

Scale Conversion and Interpolation

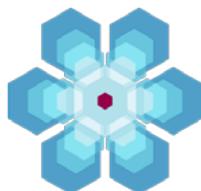
A SpaceStat Software Tutorial

Copyright 2013, BioMedware, Inc. (www.biomedware.com). All rights reserved.

SpaceStat and BioMedware are trademarks of BioMedware, Inc. SpaceStat is protected by U.S. patents 6,360,184, 6,460,011, 6,704,686, 6,738,729, and 6,985,829, with other patents pending.

Principal Investigators: Pierre Goovaerts and Geoffrey Jacquez . SpaceStat Team: Eve Do, Pierre Goovaerts, Sue Hinton, Geoff Jacquez, Andy Kaufmann, Sharon Matthews, Susan Maxwell, Kristin Michael, Yanna Pallicaris, Jawaid Rasul, and Robert Rommel.

SpaceStat was supported by grant CA92669 from the [National Cancer Institute](#) (NCI) and grant ES10220 from the [National Institute for Environmental Health Sciences](#) (NIEHS) to BioMedware, Inc. The software and help contents are solely the responsibility of the authors and do not necessarily represent the official views of the NCI or NIEHS.



BioMedware
Geospatial Research and Software

MATERIALS

[SpaceStat Tutorial Scale Conversion.SPT](#)

ESTIMATED TIME

20 minutes

OBJECTIVE

You will convert the spatial support of a population density dataset from county polygons to watershed polygons. You will also interpolate precipitation data from meteorological stations (point observations) to the same watersheds (polygons). These examples illustrate how simple converting data between geographic scales is with SpaceStat.

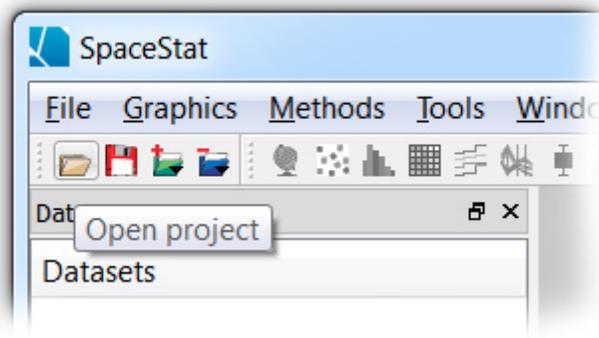
ABOUT THE DATA AND SPATIAL SUPPORT

The SpaceStat project for this tutorial contains data from three different original datasets. The first is an outline of the major watersheds of the northwestern part of the Lower Peninsula of Michigan. The second contains the population of Michigan counties from the 1990 US census. The third dataset contains precipitation measurements from 21 meteorological stations in operation in this area. The precipitation data records the average annual total precipitation at these stations from the period 1990-2000 and comes from the National Weather Service.

These data have three different spatial supports (underlying geographies): watershed boundary polygons, county boundary polygons, and meteorological station point locations. It is quite common when you have data coming from different sources that they will have different spatial supports but you will still like to conduct analyses that combine the different sources of data. Some types of analyses require the underlying data to have the same spatial support. This tutorial will show you some of the methods you can use to convert your data to the same spatial support to conduct such analyses.

STEP 1: OPEN THE STARTING PROJECT

The SpaceStat project containing these different datasets has already been created for you. Navigate through the menus to File-> Open Project or **click the Open Project button** as shown below.



Using the Open Project button

You can download this project to use from the following link:

[SpaceStat Tutorial Scale Conversion.SPT](#)

STEP 2: POLYGON TO POLYGON CONVERSION

After you have familiarized yourself with the data and perhaps made a few maps or plots to examine it, you are ready to convert the spatial support of the data. The county population data was gathered in census units and we'd like to convert the spatial support for it to our watershed units. This conversion is between two polygon geographies. The technique we'll use will estimate values in the watershed geography from those in the census geography without "borrowing strength" (interpolating) from nearby observations.

To begin the process, **select Methods->Scale Conversion/Interpolation** from the menus. You may want to maximize the SpaceStat window so that you can see the Task Manager easily. As the source dataset, **select the 'counties' geography and the 'POPDEN90' dataset. Select the 'watersheds' as the target geography.** If you choose to use geography centroids, the algorithm is much quicker but it is only allocating based on where the centroid of each polygon lies. Since we want to allocate based on the proportion of the census polygon that lies within the watershed polygon, **choose 'No' for using BOTH source and target geography centroids.** The default start and end time of the full time period are fine. Your settings should look as in the screenshot below:

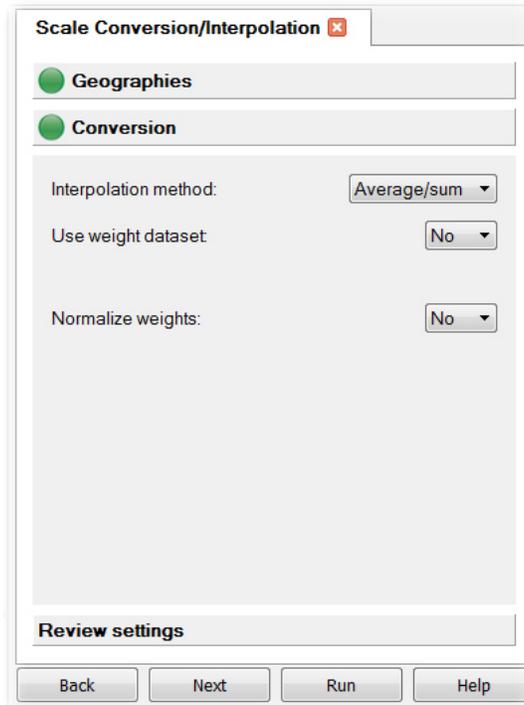
The screenshot shows a dialog box titled "Scale Conversion/Interpolation" with a close button (X) in the top right corner. The dialog is divided into three main sections: "Geographies", "Conversion", and "Review settings".

- Geographies:** This section contains several dropdown menus and buttons:
 - Source dataset: "counties" (dropdown) and "POPDEN90" (dropdown)
 - Use source geography centroids: "No" (dropdown)
 - Target geography: "watersheds" (dropdown)
 - Use destination geography centroids: "No" (dropdown)
 - Start time: "1-1-1985" (dropdown)
 - End time: "12-31-1995" (dropdown)
- Conversion:** This section is currently inactive, indicated by a greyed-out radio button.
- Review settings:** This section is also inactive, indicated by a greyed-out label.

At the bottom of the dialog, there are four buttons: "Back", "Next", "Run", and "Help".

Scale Conversion/Interpolation Settings for the first tab

Next you will specify how the scale conversion will be calculated. **Click the Conversion button** to bring up the conversion tab. Since we do not wish to borrow strength from surrounding areas now, **select 'Average/Sum'** as the interpolation method. Since you are calculating the population density in the watersheds you do not want to use a weight or normalize the result since these observations already account for the area. Make sure both of these options are set to 'No'. Your settings should match the following screenshot:



Scale Conversion/Interpolation Settings for the conversion tab

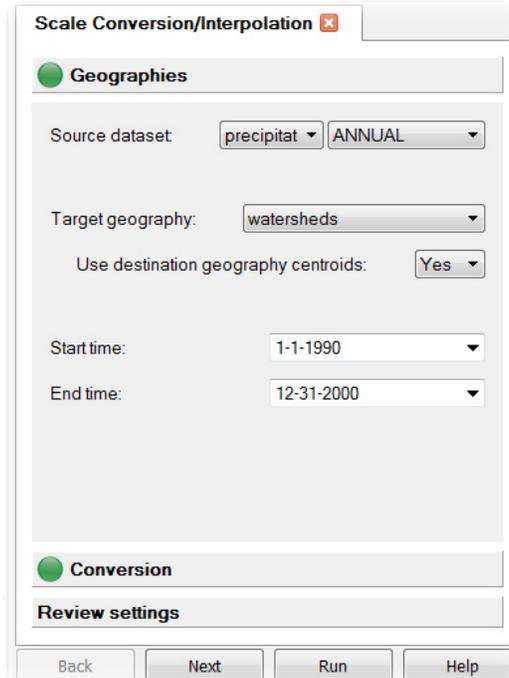
If your settings match these two dialogs, click the 'Run' button to begin the scale conversion. The method will create two datasets: a list of the polygon ID's in the counties that overlapped the watersheds and the population density in the watersheds. Find the dataset named 'POPDEN90 mean' in the watersheds geography. This is the dataset that you are concerned with. You can click on this dataset and then on the 'Create Map' button (it looks like a globe) in the toolbar to map this data.

STEP 3: INTERPOLATE FROM POINTS TO POLYGONS WITH NEAREST NEIGHBOR INTERPOLATION

The other scale conversion we'd like to perform is to convert the precipitation measurements from the sampling stations to the watershed level. For this conversion we need to go from a point geography to a polygon geography. Also, the sampling stations are discrete points and we will have to interpolate values to estimate values for the polygons. There are several ways to accomplish this including averaging over nearest neighbors, inverse distance weighting, and geostatistical kriging. You will use and compare the first two methods and kriging will be covered in later tutorials.

First you will use nearest neighbor interpolation. Start by **selecting Methods->Scale Conversion/Interpolation** from the menus. For the source **choose the 'precipitation' geography**

and the 'ANNUAL' dataset. For the target, again **choose the 'watersheds' geography.** SpaceStat can either assign interpolated values to either the polygon centroid or to the polygon itself based on the options in the 'Use destination geography centroids' dropdown. Here **keep the default of 'Yes'** for the 'Use destination geography centroids' dropdown. This will select the nearest number of neighbors to the centroid and average over these precipitation sampling stations. The default time settings are fine. Your settings should look as below:

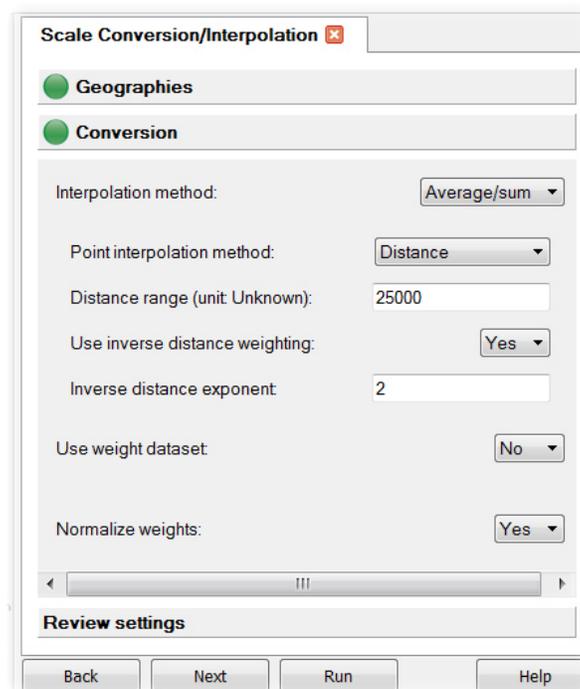


Scale Conversion/Interpolation Settings for the geographies tab

The default settings in the Conversion tab are fine: 'Average/Sum' for interpolation method, 'Nearest Neighbor' for point interpolation method, 32 for nearest neighbor count, 'Yes' for use inverse distance weighting, 1 for inverse distance exponent, 'No' for using weight dataset, and 'Yes' for normalize weights. **Click the 'Run' button** to assign the precipitation to the watershed geographies. The 'ANNUAL Mean' dataset in the watersheds geography is the watershed annual precipitation that you just calculated. Since we will also be calculating this data using the inverse distance interpolation, **rename the 'ANNUAL mean' dataset to 'PrecipNN mean' and rename the 'ANNUAL std dev' to 'PrecipNN std dev'**. To rename a dataset, click on the dataset to select it and then right-click and choose the 'Rename dataset' option. **Map the 'PrecipNN mean' dataset** to see the results.

STEP 4: INTERPOLATE FROM POINTS TO POLYGONS WITH INVERSE DISTANCE INTERPOLATION

If you have not closed the Task Manager window from the previous 'Scale Conversion/Interpolation' you can use this window again since the settings for the first tab have not changed. If you closed it, select Methods->Scale Conversion/Interpolation to reopen the method and fill out the settings as in Step 3. You will specify a different method in the Conversion tab. **Click the 'Conversion' button** to choose the Conversion tab. **Choose 'Distance' for the Point interpolation method. Enter 25000 for the Distance range and 2 for the Inverse distance exponent** (make sure that 'Yes' is selected for Use inverse distance weighting). This instructs SpaceStat to go out 25 kilometers from each polygon centroid weighting each observation by 1 over the square of the distance to the observation. This will give observations closest to the polygon centroid the largest weight. Check your settings with the screenshot below:



The screenshot shows the 'Scale Conversion/Interpolation' dialog box with the 'Conversion' tab selected. The settings are as follows:

- Interpolation method: Average/sum
- Point interpolation method: Distance
- Distance range (unit: Unknown): 25000
- Use inverse distance weighting: Yes
- Inverse distance exponent: 2
- Use weight dataset: No
- Normalize weights: Yes

At the bottom, there are buttons for 'Back', 'Next', 'Run', and 'Help', along with a 'Review settings' button.

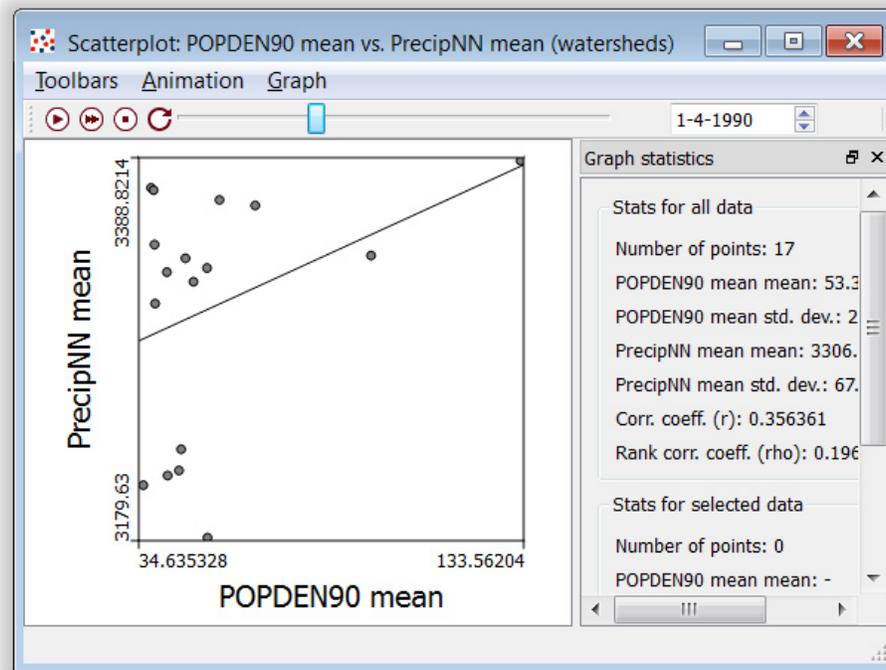
Scale Conversion/Interpolation Settings for the Conversion tab

After you verify your settings, **click the 'Run' button** to interpolate the data. **Rename the 'ANNUAL mean' dataset to 'PrecipID mean'. Map this dataset** to see the results.

STEP 5: IDENTIFY VLNERABLE WATERSHEDS

Our purpose behind these scale conversions was to locate watersheds which are vulnerable, i.e. have a high population but low precipitation. You can do this easily with brush selection. Make sure you have a map of the watersheds open. **Create a scatterplot** by pressing the 'Scatterplot' button

on the toolbar. Choose 'POPDEN90 mean' for the x-axis and 'PrecipNN mean' as the y-axis. Look at this scatterplot and use brush selection to identify the vulnerable watersheds. The most vulnerable watersheds would be in the bottom right of this scatterplot. You will have to advance the time slider to at least 1990 before data will be displayed since the precipitation data is not valid until 1990:



Scatterplot of watershed population against rainfall

STEP 6: NEXT STEPS

We used two methods to interpolate the precipitation measurements. To compare them you could create maps of each and visually look for differences or you could create a scatterplot of one versus the other.

It is a good practice to validate the interpolations you perform. One way of validating is to compare the estimates from the interpolation with the actual observations. The easiest way to accomplish this is to perform another scale conversion: here assigning the precipitation from the watershed polygons to the sampling station points (you will select 'Assignment' as the 'interpolation method'). Then you can create a scatterplot of the interpolated values versus the observations and see if there are any outliers. If the geographic scales between the two geographies increases, you would expect to see more outliers. Data that have large amounts of local variation would also show more outliers than data that is more continuous.