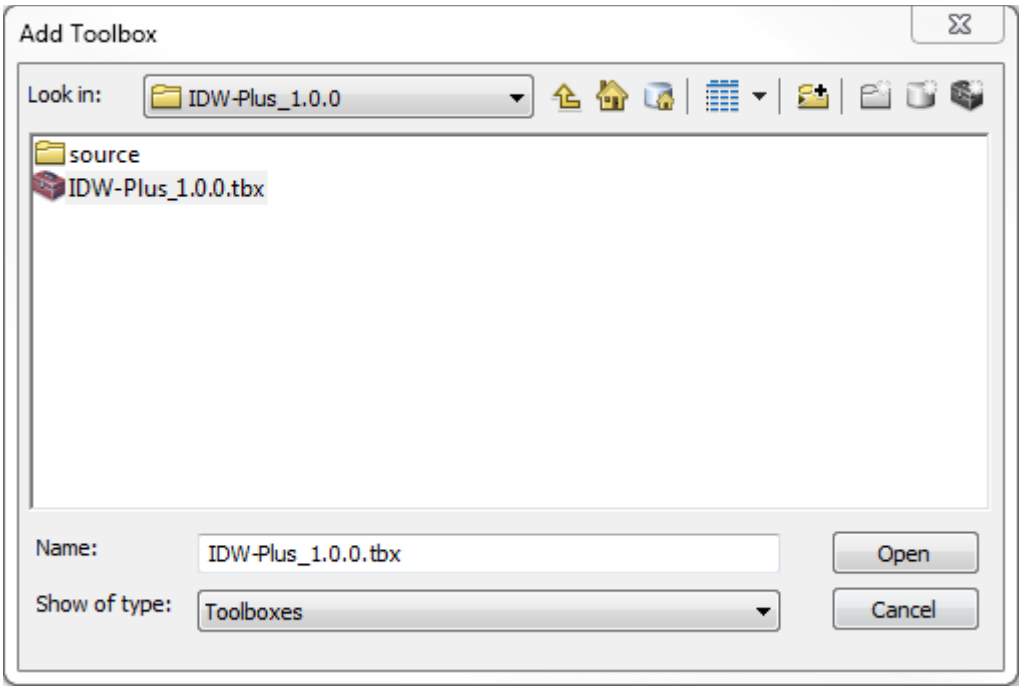
NAME	FORMAT	DESCRIPTION
sites	Point Shapefile	A shapefile of survey site locations
filldem	Raster	Filled digital elevation model
fdgrid	Raster	Flow direction raster
fagrid	Raster	Flow accumulation raster
stream	Raster	Streams, where NoData = stream and 1 = other
lugrid	Raster	Pasture land use, where 1 = pasture and 0 = other
stream_lines	Polyline Shapefile	Streams in the study area
stream_rast	Raster	A raster representation of streams in the study area, where values ≥ streams and NoData = other
raw_dem	Raster	Digital elevation model, unfilled

5. Getting Started

This toolbox can be downloaded at http://www.fs.fed.us/rm/boise/AWAE/projects/SSN_STARS/other_software.html. Save the toolbox and **make sure that there are no spaces in the pathname where the toolbox resides**.

To add the IDW-Plus toolset,

1. Open ArcMap.
2. Add the IDW-Plus toolset. In ArcToolbox, right click on ‘ArcToolbox’, and select ‘Add Toolbox’. Navigate to the IDW-Plus toolset and click OK. Then, right click on ArcToolbox, scroll down, select Save Settings, and click on To Default.



3. 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.

1. Under Processing Extent, set Extent to Same as layer “example/raw_dem.tif”.
2. Under Raster Analysis, set the cell size to Same as layer “example/raw_dem.tif”. This will automatically set the cell size. No Mask is used.

3. Click OK.

***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.

5.1 Tutorial Directory Structure

In this tutorial, we assume that the toolbox files have been saved in a new directory named "C:\IDW-Plus". All of the new data created in the tutorial will be saved in a directory called "C:/IDW-Plus/example/work" – create this directory on your computer. It is not a requirement to use this directory structure to complete the tutorial, but it may be less confusing if users employ the same file structure while working through the tutorial.

6. Pre-processing Steps

In this section, we provide detailed instructions about how to process the data before running the IDW-Plus tools. For example, we assume that the input data have certain characteristics:

1. The stream grid has NoData values for stream cells and is 0 in every other cell;
2. The cells in the land use grid are 1 where a land use type is present and 0 otherwise;
3. The DEM has had the streams "burned" in;
4. The DEM has no sinks; and
5. The flow direction and flow accumulation grids are derived from this 'corrected' DEM.

Note that (3) is not required if the streams have been derived from the same DEM used to run the tools. It is only necessary when the two layers are from different data sources.

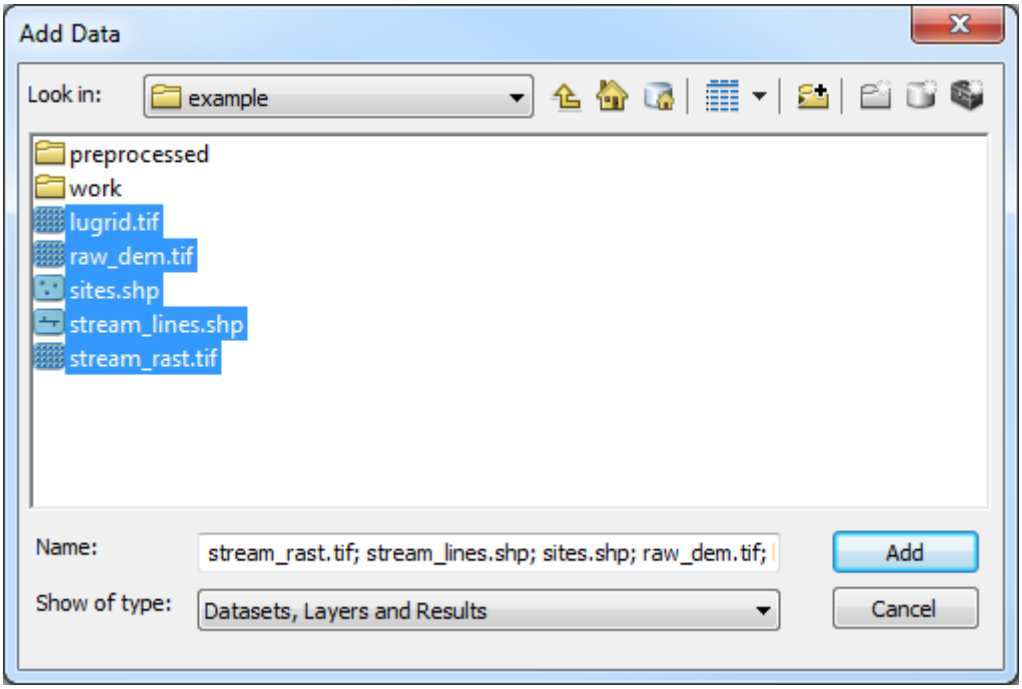
We have included these instructions for users who are not overly familiar with ArcGIS, so that they understand how to pre-process their own datasets. **However, the pre-processed input datasets have been provided in the *example/preprocessed* directory so that users who are familiar with ArcGIS may skip to Section 7**.

6.1 General Pre-processing Steps

1. Load the datasets in the example folder into ArcMap

    1. Click on the Add data tool and navigate to the *example* directory (Table 2).
    2. Select the input datasets in the *example* directory and click Add. This includes:

        - "lugrid.tif"
        - "raw_dem.tif"
        - "sites.shp"
        - "stream_lines.shp"
        - "stream_rast.tif"



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

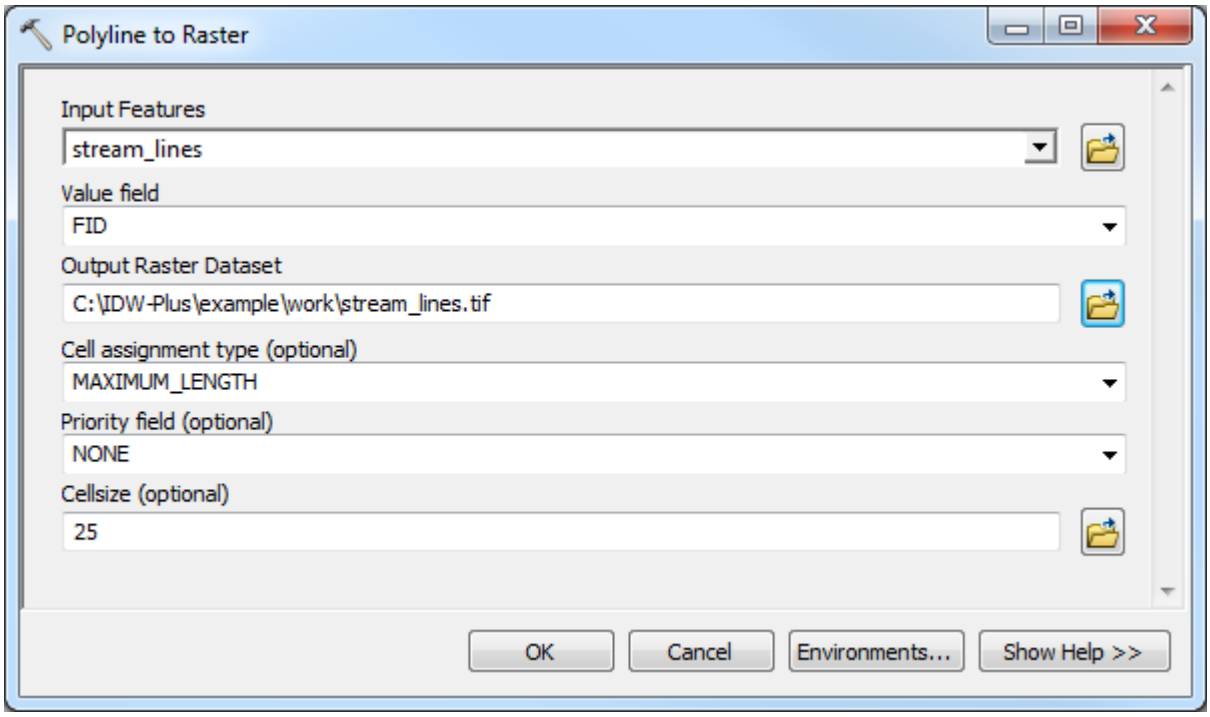
   ***Note**: the example dataset has already been projected to a common projection. See the Data Management Tools> Projections and Transformations and the ArcGIS help if you need more information.

6.2 Creating the Stream Raster

A stream raster with NoData for streams cells and 1 everywhere else is needed to calculate the iFLS, HAiFLS and iEDS metrics. This can be created from either a polylines feature class or an existing streams raster.

To create a stream raster with NoData values from a polyline feature class:

1. Go to ArcToolbox > Conversion Tools > To Raster and open the Polyline to Raster tool.
2. Fill in the arguments as shown below and click OK.



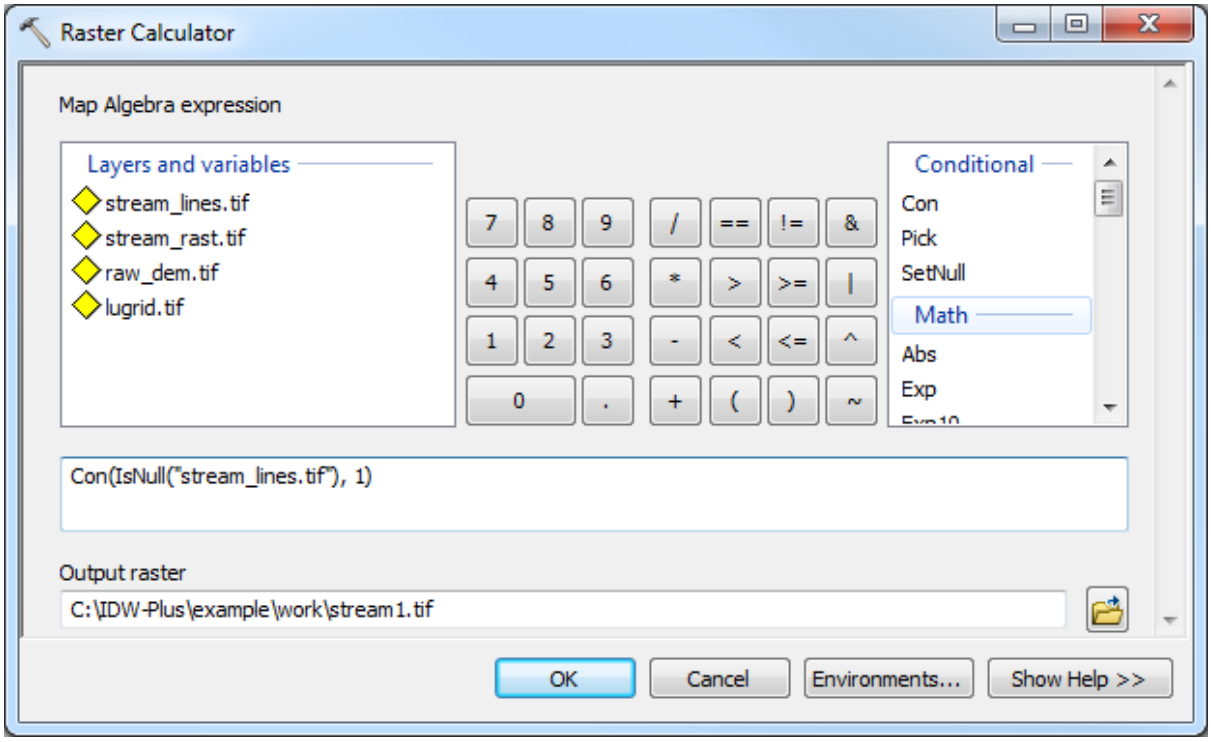
Once the tool has completed running,

3. Make sure the newly calculated “stream_lines.tif” is loaded in the ArcMap Table Of Contents.
4. Open the ArcToolbox> Spatial Analyst Tools> Map Algebra> Raster Calculator tool.
5. Enter the command

   Con(IsNull(“stream_lines.tif”),1)

   Name the Output raster “stream1.tif” and save it to the work directory.

6. Click OK and allow the tool to run. The new streams raster,“stream1.tif”, should appear in the ArcMap Table Of Contents. Examine the data - stream cells should have been assigned NoData values, with 1s for all other cells.



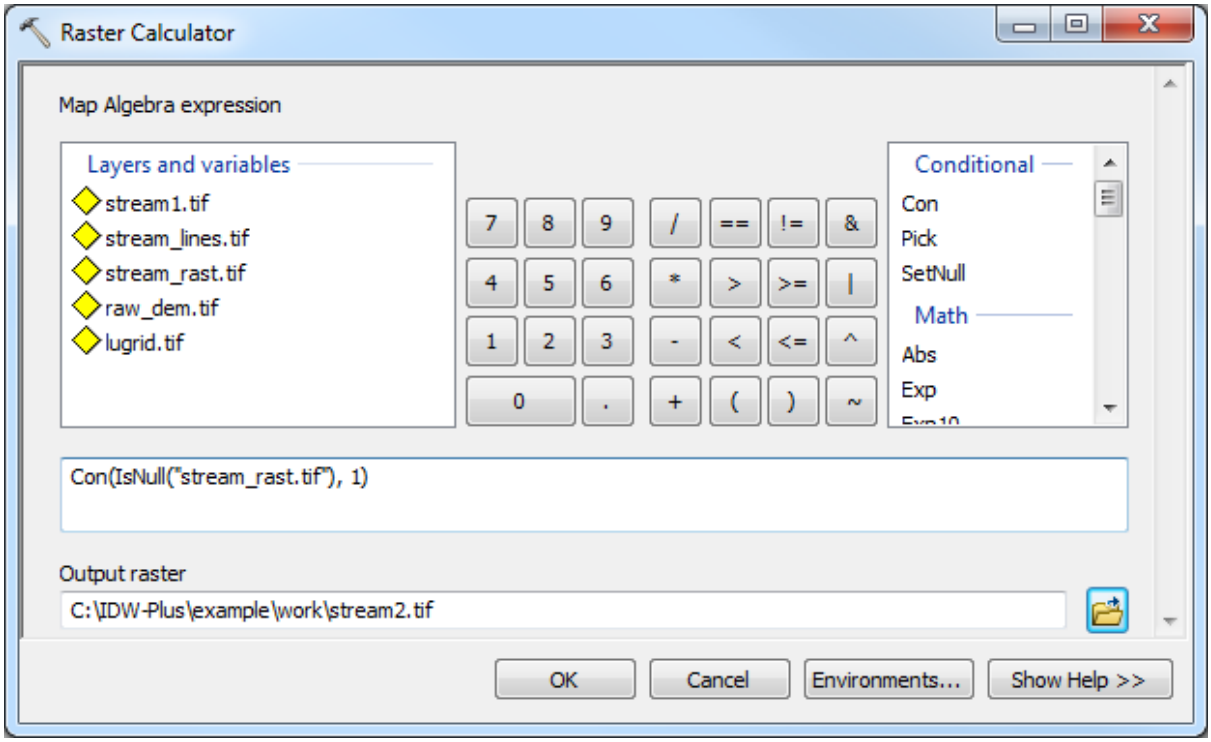
Note, sometimes an error message will appear if you attempt to paste the command into the Raster Calculator. If you get this error message, make sure to double click on the stream_lines.tif layer to enter it into the expression.

To create a stream raster with NoData values from an existing stream raster that has values assigned to stream cells and NoData in all other cells,

1. Open the ArcToolbox> Spatial Analyst Tools> Map Algebra> Raster Calculator tool.
2. Enter the command below,

   Con(IsNull("stream_rast.tif"),1)

   name the Output Raster "stream2.tif", and click OK to run the tool
3. The raster "stream2.tif" should appear in the ArcMap Table Of Contents. Examine the data - it should have NoData for streams cells and 1s everywhere else.



Note: If your streams raster does not contain NoData in non-stream cells, please see the ArcGIS help page for Conditional evaluation with Con:

http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/con-.htm.

Once you have finished with these steps, remove the following layers from the ArcMap session:

- “stream1.tif”
- “stream2.tif”
- “stream_lines.tif”
- “stream_lines.shp”
- “stream_rast.tif”

This can be done by left-clicking on the data layer, right-clicking, and then selecting Remove.

6.3 Formatting the DEM

6.3.1 Burning in the streams

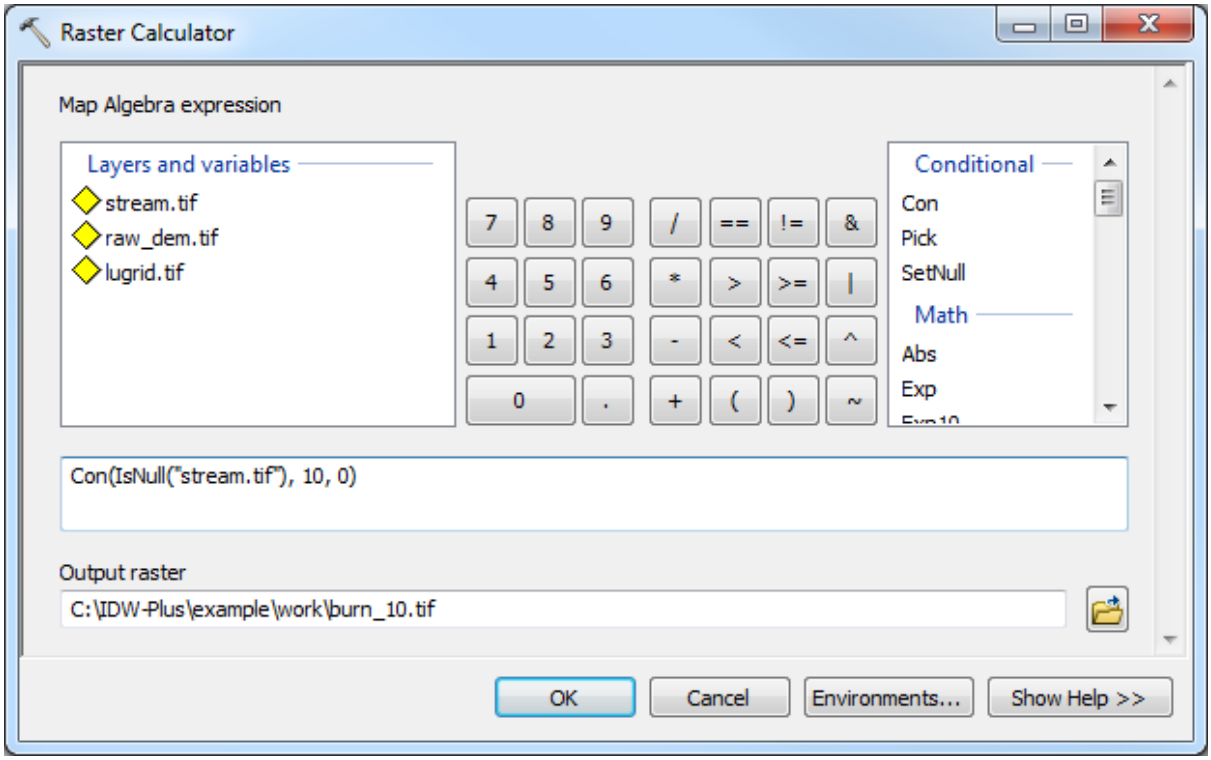
The streams must be burned into the DEM if the streams dataset was derived from a vector of streams or another DEM. If it was derived from the same DEM that will be used to create the spatially explicit watershed attributes, then skip this step.

The first step is to create a raster layer of streams that will be subtracted, or “burned into” the DEM. There is no set rule about how deep to burn the streams in, but typical values range between 5 and 10 meters. In this example, we will burn the streams in 10 meters.

1. Load the preprocessed streams dataset from “example/preprocessed/stream.tif”.
2. Open ArcToolbox> Spatial Analyst Tools> Map Algebra> Raster Calculator
3. Input the command

   Con(IsNull(“stream.tif”), 10, 0)

   and name the Output raster “burn_10.tif”
4. Press OK and wait for the tool to run.



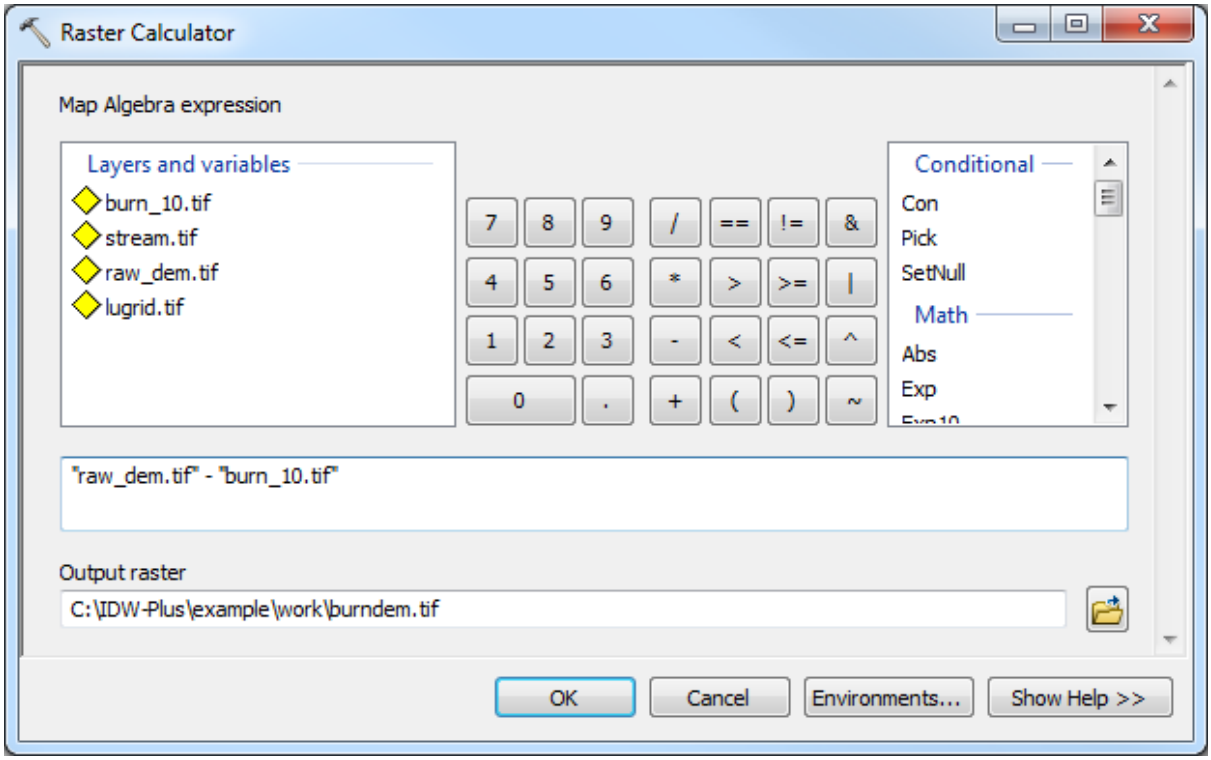
When the tool finishes running, a raster layer of streams called “burn_10.tif” is produced, where stream cells have a value of 10 and all other cells have a value of 0. Make sure “burn_10.tif” is loaded in ArcMap. The next step is to burn these streams into the DEM.

5. Open the Raster Calculator again and input the command

   “raw_dem.tif” – “burn_10.tif”

   and name the output “burndem.tif”.

6. Press OK. The new raster layer is a DEM, with streams burned in 10 meters.

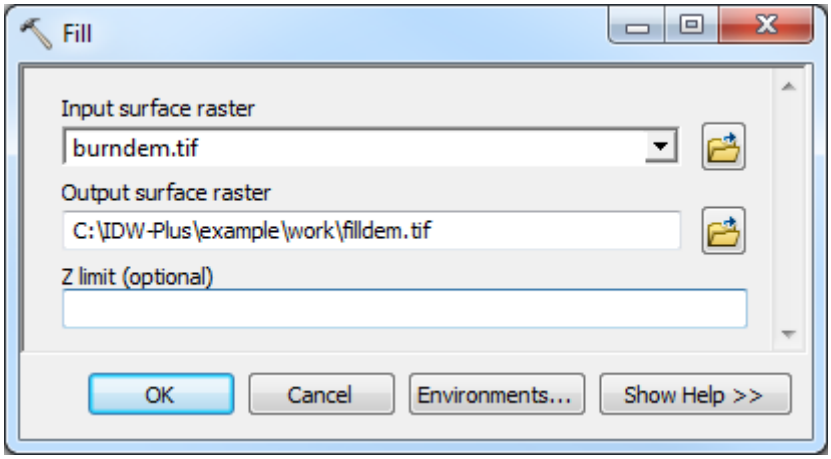


Remove “raw_dem.tif” and “burn_10.tif” from the Table of Contents.

6.3.2 Filling the DEM

DEMs may contain sinks, which are cells whose eight neighbours all flow into it. This is problematic for deriving flow direction and accumulation values, which in turn affects some of the distance-weighted watershed attributes. It is therefore necessary to identify and remove sinks in the DEM.

1. Ensure the “burndem.tif” DEM is loaded in the Table of Contents.
2. Open the ArcToolbox> Spatial Analyst Tools> Hydrology> Fill tool.
3. Enter the arguments as shown below and click OK to run the tool.



For some DEMs a z-limit may need to be specified. More information is available in the ArcGIS documentation for the Fill tool:

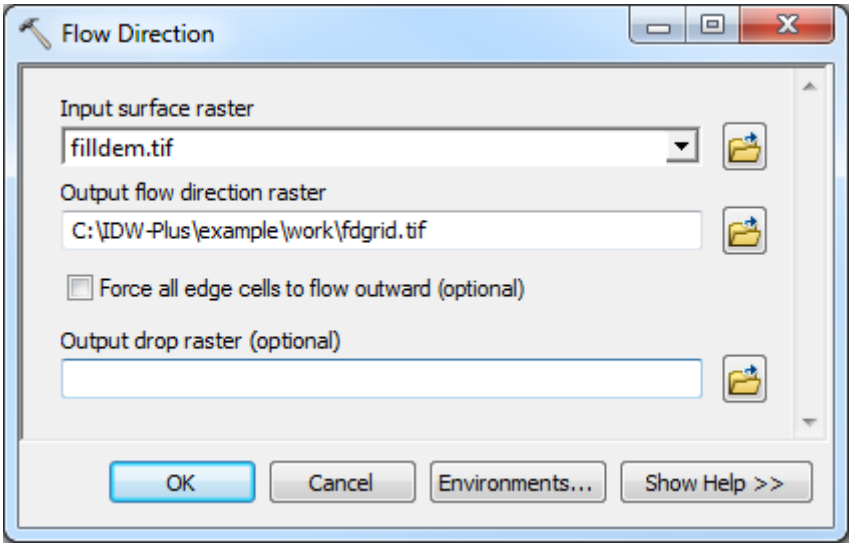
http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/fill.htm

Remove “burndem.tif” from the Table Of Contents.

6.3.3 Creating the Flow Direction and Flow Accumulation Grids

The next step is to derive the flow direction and flow accumulation grids from the filled DEM.

1. In ArcMap, load the “filldem.tif” if it is not already in the Table of Contents.
2. Open the ArcToolbox> Spatial Analyst Tools> Hydrology> Flow Direction tool and provide the tool arguments as shown below.
3. Click OK to run the tool.

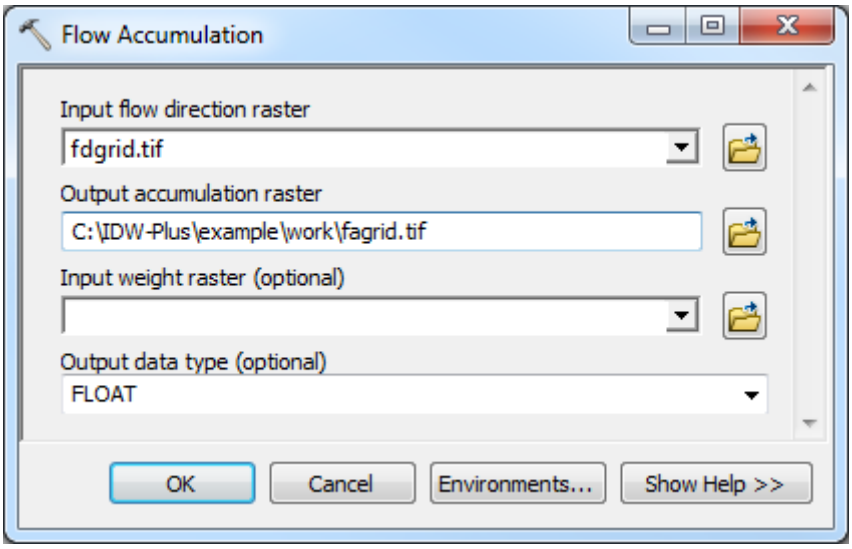


4. Ensure the "fdgrid.tif" is loaded in the ArcMap Table Of Contents.

5. Examine the values in fdgrid.tif, which may include 1, 2, 4, 8, 16, 32, 64, or 128. If any other values are present, then return to Step 6.3.1 because an error has occurred. Please see the ArcGIS help for the Flow Direction tool for additional information: http://pro.arcgis.com/en/pro-app/tool-reference/spatial-analyst/flow-direction.htm

6. Open the ArcToolbox> Spatial Analyst Tools> Hydrology> Flow Accumulation tool and provide the tool arguments as shown below.

7. Click OK to run the tool.



8. Ensure “fagrid.tif” is loaded into the ArcMap TABLE OF CONTENTS.

At this point, all of the pre-processing steps are complete and users have the data they need to run the IDW-Plus tools. Remove the following datasets from the Table Of Contents:

- “fagrid.tif”
- “fdgrid.tif”
- “filldem.tif”

7. Using the IDW-Plus tools

Each metric produces distinct weighting schemes that capture the proximity and runoff potential of the landscape in different ways (*Figure 2*). Greater weight is attributed to land use cells in close proximity to the watershed outlet by the iEDO and iFLO metrics. The percent land use values produced by these two metrics may be strongly correlated (Peterson *et. al* 2011). However, the iFLO metric may still be altered based on the topography of the watershed, which is not the case for the iEDO metric. In the iEDS and iFLS metrics, land use cells in close proximity to the stream have a stronger influence on the percent land use metric. Thus, a larger number of cells in the watershed will strongly influence the land-use metric compared to metrics based on distance to the outlet. For example, the stream usually extends up into the watershed, while the survey site (i.e. outlet) is represented by a single cell. The relative influence of land use cells in the hydrologically active metrics (HAiFLO and HAiFLS) is more complicated than the pure distance-weighted metrics. Land use in close proximity to the stream or outlet, that also has a higher chance of runoff potential (e.g.

intermittent streams and gullies) contribute more to the LU% estimate. However, this also means that the HAIDW metrics are less dependent on the spatial resolution of the streams dataset than the iFLO and iFLS metrics (Peterson *et. al* 2011).

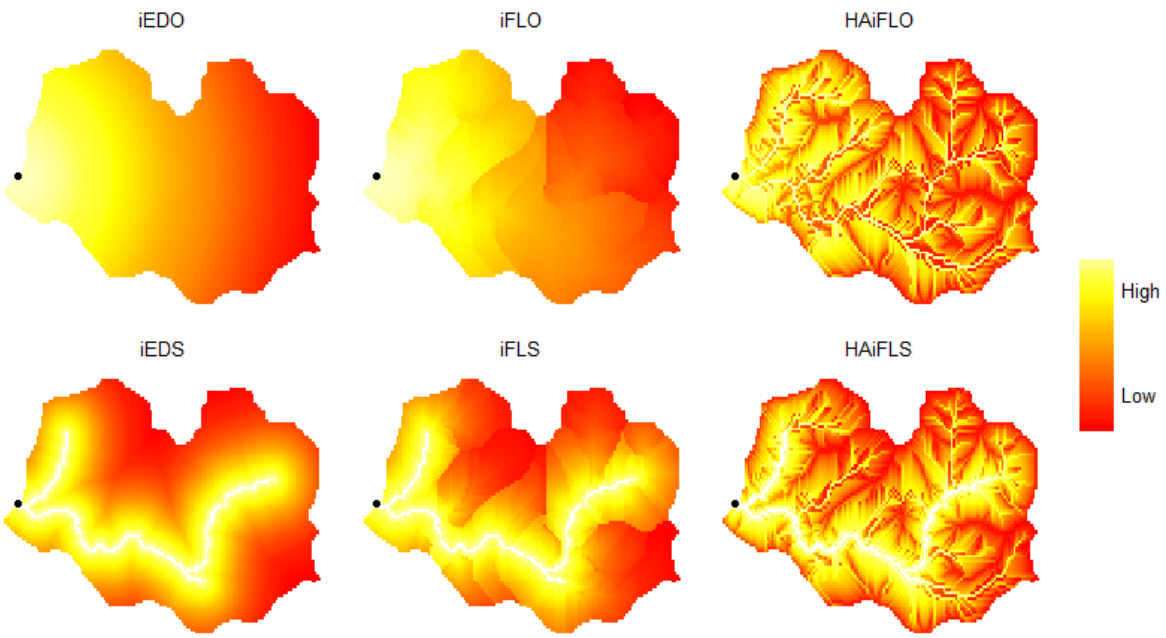


Figure 2. Weighted influence of the landscape within a watershed based on inverse Euclidean-distance to the outlet (iEDO), inverse Euclidean-distance to the stream (iEDS), inverse flow-length to the outlet (iFLO), inverse flow-length to the stream (iFLS), hydrologically active inverse-flow-length to the outlet (HAiFLO), and hydrologically active inverse-flow-length to the stream (HAiFLS). The black circle represents the watershed outlet. Note that the HAiFLO and HAiFLS weights were standardised to range from 0 to 1 to improve visualisation.

The IDW-Plus tools are set to run in the foreground, meaning that no other actions can be performed in the ArcMap session while the script is executing. Note that, the tools may take several minutes to run if the DEM, flow direction and flow accumulation rasters (see Table 4)

are large. To continue using ArcMap while running the tools, another instance of ArcMap must be opened in a separate window.

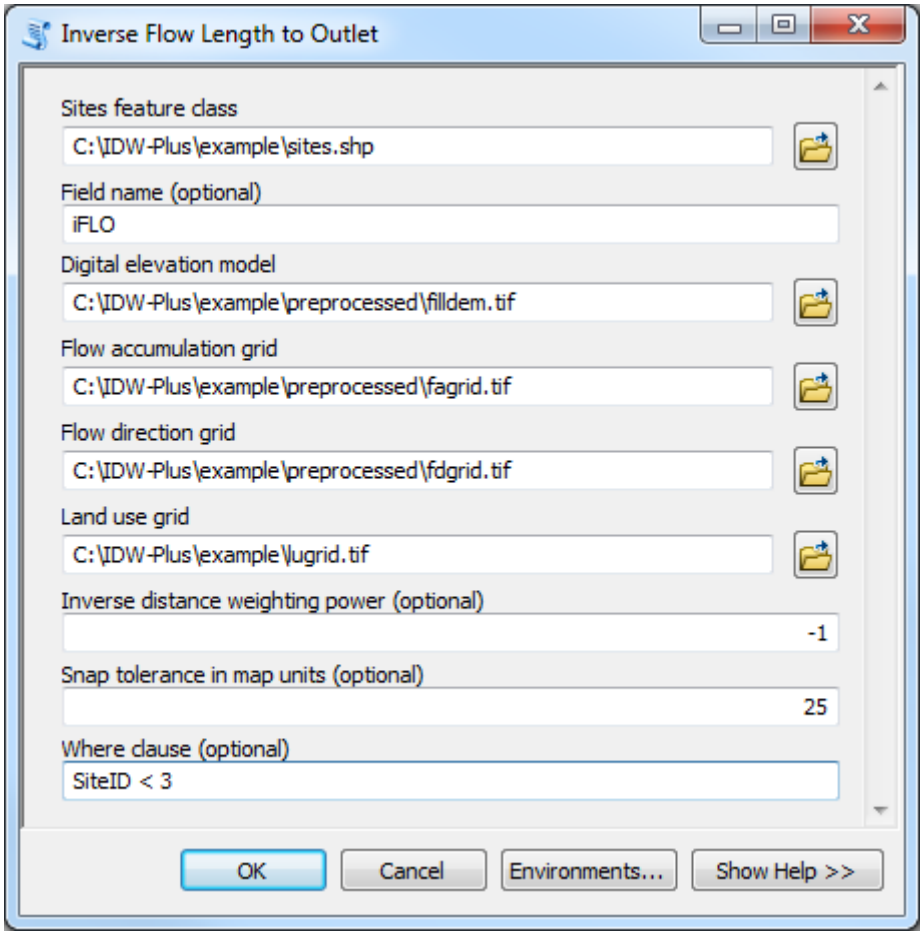
7.1 Tool Inputs and Arguments

The main data inputs to the IDW-Plus tools were described in

Table 1 and the pre-processing steps used to generate them were described in Section 6. However, there are four optional parameters that may be defined in each tool:

1. Field name,
2. Inverse distance weighting power,
3. Snap tolerance, and the
4. WHERE clause.

Details about each of the four optional arguments are provided below.



7.1.1 Field name

The user can specify the new field name that will be created in the sites feature class using the Field name argument.

There are restrictions on the types and number of letters in field names. Specifically,

- A field in a shapefile attribute table cannot be longer than 10 letters.
- A field in a feature class attribute table cannot be longer than 64 letters.
- For restrictions on characters in field names, see: http://support.esri.com/technical-article/000005588

The tool will report an error and immediately stop execution if the applicable character limit is exceeded.

If no field name is provided, a default field name is assigned based on the tool name and the land use raster name, up to a maximum of 10 letters. However, we recommend that users set their own field names.

7.1.2 Inverse distance weighting power

The contribution of each ith land use cell in the catchment to the final metric score is weighted:

$$w_i = (d_i + 1)^p,$$

where w_i is the weighting for the ith cell; d_i is the distance (either Euclidean or flow length distance) of the ith cell to the stream or the outlet; and p is the inverse-distance-weighting power (with $p \leq 0$). The default for the tools is $p = -1$ since this is true inverse distance, but the user can change p according to their needs. Small changes in p can dramatically affect the weighting scheme and the metric value. Users should consider the spatial resolution of their raster layers when making this decision, as well as the scale at which environmental processes that link the terrestrial landscape to the stream operate.

7.1.3 Snap tolerance

The snap tolerance is passed to the Snap Pour Point tool in the Hydrology toolset (see http://pro.arcgis.com/en/pro-app/tool-reference/spatial-analyst/snap-pour-point.htm; ESRI 2015). It is used to improve the delineation of watersheds around pour points (i.e. sites),

which may not be properly aligned with the stream cells. The argument is a single number expressed in map units. In this tutorial, a snap tolerance of 25m has been used because the raster data has a 25m spatial resolution and the site features were already aligned with the stream lines. When the IDW-Plus tools are run, the sites are converted to raster format in the background. However, the newly created source cells may not fall directly on the raster-based stream line if a snap tolerance is not set. Decisions about the snap tolerance value should be based on the user's knowledge about alignment of the sites to the streams, with larger snap tolerance values used when sites have not previously been snapped to stream lines.

7.1.4 WHERE clause

The "WHERE clause" is used in exactly the same way as the WHERE clause in the Select by Attributes function inside a shapefile or feature class attribute table. It is an SQL query that allows users to select and operate on subsets of the data, rather than the full dataset. The general format of the WHERE clause in SQL queries is:

COLUMN_NAME LOGICAL_CRITERIA

For example, to calculate a metric for the first two sites only, the WHERE clause would be:

SiteID < 3

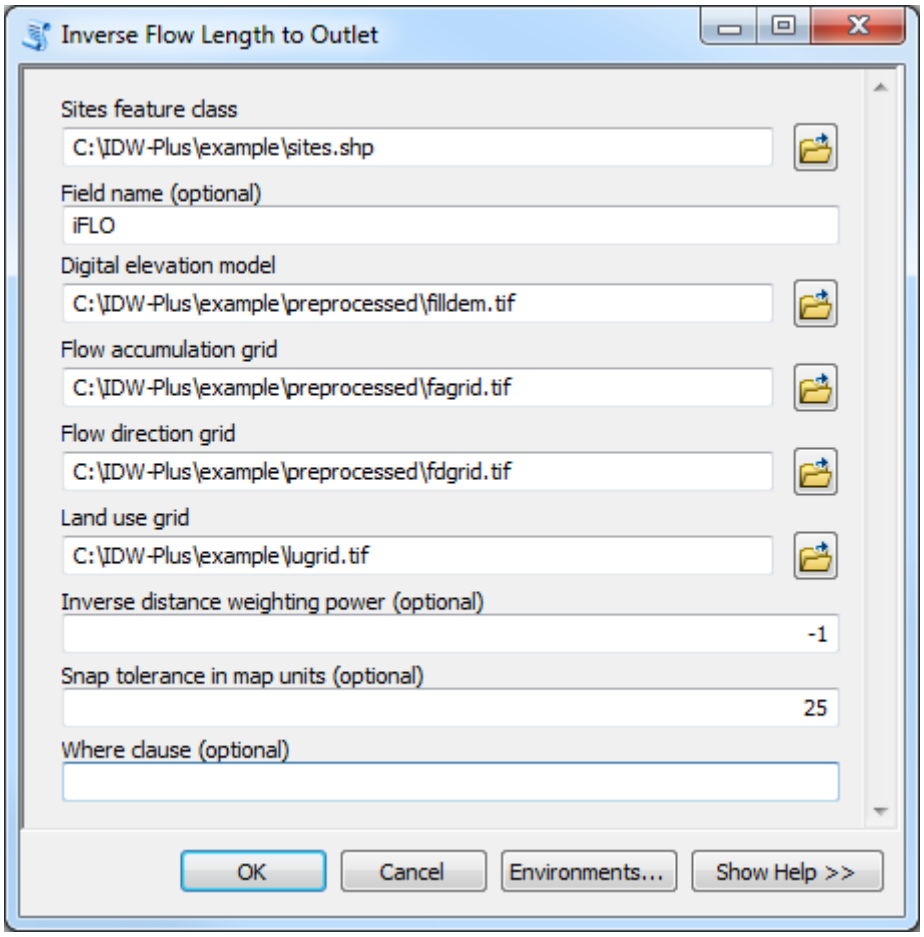
The column names are unique to the data, but examples of logical criteria can be found at http://www.w3schools.com/sql/sql_where.asp. To run the tools on the full dataset, simply leave the WHERE clause blank.

7.2 The iFLO Tool

To use the iFLO tool,

1. Open the ArcToolbox> IDW-Plus> Inverse Flow Length to Outlet tool.
2. Input the tool arguments as shown below and click OK to run the tool.

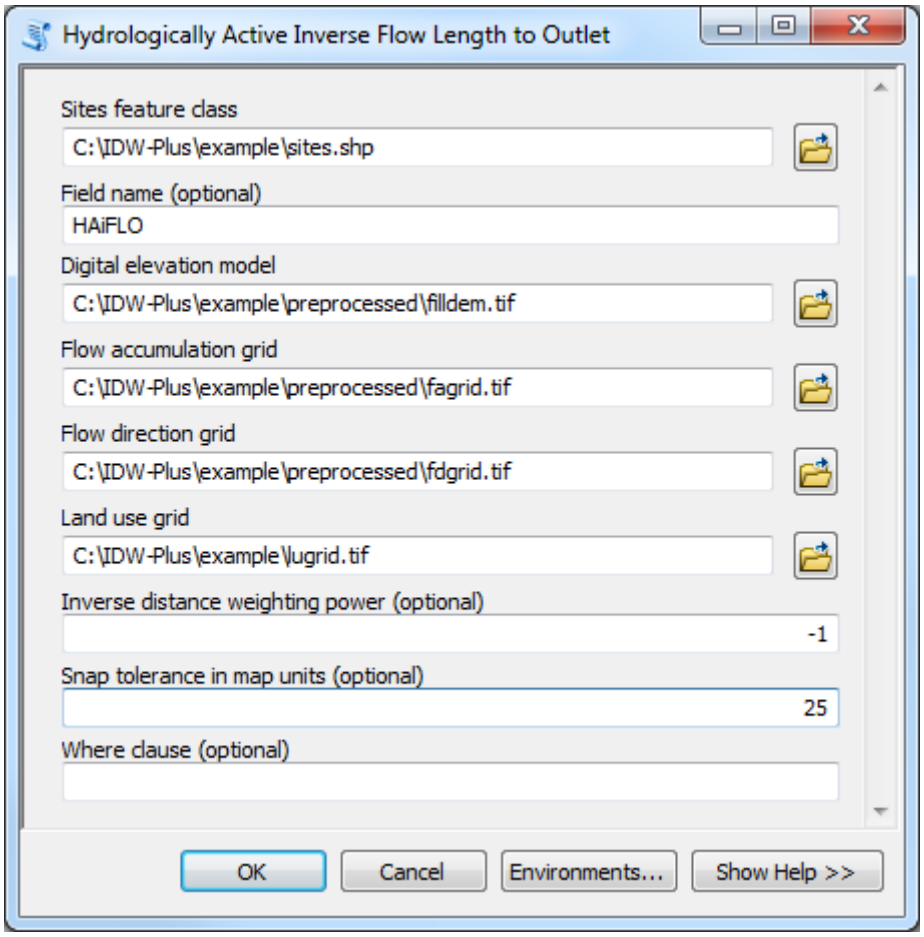
Note that, we will be using the input data provided in the example and example/preprocessed directories in this, and subsequent tool demonstrations.



7.3 The HAiFLO Tool

To use the HAiFLO tool,

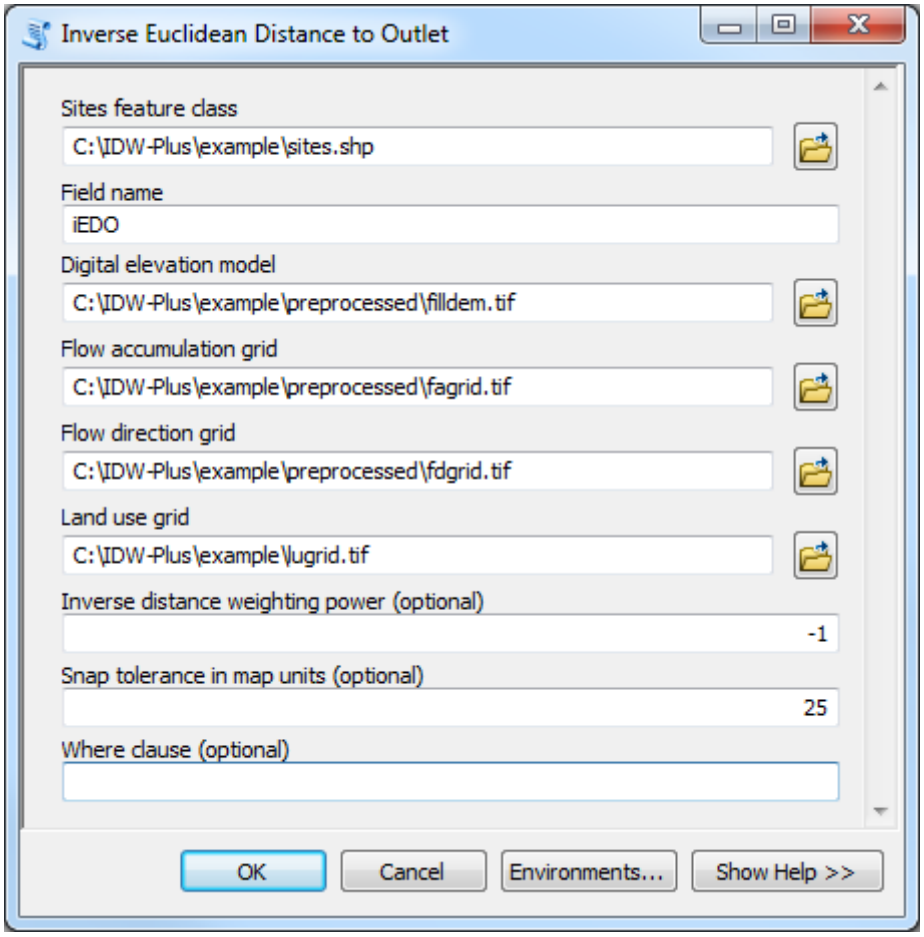
1. Open ArcToolbox> IDW-Plus> Hydrologically Active Inverse Flow Length to the Outlet tool.
2. Input the tool arguments as shown below and click OK to run the tool.



7.4 The iEDO Tool

To use the iEDO tool,

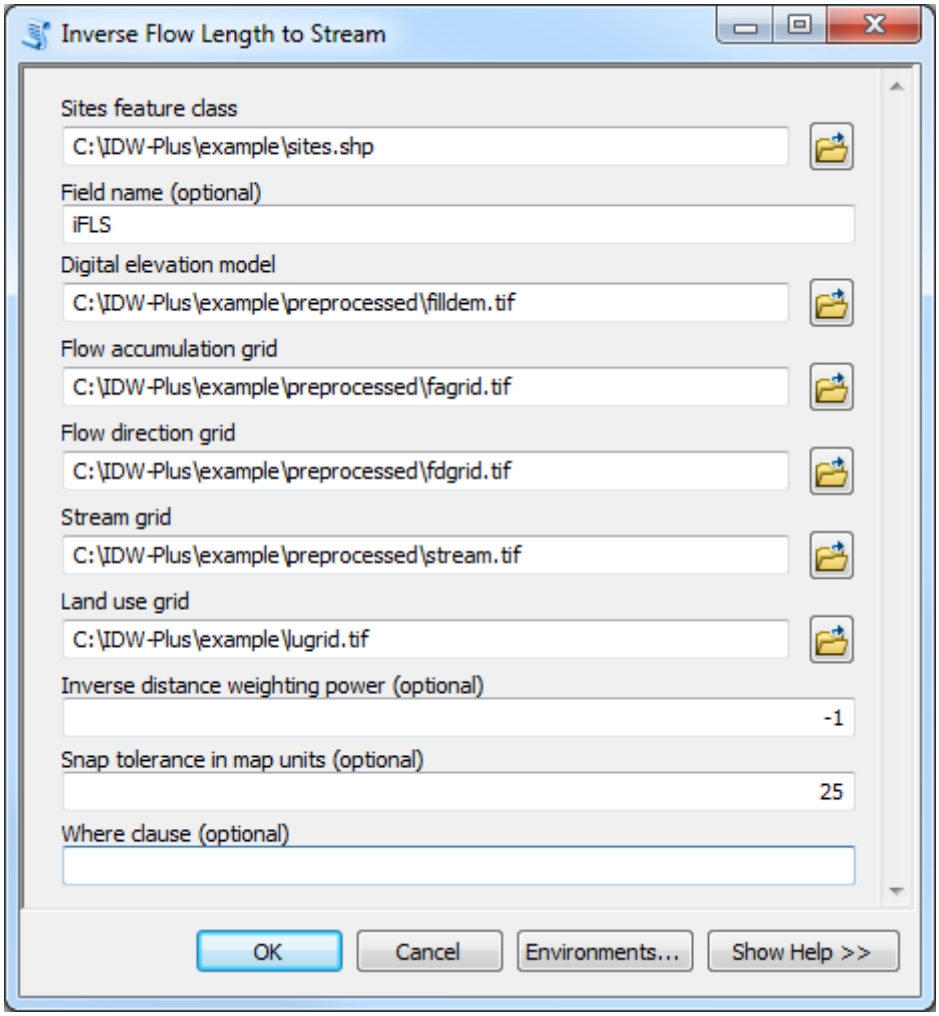
1. Open ArcToolbox > IDW-Plus > Inverse Euclidean Distance to Outlet.
2. Input the tool arguments as shown below and click OK to run the tool.



7.5 The iFLS Tool

To use the iFLS tool,

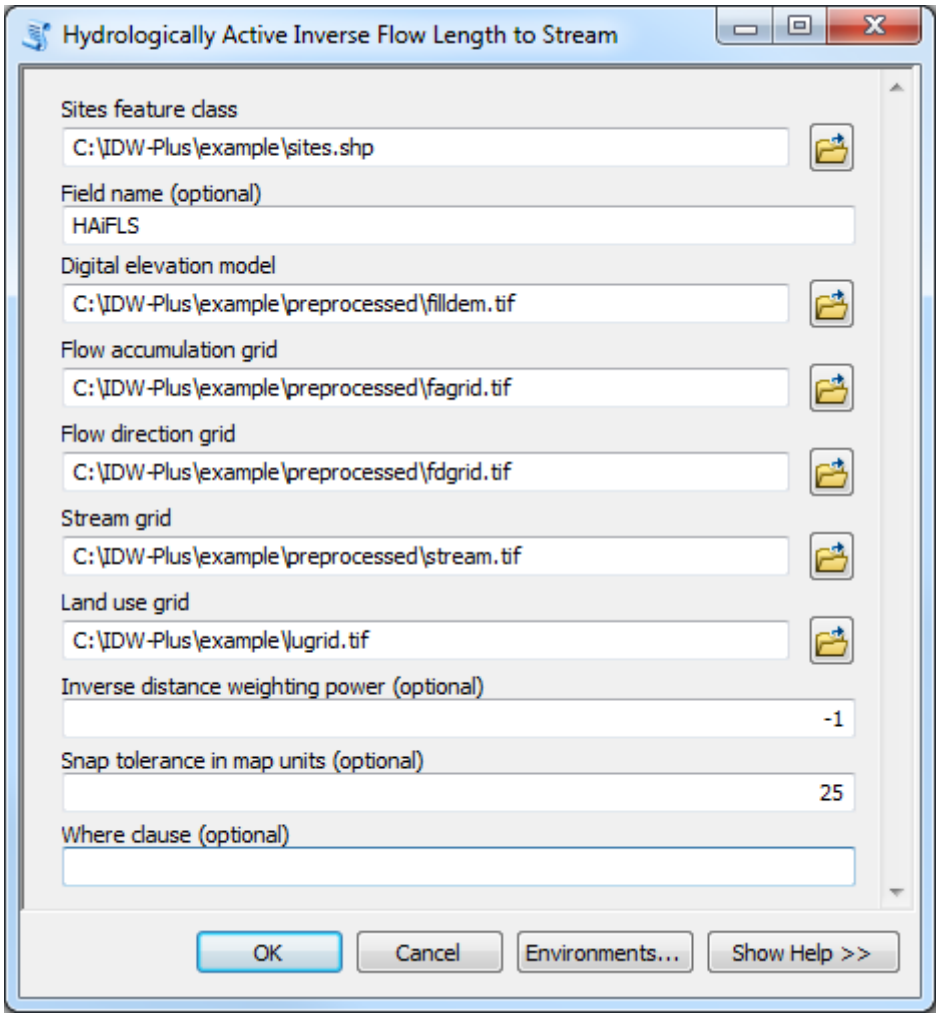
1. Open ArcToolbox > IDW-Plus > Inverse Flow Length to Stream.
2. Input the tool arguments as shown below and click OK to run the tool.



7.6 The HAiFLS Tool

To use the HAiFLS tool,

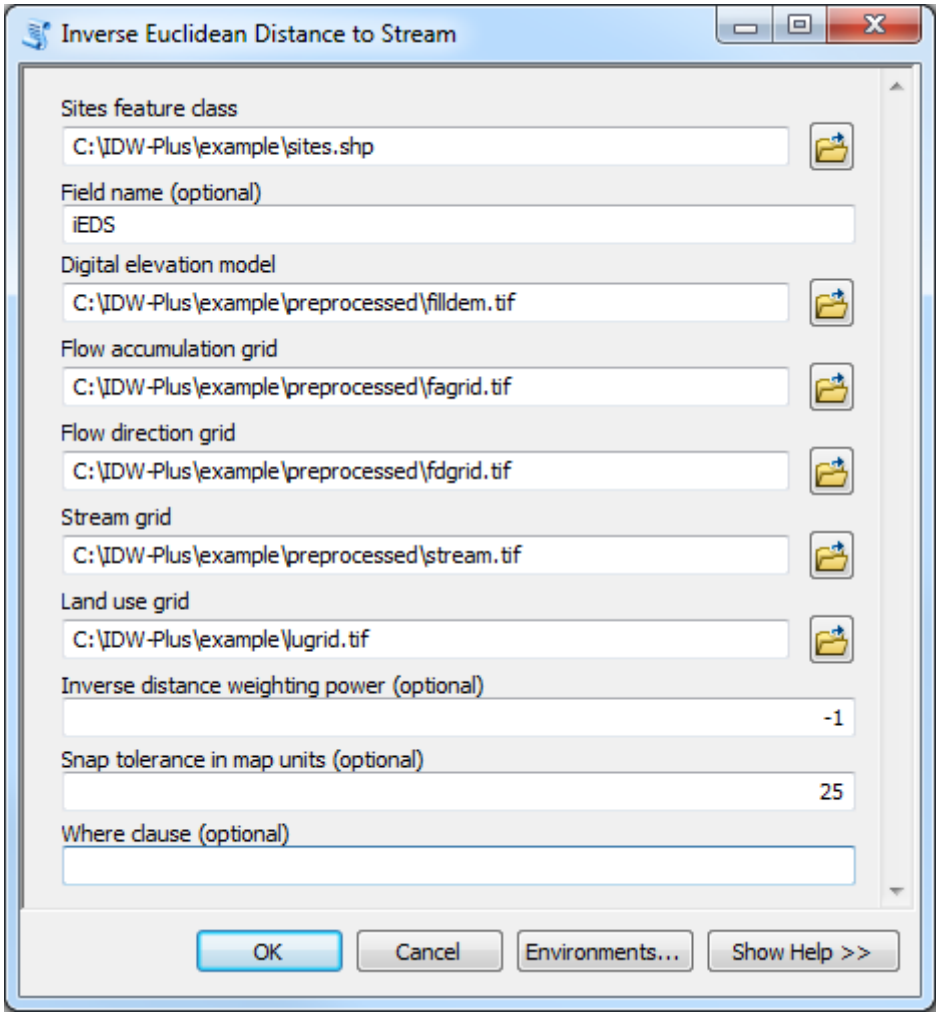
1. Open ArcToolbox > IDW-Plus > Hydrologically Active Inverse Flow Length to Stream.
2. Input the tool arguments as shown below and click OK to run the tool.



7.7 The iEDS Tool

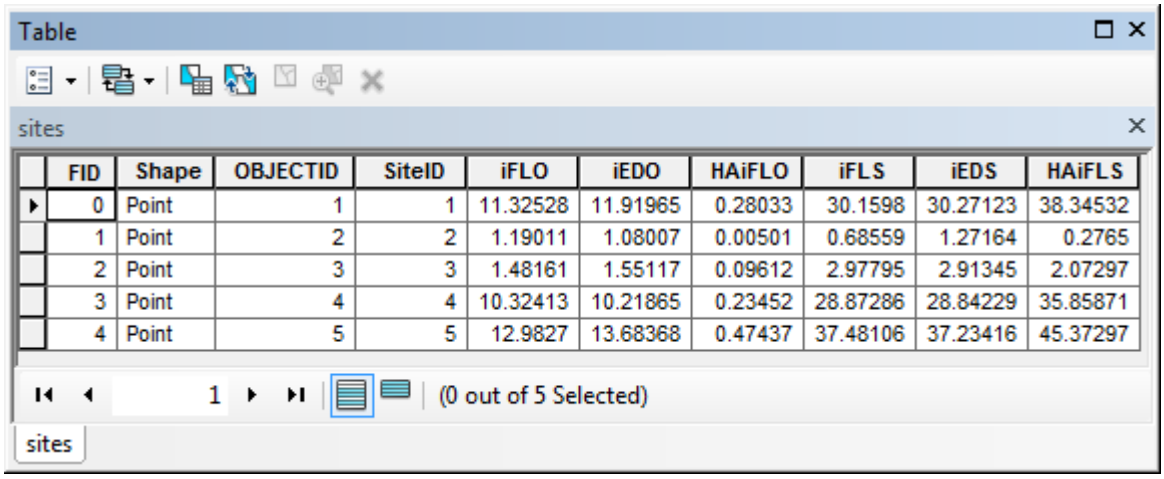
To use the Inverse Euclidean Distance to Stream tool,

1. Open ArcToolbox > IDW-Plus > Inverse Euclidean Distance to Stream.
2. Input the tool arguments as shown below and click OK to run the tool.



8. Output

The spatially explicit landscape representation metrics generated by the IDW-Plus tools are stored in the sites attribute table (Figure 3).



Table

sites

FID	Shape	OBJECTID	SiteID	iFLO	iEDO	HAiFLO	iFLS	iEDS	HAiFLS
0	Point	1	1	11.32528	11.91965	0.28033	30.1598	30.27123	38.34532
1	Point	2	2	1.19011	1.08007	0.00501	0.68559	1.27164	0.2765
2	Point	3	3	1.48161	1.55117	0.09612	2.97795	2.91345	2.07297
3	Point	4	4	10.32413	10.21865	0.23452	28.87286	28.84229	35.85871
4	Point	5	5	12.9827	13.68368	0.47437	37.48106	37.23416	45.37297

1 (0 out of 5 Selected)

sites

Figure 3. The sites attribute table, where the 6 spatially explicit landscape representation metrics have been added, including the inverse Euclidean-distance to the outlet (iEDO), inverse Euclidean-distance to the stream (iEDS), inverse flow-length to the outlet (iFLO), inverse flow-length to the stream (iFLS), hydrologically active inverse-flow-length to the outlet (HAiFLO), and hydrologically active inverse-flow-length to the stream (HAiFLS).

9. References

Bunn, S.E., E.G. Abal, M.J. Smith, S.C. Choy, C.S. Fellows, B.D.Harch, M.J. Kennard & F. Sheldon, 2010. Investigation of science and monitoring of river ecosystem health to guide in catchment protection and rehabilitation. *Freshwater Biology 55*: 223 – 240. DOI: 10.1111/j.1365-2427.2009.02375.x

Bureau of Rural Sciences, Department of Agriculture, Fisheries, and Forestry (BRS), 2002. Land use mapping at catchment scale: principles, procedures and definitions, second edition. Bureau of Rural Sciences, Department of Agriculture, Fisheries, and Forestry, Kingston, ACT, Australia.

Peterson, E.E., F. Sheldon, R. Darnell, S.E. Bunn and B.D. Harch, 2011. A comparison of spatially explicit landscape representation methods and their relationship to stream condition. *Freshwater Biology* 56(3): 590-610. DOI:10.1111/j.1365-2427.2010.02507.x

Queensland Natural Resources and Water (QNRW), 2000. Southeast Queensland 25 meter Digital Elevation Model. Queensland Natural Resources and Water, Indooroopilly, QLD, Australia.