Doc for InVEST Helper Toolbox

These tools were created to help automate common InVEST pre- and post-processing functions done in ArcGIS.

* **Calculate change** - Calculates absolute change and/or percent change for InVEST pixel, subwatershed and watershed outputs.
* **Create servicesheds** (ArcGIS 10 only) - Uses Arc Hydro to create potentially overlapping servicesheds/watersheds providing water services to a set of points of interest.
* **Multi service rank** – Calculates the level of combined ecosystem services on a landscape provided by up to 7 services.
* **Prepare DEM** - Prepares a Digital Elevation Model (DEM) to be used in InVEST terrestrial hydrology models.

**Calculate change**

This tool simply calculates change and/or percent change for per-pixel raster data, per-subwatershed raster data or per-watershed table data.

Change is calculated as .

Percent change is calculated as .

**Inputs**

*Workspace:* The folder where all files generated by the tool are saved. The path to this folder must contain no spaces. Required input.

*Calculate change for raster data?:* If the data to be analyzed are in raster datasets, check this box. This includes per-pixel output from all terrestrial models, as well as per-subwatershed output from the terrestrial hydrology models. Optional input.

*Raster Scenario 1*: Raster dataset containing the first scenario used to calculate change. Required if *Calculate change for raster data?* Is selected.

*Raster Scenario 2*: Raster dataset containing the second scenario used to calculate change. Required if *Calculate change for raster data?* Is selected.

*Are the raster scenarios subwatershed data?*: If *Raster Scenario 1* and *Raster Scenario 2* contain values per subwatershed, check this box. Optional input.

*Subwatersheds*: Shapefile with polygons defining the subwatersheds used for creating the *Raster Scenario 1* and *Raster Scenario 2* subwatershed inputs. This shapefile must contain an ID field with unique values for identifying each subwatershed. Required if *Are the raster scenarios subwatershed data?* is selected.

*Subwatershed ID field*: Integer field in the *Subwatersheds* shapefile containing a unique ID for each subwatershed. Required if subwatersheds are entered.

*Aggregate within an area of interest?:* If change in the raster data should be aggregated over a particular area of interest (such as ownership parcels, park boundaries or political districts), check this box. This option may not be used if subwatershed scenarios are entered. If whole watershed data is to be analyzed, use the *Calculate change for watershed output tables?* option. Optional input.

*Area of interest*: Shapefile with a polygon defining the area of interest (such as ownership parcels, park boundaries or political districts.) This shapefile must contain an ID field with unique values for identifying each area of interest. Required if *Aggregate within an area of interest?* is selected.

*Area of interest ID field*: Integer field in the area of interest shapefile containing a unique ID for each polygon. Required if an *Area of interest* is entered.

*Calculate change for watershed output tables?:* If change should be calculated for watershed data contained in tables, check this box. If selected, two table scenarios must be provided. Optional input.

*Table Scenario 1*: Table (.dbf or .mdb) containing per-watershed data to be analyzed for change. Both Scenario 1 and Scenario 2 tables must contain two fields named the same in each scenario, one containing unique watershed IDs, the other containing the data to be analyzed for change. Required if *Calculate change for watershed output tables?* is selected.

*Table Scenario 2*: Table (.dbf or .mdb) containing per-watershed data to be analyzed for change. Both Scenario 1 and Scenario 2 tables must contain two fields named the same in each scenario, one containing unique watershed IDs, the other containing the data to be analyzed for change. Required if *Calculate change for watershed output tables?* is selected.

*Watershed ID field*: Integer field in the watershed scenario tables containing a unique ID for each watershed. The name of this field must be the same in both Scenario 1 and Scenario 2. Required if *Calculate change for watershed output tables?* is selected.

*Watershed table data field*: Numerical field in the watershed scenario tables containing the data to be analyzed for change. The name of this field must be the same in both *Table Scenario 1* and *Table Scenario 2*. Required if *Calculate change for watershed output tables?* is selected.

*Calculate change?:* If absolute change () is to be calculated, check this box. Change will be calculated for all types of input (raster, subwatershed, watershed) entered. At least one of *Calculate change?* and *Calculate percent change?* must be selected.

*Calculate percent change?:* If percent change ( is to be calculated, check this box. Percent change will be calculated for all types of input (raster, subwatershed, watershed) entered. At least one of *Calculate change?* and *Calculate percent change?* must be selected.

*Split raster results?*: If raster-based change results should be split into two output rasters, one containing values less than zero and the other containing values greater than or equal to zero, check this box. This provides for more accurate visualization of results by allowing each to have its own color gradient assigned for mapping. Note that in the event that one of these split rasters has no data (i.e. all change results are either positive or negative), the raster with no data will show extreme high/low values such as 10e38. Optional input.

*Suffix*: String that will be appended to the end of all output files, in the form of <Filename>\_<*Suffix*>.tif or <Filename>\_<*Suffix*>.dbf. This allows for differentiation of results for different sets of scenarios. Optional input.

**Outputs**

All output files are found in the folder **<*Workspace*>/Post\_process**.

*change.tif*: Absolute change () for per-pixel or per-subwatershed data. Output if *Calculate change?* and *Calculate change for raster data?* are selected.

*change\_lt\_zero.tif*: All pixels in the *change.tif* output raster that have change values less than zero. Output if *Split raster results?* and *Calculate change for raster data?* are selected.

*change\_gte\_zero.tif*: All pixels in the *change.tif* output raster that have change values greater than or equal to zero. Output if *Split raster results?* and *Calculate change for raster data?* are selected.

*percent\_change.tif*: Percent change ( for per-pixel or per-subwatershed data. Output if *Calculate percent change?* and *Calculate change for raster data?* are selected.

*percent\_change\_lt\_zero.tif*: All pixels in the *percent\_change.tif* output raster that have percent change values less than zero. Output if *Split raster results?* and *Calculate change for raster data?* are selected.

*percent\_change\_gte\_zero.tif*: All pixels in the *percent\_change.tif* output raster that have percent change values greater than or equal to zero. Output if *Split raster results?* and *Calculate change for raster data?* are selected.

*change\_subwatershed.dbf*: Table containing change and/or percent change output values per subwatershed. If *Calculate change?* is selected, change values are found in the output table’s **change** field, and if *Calculate percent change?* is selected, percent change values are in the **pchange** field. Subwatershed ID values are in the **subws\_id** field.

*change\_area.dbf*: Table containing change and/or percent change output values per area of interest. Absolute change is calculated by summing *change.tif* output within each area of interest. Percent change is determined by summing each of *Raster Scenario 1* and *Raster Scenario 2* within each area of interest then calculating percent change using the sums. If *Calculate change?* is selected, change values are found in the output table’s **change** field, and if *Calculate percent change?* is selected, percent change values are in the **pchange** field. Area of interest ID values are in the **area\_id** field.

*change\_watershed.dbf*: Table containing change and/or percent change output values per watershed. . If *Calculate change?* is selected, change values are found in the output table’s **change** field, and if *Calculate percent change?* is selected, percent change values are in the **pchange** field. Watershed ID values are in the *Watershed ID field*entered by the user.

**Create Servicesheds (ArcGIS 10+ only)**

This tool uses Arc Hydro to create servicesheds for a set of points of interest. In this case, a serviceshed is the upstream watershed that drains into a point of interest, called an outlet. Outlets may be towns, dams, water treatment plants or any other points where water-related ecosystem services are used by people. The generated servicesheds may be nested, for example if the outlets lie sequentially along a particular stream.

**Requirements**

At the moment, this tool requires a beta version of Arc Hydro, which is included in the InVEST installer. After the Helper Toolbox is installed, the next time ArcGIS is opened, the Arc Hydro toolbox will be accessible from ArcToolbox, or via the menu Customize -> Toolbars -> Arc Hydro Tools.

**Inputs**

*Workspace:* The folder where all files generated by the tool are saved. The path to this folder must contain no spaces. Required input.

*DEM*: Digital Elevation Model raster, which provides the terrain information for creating the servicesheds. This raster must be in a projected coordinate system. It should have sinks filled, have no areas of missing data (NoData cells within the area of interest), and if necessary have streams burned in. See the **Prepare DEM** tool for help with this process. Required input.

*Serviceshed outlets*: Point shapefile containing the locations of serviceshed outlets (towns, dams, etc.) One serviceshed will be created for each outlet, and the resulting servicesheds may be nested (and thus overlap.) This shapefile must have an ID field with a unique value for each outlet. Required input.

*Serviceshed ID field*: Integer or text field in the *Serviceshed outlets* shapefile containing a unique ID for each outlet. This ID will be used to identify the resulting servicesheds. Required input.

*Stream threshold*: Integer value defining the number of upstream raster cells that must flow into any given cell for that cell to be considered part of a stream network. Larger values produce coarser stream networks (major rivers only), smaller values produce finer networks. Choose a value that creates a stream network that is close to an actual perennial stream map for the area of interest. The tool creates a *streams\_<Stream threshold>.tif* output raster showing the Arc Hydro-generated stream layer corresponding to this stream threshold. Compare this raster with the actual stream map and adjust the stream threshold value until the two layers' stream networks look the same. Required input.

*Suffix*: String that will be appended to the end of all output files, in the form of <Filename>\_<Suffix>.tif. This allows for differentiation of results for different input parameters. Optional input.

**Outputs**

All final output files are found in the folder **<*Workspace*>/Output**.

*streams\_<Stream threshold>.tif*: Raster showing the stream network generated by Arc Hydro, based on the *Stream threshold* input. Compare this raster with a known accurate perennial stream network map, and adjust the *Stream threshold* value until the two stream networks look the same.

*servicesheds.shp*: Shapefile containing the generated servicesheds, one per outlet. The **Name** field in this shapefile contains the user-specified *Serviceshed ID field*.

There are also many intermediate steps involved in making servicesheds, some of which may be of interest. Rasters are found in the folder **<*Workspace*>/Intermediate** and feature classes in the geodatabase **<*Workspace*>/Intermediate/ssheds.gdb**. They are listed here, in the order they are created, with a brief description. For more detailed information, see the Arc Hydro documentation.

*flow\_dir*: Flow direction raster, with a value designating in which direction water flows off of this cell downslope.

*flow\_acc*: Flow accumulation raster, with a value for how many upstream cells flow into this cell.

*streams*: Stream network raster, based on the input *Stream threshold*. This is the same as the *streams\_<Stream threshold>.tif* output.

*stream\_link*: Raster of stream segments, each with a unique ID.

*catch\_grid*: Catchment raster, with values describing to which catchment each cell belongs.

*sshed.gdb/catch\_poly*: Polygon version of the Catchment raster.

*sshed.gdb/drainage\_line*: Feature class with information about each stream segment: which catchment it belongs to and how it connects to downstream features.

*sshed.gdb/adj\_catch*: Feature class with aggregated upstream catchments.

*sshed.gdb/batchpoints*: ArcHydro requires that the watershed outlets (called batchpoints) that are entered into its watershed delineation tool have particular fields and values. The *Serviceshed outlets* shapefile is used to create this batchpoint input.

*sshed.gdb/servicesheds*: Final servicesheds. This is the same as the *servicesheds.shp* output.

*sshed.gdb/ssheds\_points*: Points corresponding to the outlets, snapped (if appropriate) to the nearest stream.

**Notes**

If serviceshed outlets are not located directly on a generated stream, very small servicesheds may be created, which may not be correct or desired. To cause outlets to ‘snap’ to (be moved to lie on top of) the nearest generated stream, the following must be done by the user outside of the Create Servicesheds tool:

1. If the Arc Hydro toolbox has not already been added to ArcToolbox, click on the ArcGIS Customize menu, then Toolbox, and check the box next to Arc Hydro Tools.
2. In the Arc Hydro tool bar, click on the ApUtilities drop-down menu and select XML Manager.
3. Click on the + next to HydroConfig, then ProgParams, then ApFunctions, then ApFunction(BatchWatershedDelineation)
4. Right-click on SnapToleranceNumCells and select Edit Text.
5. Enter a Value for how many cells around an outlet the tool should look for a stream to snap to. The value chosen will depend on how large the cell size is, how far away the outlets are from streams and knowledge of the area of interest. This value might need to be adjusted several times.
6. Click OK and click on the red X button in the upper right hand corner of the window to exit out of the XML viewer

Then re-run the Create Servicesheds tool and see if the servicesheds look better. Also look at the *ssheds\_points* intermediate output to see which points have been snapped and which haven’t and adjust the SnapTolerance accordingly.

**Multi-service rank**

This tool determines the total relative level of ecosystem services provided by each cell on a landscape, by combining InVEST output data for multiple services. For each service, a service quantity raster is entered (such as carbon sequestration or sediment retention), along with a numerical weight indicating how much importance is placed on that service, relative to the other services entered. The service raster will be normalized by its highest value, providing a relative ranking of service provision per cell, and the normalized result will be multiplied by the weight to get a weighted ranking raster for that service. All weighted ranking layers are then added together to create the final combined ranking raster, where higher numbers indicate higher levels of combined ecosystem services. Optionally, the final ranking raster values can be grouped by a user-defined percent, to show which areas of the landscape are providing the highest (and lowest) X% of combined ecosystem services.

**Inputs**

*Workspace:* The folder where all files generated by the tool are saved. The path to this folder must contain no spaces. Required input.

*Services*: List of rasters with service quantity data, output by an InVEST tool, such as carbon sequestration or sediment retention. At least one Service raster is required. If only one Service is entered, ranking and grouping will be still be done on that service alone.

*Weights*: Floating point values, assigning a weight to each *Service*, relative to other services. One *Weight* is required for each *Service* entered, and the *Weights* must be entered in the same order as the corresponding *Services*.

*Grouping percent*: Integer value indicating that the final ranked output should be grouped by X%. Optimally, this will create (100/X%) output groups, but depending on the value distribution some groups may be missing (for example, if no values fall within that group.) Optional input.

*Group by area?*: Check this box if groups should be created by slicing the ranked output by X% of total area. Using this option, a similar number of cells will be assigned to each group. If *Grouping percent* is defined, at least one of *Group by pixel value?* and *Group by distribution?* must be selected. If both are selected, two different outputs will be created.

*Group by distribution?*: Check this box if groups should be created by evenly slicing the ranked output value distribution by X%. Note that with this option, it is possible that very small or very large numbers of cells may fall within each group, and there may be groups containing no cells at all. If *Grouping percent* is defined, at least one of *Group by pixel value?* and *Group by distribution?* must be selected. If both are selected, two different outputs will be created.

*Suffix*: String that will be appended to the end of all output files, in the form of <Filename>\_<Suffix>. This allows for differentiation of results for different input parameters. Optional input.

**Outputs**

All output files are found in the folder **<*Workspace*>/Post\_process**.

*combined\_service\_rank.tif*: Final ranking raster, combining normalized, weighted rasters from all services. High values indicate high levels of combined ecosystem service provision.

*combined\_service\_group\_by\_distribution.tif*: Final ranking raster, grouped by X% of the value distribution. A value of 1 indicates where the top X% of combined ecosystem service provision is located. Larger numbers indicate lower levels of provision.

*combined\_service\_group\_by\_distribution.shp*: Shapefile version of *combined\_service\_group\_by\_distribution.tif.*

*combined\_service\_group\_by\_value.tif:* Final ranking raster, grouped by X% of total area. A value of 1 indicates where the top X% of combined ecosystem service provision is located. Larger numbers indicate lower levels of provision.

*combined\_service\_group\_by\_value.shp*: Shapefile version of *combined\_service\_group\_by\_value.tif.*

**Prepare DEM**

This tool performs several functions on a Digital Elevation Model (DEM) to prepare it for use in the InVEST terrestrial hydrology models:

* Fills small holes (areas of NoData cells) within the DEM, using 2 passes of a Nearest Neighbor function, the first with a 5x5 neighborhood around all NoData cells to fill the smallest holes, the second with an 11x11 neighborhood.
* Fills large holes within the DEM, using interpolation to create a new elevation surface without holes, then the NoData cells in the original DEM are assigned values from the interpolated surface.
* ‘Burns’ streams into the DEM, making DEM cells that lie along user-defined streams 3 map units lower than they are currently. This forces the DEM to generate the desired stream network.
* Fills sinks in the DEM, to fix areas that have undefined or problematic flow directions, usually caused by errors in the original DEM data. After sinks are filled, a Flow Direction raster is created from the filled DEM, to see if all cells are assigned one of the 8 cardinal flow directions. If there are non-cardinal values found, sinks will be filled again, up to three times. If non-cardinal values are still found, a warning will be given and the user should look closely at the quality of the DEM.

**Inputs**

*Workspace:* The folder where all files generated by the tool are saved. The path to this folder must contain no spaces. Required input.

*DEM*: Digital Elevation Model raster, which must be in a projected coordinate system. Required input.

*Fill sinks*: If sinks in the DEM should be filled, check this box. Optional input.

*Burn streams*: If streams should be burned into the DEM, check this box. A *Streams* shapefile must also be provided. Optional input.

*Streams*: Line shapefile containing a perennial stream network that is known to be accurate, which will be burned into the DEM. Required if *Burn streams* is selected.

*Fill small holes*: If small holes in the DEM (just a few cells across) should be filled, check this box. Optional input.

*Fill large holes*: If large holes in the DEM should be filled, check this box. The interpolation process can take a very long time, especially with large DEMs, so only use this option if necessary. If unsure, run the tool once selecting only *Fill small holes*, and if areas of NoData are still found in the output DEM, do a second run selecting both *Fill small holes* and *Fill large holes*. Optional input.

*Suffix*: String that will be appended to the end of all output files, in the form of <Filename>\_<Suffix>.tif. This allows for differentiation of results for different input parameters. Optional input.

**Outputs**

*dem\_prep.tif:*  Final prepared DEM that has had all of the selected processes performed on it.

**Watershed table data field (optional)**

Numerical field in the watershed scenario tables containing the data to be analyzed for change.

The name of this field must be the same in both Scenario 1 and Scenario 2.

**Calculate change? (optional)**

If absolute change is to be calculated, check this box.

Absolute change = (Scenario 1 - Scenario 2)

At least one of Calculate Change and/or Calculate percent change must be selected.

**Calculate percent change? (optional)**

If percent change is to be calculated, check this box.

Percent change = ((Scenario 1 - Scenario 2) / Scenario 2) \* 100

At least one of Calculate Change and/or Calculate percent change must be selected.

**Split raster results? (optional)**

If raster results should be split into two rasters, one containing values less than zero and the other containing values greater than or equal to zero, check this box.

This allows for more accurate visualization of results by allowing each to have its own color gradient assigned for mapping.

**Suffix (optional)**

Optional string that will be appended to the end of all final output files, in the form of &lt;Filename&gt;\_&lt;Suffix&gt;