Documentation
Areal-Weighted Reaggregation Tool

D e s c r i p t i o n

The Areal-Weighted Reaggregation (AWR) tool is a custom Python script that allows ArcGIS users to perform an areal-weighted aggregation using any two input polygon feature classes. Users of the tool are required to input the following parameters:

- **Workspace** - ***the path telling the computer where you're storing the files of interest.*** The script will use this path to access the input polygon feature classes and store generated outputs.

- **Data Feature Class** - ***the feature class holding your data of interest.*** This is the file that contains the **aggregated data** that you will be working to reaggregate. For example, if you have data by Census **block groups** and you want to reaggregate these data by **county boundaries**, your **block groups** feature class would be the input for this parameter.

- **Areal Units Feature Class** - ***the feature class holding your areal unit of interest.*** This is the file that contains the **areal units** by which you are reaggregating. For example, if you have data by Census **block groups** and you want to reaggregate these data by **county boundaries**, your **county boundaries** feature class would be the input for this parameter.

- **Dissolve Field** - ***the field in your areal unit feature class that uniquely identifies each feature in the feature class.*** Your data will be reaggregated (**dissolved**) based on this field. For example, in your **county boundaries** feature class, there is likely a field (i.e. **COUNTY_ID**) that uniquely identifies each county. This would be the input for this parameter.

- **Intersect Feature Class** - ***the intersect feature class created as part of the areal-weighted reaggregation.*** This is one of two outputs generated by the tool and includes the result of an intersect analysis between your data feature class and areal unit feature class. This feature class isn't imperative for analysis, but it does have additional applications beyond running an AWR. For this parameter you provide a file name and output path, and the script will generate the intersect feature class with that name.

- **Statistics Field(s)** - ***the field(s) that you want to reaggregate (SUM) in the output feature class.*** These fields are the **data of interest** contained in your data feature class. For example, if I wanted to aggregate the number of **Black**, **White**, **Hispanic**, and **Asian** people by county, I would select fields **"BLACK"**, **"WHITE"**, **"HISPANIC"**, and **"ASIAN"** for this parameter. Note that this parameter can be any numerical input.

- **Output Feature Class** - ***the output of your areal-weighted reaggregation.*** The output feature class of our AWR will provide the statistics specified in the Statistics Field(s) parameter reaggregated by the areal units specified in the Dissolve Field parameter. These statistics will appear as new fields with a prefix **"SUM_"** and a **suffix "_awr"**.

F u n c t i o n a l i t y

What is areal-weighted reaggregation and how does it work? AWR is a geospatial technique developed to take data that is aggregated according to the geometries of one areal unit (i.e. county, state, Census block groups, protected areas) and reaggregate that data based on the geometires of a different areal unit **when these areal units are not nested.** If the areal units are nested, simply using the Dissolve tool will perform a more rapid analysis. The technique, as performed by this script, involves the following steps:

(a) Create a new field in your **Data Feature Class** and populate this field with the **area** value of each of the features in the feature class.

(b) Intersect your **Data Feature Class** with your **Areal Units Feature Class**.

(c) Create a new field in your **Intersect Feature Class** and populate this field with the **area** value of each of the intersected features (each of the features in your **Intersect Feature Class**).

(d) Create a new field in your **Intersect Feature Class** and populate this field by dividing the **area** of your **Intersect Feature Class** features by the **area** of your **Data Feature Class** features. Place this value in a field called **AREA_RATIO**. **(e)** Determine the **Statistic Field(s)** you are interested in reaggregating. For **x** number of statistic fields, create **x** number of new fields to be populated.

(f) Multiply the values in each **Statistic Field(s)** by **AREA_RATIO** and store these values in the fields created in step (e).

(g) Dissolve your **Intersect Feature Class** along the areal units defined in the input **Dissolve Field.** As part of this Dissolve, perform a **SUM summary statistic** on each **Statistic Field(s)** to reaggregate the data of interest along the new areal units.

Setting up the tool in ArcGIS

This AWR script runs using the ArcPy site package and is therefore only compatible with systems that are capable of importing ArcPy. Future development could be targeted towards creating an open source alternative. To use the tool in your own analysis, follow these steps:

(a) Download **<ArealWeightedReaggregation.py>** from http://parkerziegler.com/senior-research-programming-for-gis/, Post #3.

(b) Open up ArcGIS. From ArcCatalog, expand **Toolboxes**. Right click on **MyToolboxes** and create create a new **Toolbox (.tbx)**.

(c) Right click on your new toolbox. Navigate to **Add > Script**. This will pop open the script wizard.

(d) The first panel of the script wizard asks you to give the script **a name, an alias, and a description. DO NOT** use spaces in the script name – this is the actual path name that ArcGIS will use to find and run your script. Spaces and other special characters are ok to use in your script alias. Add a description if you like to remind you of how the tool works in the future.

(e) Continue on to the second panel of the script wizard. Here, you will point ArcGIS to the location of the actual Python script file on your hard drive.

(f) Continue on to the third (and final) panel of the script wizard. In this panel, you will set the **parameters** of the script (see **Description**) as well as the **properties** of each parameter. These are described in the **Parameters** section below.

(g) Once you've finished setting the parameters, select OK to finish creating the script. Your AWR tool is now ready for use in geoprocessing tasks.

Parameters

It is extremely important that you enter parameters in the exact order outlined here. ArcGIS relies on indexing each of the parameters in order to access the correct feature classes and fields. This means that if even one parameter is out of place, the script will fail to run or give an improper output. Moreover, each parameter requires that you explicitly state its properties. Properties include:

(a) Type – Required, Optional, or Derived.

If a parameter is **Required**, it must be input by the user in order for the script to run. If a parameter is **Optional**, it may be input by the user but isn't necessary for the script to run. If a

parameter is **Derived**, it is created somewhere in the course of the script but not input explicitly by the user.

(b) Direction – Input, Output.

A parameter will have direction **Input** if it refers to an existing feature class or field that the script will use. A parameter will have direction **Output** if it refers to a feature class or field that the script will create in the course of processing.

(c) Multivalue – Yes, No.

A parameter will be set to **Yes** for **MultiValue** if it accepts multiple inputs (either fields or feature classes). A parameter will be set to **No** for **Multivalue** if it only accepts one input. **(d) Default** – Set the **default** value for the parameter.

Use this property when you want to set a default value for a parameter upon load of the tool dialog box.

(e) Environment – Have the tool pull a particular property from the layer **environments**.

If you want the tool dialog box to pull this parameter from information stored in one of your environments, set this parameter to that environment (i.e. workspace, processing extent, snap raster).

(f) Filter – Select the **data type(s)** for the parameter.

When you want to limit a parameter to a particular **data type**, use this property to set that data type. This will influence what feature classes or fields the tool will allow you to select when choosing an input for this parameter from the tool.

(g) Obtained from – Connect this parameter with another parameter.

When a parameter is closely related to another parameter, use this property to establish the connection. For example, if you have a parameter that refers to a field in a feature class, set the **Obtained from** property to that feature class.

(h) Symbology – Import the **symbology** of another layer.

When you want to import the symbology from another layer or feature class, set this property to that layer or feature class.

Below are the list of parameters and the associated properties of each required for the AWR tool to run properly. Each entry is formatted as **Parameter Name | Type** – Properties.

Workspace | Workspace – Required, Input, No, <Null>, Workspace, Workspace, <Null>, <Null>

Data Feature Class | Feature Class – Required, Input, No, <Null>, <Null>, Polygon, <Null>, <Null>

Areal Unit Feature Class | Feature Class – Required, Input, No, <Null>, <Null>, Polygon, <Null>, <Null>

Dissolve Field | Field – Required, Input, No, <Null>, <Null>, *All Numerical Field Types*, Areal Unit Feature Class, <Null>

Intersect Feature Class | Feature Class – Required, Output, No, <Null>, <Null>, Polygon, <Null>, <Null>

Statistic Field(s) | Field – Required, Input, Yes, <Null>, <Null>, *All Numerical Field Types*, Data Feature Class, <Null>

Output feature Class | Field – Required, Output, No, <Null>, <Null>, Polygon, <Null>, <Null>

L i m i t a t i o n s

The areal-weighted reaggregation tool has several limitations. The following list outlines several of these, such that future developers might tackle these problems:

• **The distribution of the data of interest is assumed to be uniform within each polygon feature in the Data Feature Class.** The spatial distribution of people, pollutants, crimes, or any other data of interest is rarely (if ever) uniform, even across small extents. This tool relies on an assumption of even distribution in order to allocate areally-weighted data counts to the new areal units. A more precise algorithim would take into account the existing spatial heterogeneity of the data of interest to generate a more precise reaggregation.

• **The script generates a SUM summary statistic for the reaggregated areal features, but it does not perform other summary statistics.** When reaggregating data, we may want to know more than the sum of the data of interest in our new areal units. An expansion of this tool would allow users to set different statistical merge rules on different fields (i.e. Median, Standard Deviation).

• **The script relies on calling existing ArcPy functions, which are not open to user alteration.** A future script should use an open source architecture to extend the tool's useability.

F i g u r e s

Figure 1. The graphical user interface (GUI) of the Areal-Weighted Reaggregation tool.

Notice that the order of input parameters corresponds exactly to to the order in which we defined each parameter in the script wizard. Interested developers should feel free to design a custom GUI interface.

Figure 2. Setting the parameters of the Areal-Weighted Reaggregation tool.

The third panel of the script wizard (Parameters) allows you to set the parameters and associated parameter properties for the AWR tool. The top portion allows you to set the display name and data type. The bottom portion allows you to set the Direction, MultiValue, Default, Environment, Filter, Obtained from, and Symbology properties. For a more detailed description of all of these inputs, see the **Description** and **Parameters** sections.