

Appendix 14. As a statistical tool within the ArcView (see Lee and Wong 2001, Hutchinson 2004), **CORRELATION** calculates the cross correlation between two input grids, prints the correlation coefficient to the screen, and sets an AML variable with the output value.

Discussion: Spatial autocorrelation is a measure of the similarity of objects within an area. An object can have at least two types of descriptive elements; 1) aspatial attributes (i.e., soil parameters, elevation, slope), and 2) spatial attributes (i.e., the x, y location in a specified coordinate system). From an object based view of spatial data, spatial autocorrelation measures the relationship between the difference of the aspatial attributes of objects with the distance between the objects.

For example, two objects which are close together and that have very similar aspatial descriptors are highly spatially correlated. Two objects close together that have very different aspatial descriptors are not very spatially autocorrelated (the objects are considered spatially independent). However, a negative spatial autocorrelation exists when objects that are close together have very different aspatial attributes than objects that are farther apart (Goodchild, 1986).

In the context of GRID, the objects correspond to cells and the aspatial descriptors or attributes correspond to cell values. The importance and usefulness of the spatial autocorrelation indices are:

The spatial autocorrelation index is a single value describing the spatial distribution of the objects in space.

It can be used to explore the cause of the spatial distribution of objects in space.

The index can be used to determine the degree of adjustment necessary when modeling phenomena (Goodchild, 1986).

The spatial autocorrelation index can aid in the process of extracting random, unbiased samples that will be fed into a classical statistical package.

GRID provides several spatial autocorrelation techniques. Two of the most widely accepted spatial autocorrelation indices are the Geary and Moran indices (both of which are available in GRID). The GEARY and MORAN indices for a grid test for spatial autocorrelation between directly adjacent cells in a grid and do not yield correlation between cells as a function of lag (beyond a cell's immediate neighborhood).

A third autocorrelation technique is incorporated in the CORRELATION atool. The CORRELATION atool measures the cross correlation between two grids, allowing for one of the grids to be offset with respect to the other. By using the same grid for both inputs, an (auto) correlation index as a function of lag can be computed. This technique

is particularly useful for determining the interval sample distance for purposes of random sampling.

The general notation used in correlation formulas and their GRID interpretation are the following:

n - the total number of cells in a grid: $NROWS * NCOLS$

i - any cell on the first input grid

j - any cell on the second input grid that is offset from i 's location by the specified x-, y-offset

z_i -the value of the attribute of cell i

z_j -the value of the attribute of cell j

\bar{z}_i -the mean value of the attribute of the first grid

\bar{z}_j -the mean value of the attribute of the second grid

c_{ij} -the similarity of i 's and j 's attributes: $(z_i - \bar{z}_i) * (z_j - \bar{z}_j)$

In the terms of the above notation, spatial autocorrelation is simply a measure of the attribute similarities in the set of c_{ij} with the locational similarities, and then summing the results into a single coefficient (Goodchild, 1986).

The formula for calculating the CORRELATION index is:

$$c = \frac{\sum_k^n c_{ij}}{\left(\sqrt{\sum_k^n (z_i - \bar{z}_i)^2} \right) * \left(\sqrt{\sum_k^n (z_j - \bar{z}_j)^2} \right)}$$

CORRELATION <grid1> {grid2} {xoffset} {yoffset}

Arguments

<grid1> - an input integer or floating-point grid.

{grid2} - an input integer or floating-point grid. If no second input grid is specified, the cross correlation will be performed on the first grid shifted on itself by the specified offset.

{xoffset} - the number of cells in the x-axis or direction to offset the second input grid from the first input grid. The default x-offset is zero cells.

{yoffset} - the number of cells in the y-axis or direction to offset the second input grid from the first input grid. The default y-offset is zero cells.

Notes:

- (1) The output from the CORRELATION atool is a coefficient (a floating-point value) that is printed to the screen and set into a global AML variable .correlation_out.
- (2) The spatial autocorrelation index is calculated best on non-categorical data.
- (3) The use of the grid.item syntax is not supported for this command. The input values must reside in the value attribute of the grid(s).
- (4) When the second input grid is the same as the first, the CORRELATION atool returns a spatial autocorrelation index for the input grid based on the specified x-, y-offset.
- (4) The resulting correlation coefficient will be from -1 to 1. If the two grids at the given offset are highly cross correlated the coefficient will equal one, if they are independent, zero, and if there is a strong negative correlation the output value will equal -1.

References

- Goodchild, M.F., 1986, Spatial Autocorrelation, Catmog 47, Geo Books, Norwich.
- Hutchinson, S., 2004, Inside ArcView GIS 8.3, Thomson/Delmar Learning, Clifton Park, NY, 510pp.
- Lee, J., and D.W.S. Wong, 2001, Statistical analysis with ArcView GIS, John Wiley and Sons, New York, 192pp.