Basic Geostatistics: Pattern Description

A SpaceStat Software Tutorial

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_Tutorial_Basic_Geostatistics_Part_1.SPT

ESTIMATED TIME

15 minutes

OBJECTIVE

The Basic Geostatistics tutorials will guide you to using basic geostatistical procedures in SpaceStat. In this first tutorial, you will look at tools to describe the spatial patterns within your data. You will use the mapping and basic statistical graphs to look at the underlying patterns in the data as well the variogram cloud to assess how variability changes as distance increases among observations. The Basic Geostatistics tutorials will use data on the arsenic concentrations in Michigan groundwater.

ABOUT THE DATA

The SpcaeStat project for this tutorial contains data on the measured arsenic concentrations from well water. The data is split into 545 prediction wells and 2160 validation wells. Arsenic concentrations were measured over the period from 1990-2003. The geographic coordinates are the wells have been slightly modified to preserve confidentiality. The Michigan township boundaries for this area are also included in the project for visualization purposes. Our main goals with the data will be to: explore the spatial variability of the prediction set, fit alternate models of arsenic concentration to the prediction set, estimate the arsenic concentrations at the locations of the validation level, and compare the estimated values of the arsenic with the actual observations.

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.
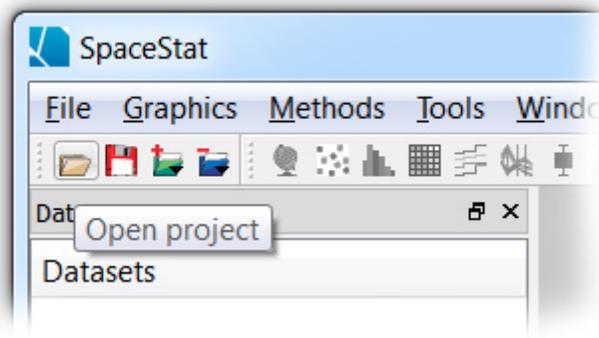
Using the Open Project button

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

SpaceStat_Tutorial_Basic_Geostatistics_Part_1.SPT

The first step to understanding the spatial patterns of the data is to create a map to visualize the data. First **right click on the 'Michigan Townships' entry in the Data view** (the Data view is the large panel on the left side of the SpaceStat interface). **Select 'Create Map' from the pop-up menu.** This will create a map with the polygon outlines of the Michigan townships.

Next you will add the data from the prediction wells to the map. To do this**, right click on the 'Arsenic (ppb)' entry** under the Prediction wells heading in the Data view (you might have to expand the Prediction wells heading by clicking on the small triangle to its left if you do not see the entries) and **select the 'Add to Map' -> 'Map: Michigan townships'** option. This will create a new layer on the map with points that show the arsenic values from the prediction wells. Take some time to explore the map and look for spatial patterns in the data. This map will also be used in Step 5, so be sure not to close it. Your map should look similar to the one below:
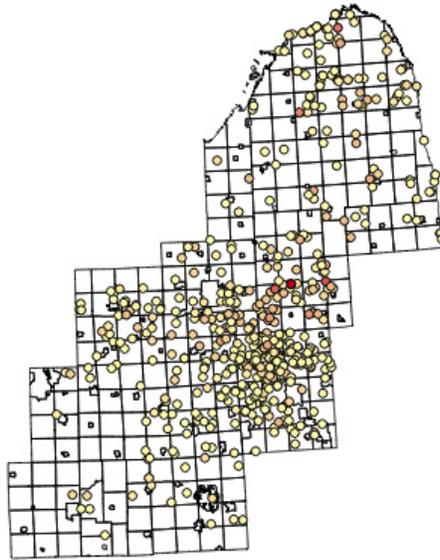
Map of arsenic at the prediction wells

STEP 3: CREATING HISTOGRAMS OF THE DATA

The next step in developing a description of how the well arsenic concentrations are distributed is to create a histogram of the values. **Press the Histogram button on the SpaceStat toolbar** (it is the icon that looks like a series of vertical bars – the third in the second grouping of icons in the SpaceStat toolbar). Make sure that **your geography is 'Prediction wells'** and **select 'Arsenic (ppb)' as the dataset** from the dropdown if it is not selected for you. **Click the 'Ok' button** after you have verified these two settings. Note that a panel of statistics is displayed by default to the right of the histogram that is created. The histogram should look like the figure below (you may want to maximize the window to more easily see the distribution):
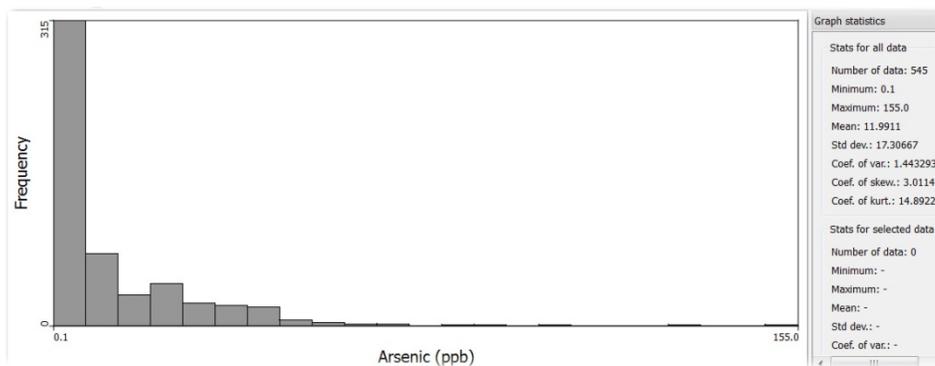
Histogram of arsenic concentrations

In the data that you have been provided there is also a transformed dataset for later analysis. The Log(Arsenic) is a log transformation of the arsenic concentrations. Repeats the steps above to create a new histogram, but this time **choose 'Log(Arsenic)' as the dataset.** The resulting histogram should look like the figure below:
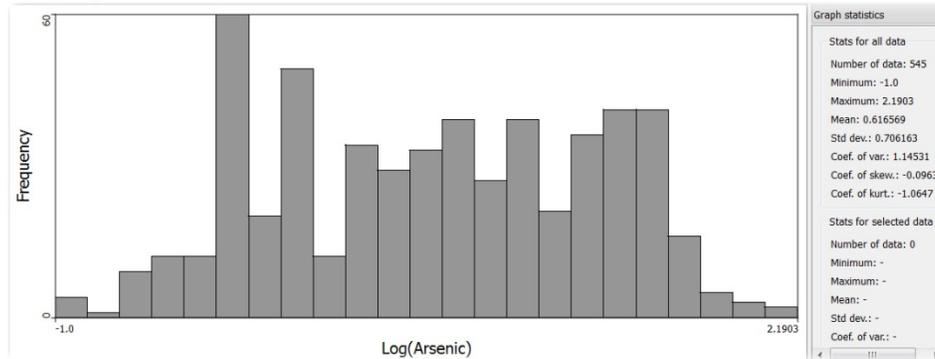
Histogram of log-transformed arsenic concentrations

Note the difference in the two distributions. Also pay particular attention to the coefficient of skewness shown in the 'Graph statistics' panel of both histograms. The concentration values are strongly positively skewed, but the log transformed values have a coefficient of skewness close to zero.

STEP 4: CREATING A VARIOGRAM CLOUD OF THE DATA

Variogram clouds are diagnostic tools that can be used to assess how the variability among points changes with increasing distance. Anomalies and non-homogeneous areas can be detected by looking at short distances that show high dissimilarities. With linking and brushing you can see where locations are on the map from Step 2.

Click the Varigram cloud button on the SpaceStat toolbar (this is the icon that looks like a few small blue dots merging into a gradient of solid blue). In the dialog that opens, make sure the **'Prediction wells' is the geography** and **'Arsenic (ppb)' is selected as the dataset**. Keep the default of the isotropic variogram cloud (do not check the 'Anisotropic' option). This assumes that the variation is comparable in all directions and should be used unless there is a directional trend to the data.

Next change the appearance of the variogram cloud. **Choose 'Properties' from the 'Graph' menu** of the variogram cloud window. The first change you will make is to **check the 'Fixed bounds' option**. Change the **horizontal maximum to 10000**. This will focus the window on observations that are up to 10km apart. Next towards the bottom of the dialog, **check the 'Connect point pairs in linked maps'** and **uncheck 'Show 2-D histogram'**. The first option draws a line connecting selected point pairs in the map when you do cartographic brushing. The second option shows a point for each pair

of observations instead of placing all pairs of observations into bins and using shade to indicate the number of observations in each bin. Check your settings with the screenshot below:
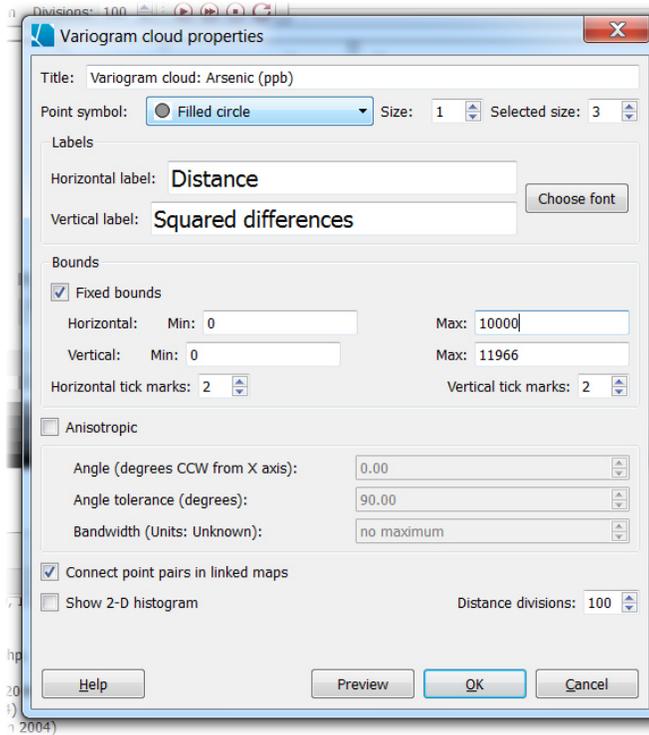
Variogram cloud settings

After you verify your settings, **click the 'Ok' button** to modify the variogram cloud appearance. Note that every possible pair of observations within 10km distance from each other is shown as a single point on the variogram cloud. The distance between the pair is shown on the x-axis and the squared difference in arsenic concentration is shown on the y-axis. Your variogram should now look similar to the following screenshot:
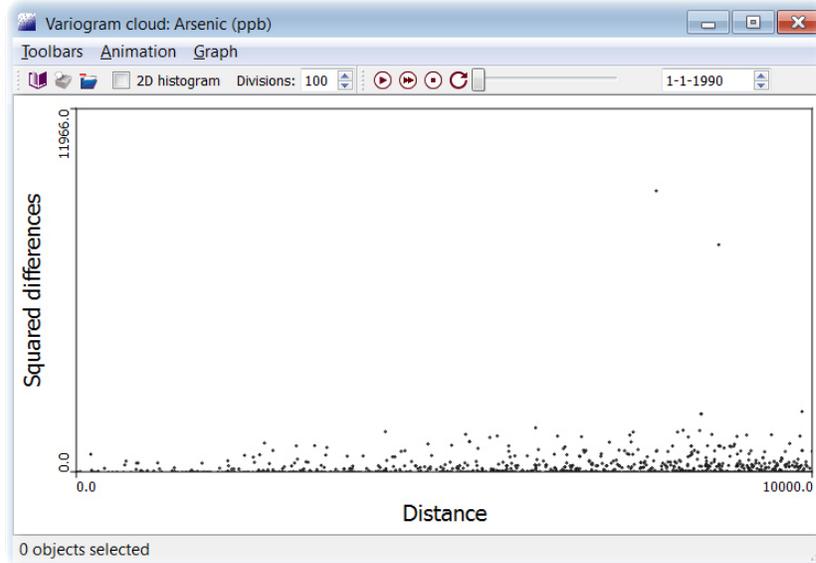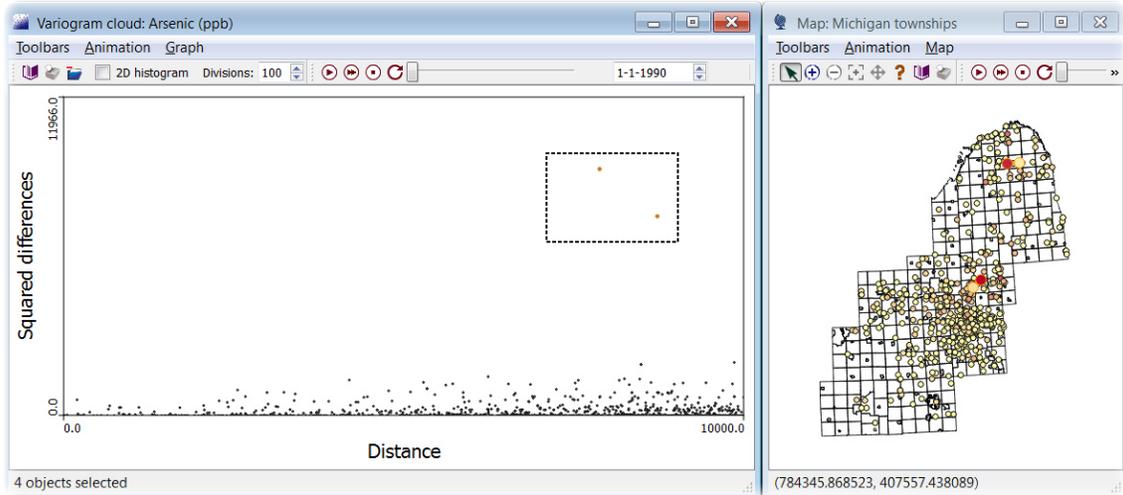
Variogram cloud

STEP 5: CARTOGRAPHIC BRUSHING WITH THE VARIOGRAM CLOUD

While the variogram cloud is a useful visualization on its own, using cartographic brushing allows you to understand where the patterns occur. For this you will need to have the map open that you created in Step 2. If you have closed this map, please revisit Step 2 to recreate the map for use in this step.

Cartographic brushing involves selecting objects in one visualization and seeing the same objects selected in another visualization. In this case you will select wells in the variogram cloud and see which wells are selected on the map. To do this you will need to rearrange the visualizations so that you can see both visualizations at the same time. You can either manually move and resize the windows so that you can see both or you can choose **Window->Tile** from the main SpaceStat menu (SpaceStat tiles all visualizations that are open so close any open histograms before you tile the windows).

Next click and drag on the variogram cloud. This will create a box, selecting anything inside. You can select the points that you wish to find more about. Remember that each point in the variogram cloud corresponds to a pair of wells. The wells will be shown on the map outlined in orange with a line connecting the two wells of a pair. Take some time to explore the variogram cloud. For instance in the following screenshot two outlying pairs have been selected. You can see that one pair is in the northern part of Michigan's Lower Peninsula and the other pair is near the center of the study area. The outliers correspond to pairs of wells that are relatively close together but have very different arsenic concentrations – potentially wells that you would want to investigate further. The next tutorial in this Basic Geostatistics series will show how to use kriging to estimate arsenic concentrations at the validation wells.

6

Catographic brushing on the variogram cloud