

Spatial Analyst

Kriging

[related topics](#)
[open tool...](#)

Interpolates a surface from points using kriging.

[Learn more about how Kriging works](#)

Usage tips

Command line and Scripting

- Kriging is a processor-intensive process. The speed of execution is dependent on the number of points in the input dataset and the size of the search window.
- The Semivariogram Properties dialog box has several models to choose from. When the Kriging method is set to Ordinary, the available models are Spherical, Circular, Exponential, Gaussian, and Linear. When the Kriging method is set to Universal, the available models are Linear with linear drift and Linear with quadratic drift.
- The Universal Kriging types (Linear with linear drift and Linear with quadratic drift) assume that there is a structural component present and that the local trend varies from one location to another.
- The advanced parameters allow control of the semivariogram used for kriging. A default value for lag size is initially set to the default output cell size. For Major range, Partial sill, and Nugget, a default value will be calculated internally if nothing is specified.
- The optional output variance of prediction raster contains the kriging variance at each output raster cell. Assuming the kriging errors are normally distributed, there is a 95.5 percent probability that the actual z-value at the cell is the predicted raster value, plus or minus two times the square root of the value in the prediction raster.
- Some input datasets may have several points with the same X,Y coordinates. If the values of the points at the common location are the same, they are considered duplicates and have no effect on the output. If the values are different, they are considered 'coincident' points. Interpolation functions may handle this data condition differently, in some cases using the first coincident point encountered for the calculation, or sometimes the last. This may cause some locations in the output raster to have different values than what you might expect. The solution is to prepare your data by removing these coincident points.
- Low values within the output variance of prediction raster indicate a high degree of confidence in the predicted value. High values may indicate a need for more data points.

Map Algebra

- Kriging is a computer intensive process. The speed of execution is dependent on the number of points in the input dataset and the size of the search window. The GRAPH option is relatively quick and should be used before generating the final output surface to help determine the most appropriate options.
- GENERATE input files in the POINT format provide an easy method to create surfaces and two-dimensional coverages from ASCII files containing x,y coordinates and z-values. The x,y coordinates and z-values can be expressed as integers or single- or double-precision floating-point values. The x,y coordinates for the output raster are stored in DOUBLE precision. The z-values are stored in SINGLE precision. GENERATE files in the POINT format contain the feature ID, x,y coordinate, and z-values for point features. The ID values are not used by the kriging function. The format for GENERATE files containing points is:

```
ID, X, Y, Z
...
END
```

The end of the file is indicated with the END keyword. A typical GENERATE file containing points is shown below. The file can be delimited by a comma or space.

```
1, 234.065, 837.123, 125.845
2, 735.774, 729.345, 93.739
37, 840.934, 385.700, 157.037
5, 385.888, 286.602, 25.764
END
```

- The barriers option is used to specify the location of linear features known to interrupt the surface continuity. These features do not have z-values. Cliffs, faults, and embankments are typical examples of barriers. Barriers limit the selected set of the input sample points used to interpolate output z-values to only those samples on the same side of the barrier as the current processing cell. Input sample points that lie exactly on the barrier line will be included in the selected sample set for both sides of the barrier.
- The barriers options will significantly extend the processing time required to interpolate the output raster.
- Barrier features can be input as either a coverage or a GENERATE file in the LINE format. Kriging will look for the {barrier_cover} first, and if not found, it will look for the {barrier_file}. Kriging only uses the x,y coordinates for the linear feature; therefore, it is not necessary to provide z-values for the left and right sides of the barrier. Any z-values provided will be ignored. LINE input files contain the feature ID, followed by the x,y coordinates for each vertex in the file. Each line feature must have at least two, and no more than 20,000, vertices. The general LINE input file format is:

```
ID
X, Y, (Z)
...
X, Y, (Z)
END
END
```

The end of each line is indicated with the END keyword, and the end of the file is indicated with a second END keyword. The following is an example of a 3D GENERATE file in the LINE format. In this case, only the x,y coordinates are used by Kriging. The feature IDs and z-values are ignored.

```

345321.5 678543.1 123
345879.0 689231.2 145
345999.9 690579.3 157
END
2
367809.0 712004.0 209
368098.9 710776.8 214
369890.6 709768.2 206
370023.4 716098.3 198
END
2
390879.0 602090.0 133
391000.6 606668.2 169
397547.3 609608.4 125
END
END

```

- The {output_variance} contains the kriging variance at each output raster cell. Assuming the kriging errors are normally distributed, there is a 95.5 percent probability that the actual z-value at the cell is the predicted raster value \pm two times the square root of the value in the {output_variance}, that is:

```
<out_grid>  $\pm$  (2 * SQRT{output_variance})
```

- When using the <num_points> method, the resolution of the output raster and {output_variance} is determined by dividing the longest dimension of the extent of the input sample points by the specified number of sample points. The aspect ratio of the input point cover or file is automatically preserved by applying the mesh-point spacing derived along the longest side to determine the number of mesh points required on the shorter side. Therefore, unless the extent of the input sample points forms a square, the number of points in the x and y coordinates will be different. Low values within the {output_variance} raster indicate a high degree of confidence in the predicted value. High values may indicate a need for more data points.
- You can accept the input point cover or file XMIN,YMIN as the default origin and the XMAX,YMAX as the upper right corner of the extent to generate a raster covering the full extent of the input sample points. Alternatively, you can specify the out_grid origin, extent, and resolution or mesh-point spacing to create a raster covering only a portion of the input sample points.
- The GRAPH and BOTH options write a set of semivariance data to an INFO table called out_grid.SVG. A semivariogram graphically illustrates how the actual semivariance and the predicted semivariance vary with the distance between the pairs of original sample points. Four items are written to the out_grid.SVG file. They are defined as follows:

NAME	WIDTH	OUTPUT	TYPE	N.DEC
METHOD	4	8	B	-
TYPE	6	6	C	-
DISTANCE	8	16	F	4
SEMIVARIANCE	8	16	F	4

- METHOD — An integer code value specifying the type of mathematical function used to model the semivariance. The code values are:

CODE	{method}
1	SPHERICAL
2	CIRCULAR
3	EXPONENTIAL
4	GAUSSIAN
5	LINEAR

- TYPE —The type of data represented by the record.
The semivariance INFO table contains two types of data: ACTUAL and FITTED. ACTUAL records contain the actual semivariance values. FITTED records contain the predicted semivariance values calculated by the specified modeling {method}.
- DISTANCE —The distance between pairs of input sample points at which the semivariance values have been measured.
The actual and predicted semivariance is calculated at each interval of the mesh-point spacing. The maximum DISTANCE value is equal to twice the specified RADIUS {r} distance, or it is the maximum distance between pairs of input sample points.
- SEMIVARIANCE — The actual or predicted semivariance value at the current DISTANCE.
- The graphing option is not available with the UNIVERSAL1 and UNIVERSAL2 kriging methods. Universal kriging assumes that there is a structural component present and that the local trend varies from one location to another. Plotting a semivariogram for each input sample point is not practical.
- The input feature dataset in the Raster Calculator can only be a point shapefile.
- Some input datasets may have several points with the same X,Y coordinates. If the values of the points at the common location are the same, they are considered duplicates and have no affect on the output. If the values are different, they are considered 'coincident' points. Interpolation functions may handle this data condition differently, in some cases using the first coincident point encountered for the calculation, or sometimes the last. This may cause some locations in the output raster to have different values than what you might expect. The solution is to prepare your data by removing these coincident points.
- Kriging does not respect the Mask setting.
- Learn more about [how to specify the input feature dataset](#) in the Map Algebra expression of Raster Calculator.

Command line syntax

[An overview of the command line window](#)

```
Kriging_sa <in_point_features> <z_field> <out_surface_raster> <SPHERICAL | CIRCULAR | EXPONENTIAL | GAUSSIAN | LINEAR | LINEARDRIFT | QUADRATICDRIFT | NONE> {Lag size} {Major range} {Partial sill} {Nugget} {cell_size} {search_radius} {out_variance_prediction_raster}
```

Parameter	Explanation	Datatype
<in_point_features>	The input point features containing the z-values to be interpolated into a surface raster.	Composite Geodataset

<z_field>	Field that holds a height or magnitude value for each point. This can be a numeric field or the Shape field if the in_point_features contains z-values.	Field
<out_surface_raster>	Output surface raster dataset.	Raster Dataset
<SPHERICAL CIRCULAR EXPONENTIAL GAUSSIAN LINEAR LINEARDRIFT QUADRATICDRIFT NONE>	Semivariogram model to be used. <ul style="list-style-type: none"> • SPHERICAL — Spherical semivariogram model. This is the default. • CIRCULAR — Circular semivariogram model. • EXPONENTIAL — Exponential semivariogram model. • GAUSSIAN — Gaussian (or normal distribution) semivariogram model. • LINEAR — Linear semivariogram model with a sill. • LINEARDRIFT — Universal kriging with linear drift. • QUADRATICDRIFT — Universal kriging with quadratic drift. 	SemiVarioqram
{Lag size} {Major range} {Partial sill} {Nugget}	These advanced parameters must be enclosed in double quotation marks along with the semivariogram model option. <ul style="list-style-type: none"> • Lag size — The default is the output raster cell size. • Major range — Represents a distance beyond which there is little or no correlation. • Partial sill — The difference between the nugget and the sill. • Nugget — Represents the error and variation at spatial scales too fine to detect. The nugget effect is seen as a discontinuity at the origin. 	
{cell_size}	The cell size at which the output raster will be created. This will be the value in the environment if it is explicitly set. Otherwise, it is the shorter of the width or the height of the extent of in_point_features in the input spatial reference, divided by 250.	Analysis cell size
{search_radius}	Defines which surrounding points will be used to control the raster. There are two options: VARIABLE and FIXED. Variable is the default. <ul style="list-style-type: none"> • VARIABLE {number_of_points} {maximum_distance} <p>{number_of_points} — An integer value specifying the number of nearest input sample points to be used to perform interpolation. The default is 12 points.</p> <p>{maximum_distance} — Specifies the distance, in map units, by which to limit the search for the nearest input sample points. If the number of points for the VARIABLE option cannot be satisfied within that maximum distance, a smaller number of points will be used.</p> • FIXED {distance} {minimum_number_of_points} <p>{distance} — The distance, in map units, specifying that all input sample points within the specified radius will be used to perform interpolation. The default radius is five times the cell size of the output raster.</p> <p>{minimum_number_of_points} — An integer defining the minimum number of points to be used for interpolation. If the required number of points is not found within the specified radius, the search radius will be increased until the specified minimum number of points is found. The default is zero.</p> 	Radius
{out_variance_prediction_raster}	Output variance of the prediction raster dataset.	Raster Dataset

[Data types for geoprocessing tool parameters](#)

Command line example

```
Kriging_sa C:/data/pts.shp z_value C:/data/final_1 "Spherical 0.4496 2.6185 542.65 0" 1 "VARIABLE 12" ""
```

Scripting syntax

[An overview of scripting in geoprocessing](#)

```
Kriging_sa (in_point_features, z_field, out_surface_raster, semiVarioqram_props, {Lag size} {Major range} {Partial sill} {Nugget}, cell_size, search_radius, out_variance_prediction_raster)
```

Parameter	Explanation	Datatype
in_point_features (Required)	The input point features containing the z-values to be interpolated into a	Composite Geodataset

	surface raster.	
z_field (Required)	Field that holds a height or magnitude value for each point. This can be a numeric field or the Shape field if the in_point_features contains z-values.	Field
out_surface_raster (Required)	Output surface raster dataset.	Raster Dataset
semiVariance_props (Required)	Semivariogram model to be used. <ul style="list-style-type: none"> • SPHERICAL — Spherical semivariogram model. This is the default. • CIRCULAR — Circular semivariogram model. • EXPONENTIAL — Exponential semivariogram model. • GAUSSIAN — Gaussian (or normal distribution) semivariogram model. • LINEAR — Linear semivariogram model with a sill. • LINEARDRIFT — Universal kriging with linear drift. • QUADRATICDRIFT — Universal kriging with quadratic drift. 	SemiVariance
{Lag size} {Major range} {Partial sill} {Nugget} (Optional)	These advanced parameters must be enclosed in double quotation marks along with the semivariogram model option. <ul style="list-style-type: none"> • Lag size — The default is the output raster cell size. • Major range — Represents a distance beyond which there is little or no correlation. • Partial sill — The difference between the nugget and the sill. • Nugget — Represents the error and variation at spatial scales too fine to detect. The nugget effect is seen as a discontinuity at the origin. 	
cell_size (Optional)	The cell size at which the output raster will be created. This will be the value in the environment if it is explicitly set. Otherwise, it is the shorter of the width or the height of the extent of in_point_features in the input spatial reference, divided by 250.	Analysis cell size
search_radius (Optional)	Defines which surrounding points will be used to control the raster. There are two options: VARIABLE and FIXED. Variable is the default. <ul style="list-style-type: none"> • VARIABLE {number_of_points} {maximum_distance} <ul style="list-style-type: none"> {number_of_points} — An integer value specifying the number of nearest input sample points to be used to perform interpolation. The default is 12 points. {maximum_distance} — Specifies the distance, in map units, by which to limit the search for the nearest input sample points. If the number of points for the VARIABLE option cannot be satisfied within that maximum distance, a smaller number of points will be used. • FIXED {distance} {minimum_number_of_points} <ul style="list-style-type: none"> {distance} — The distance, in map units, specifying that all input sample points within the specified radius will be used to perform interpolation. The default radius is five times the cell size of the output raster. {minimum_number_of_points} — An integer defining the minimum number of points to be used for interpolation. If the required number of points is not found within the specified radius, the search radius will be increased until the specified minimum number of points is found. The default is zero. 	Radius
out_variance_prediction_raster (Optional)	Output variance of the prediction raster dataset.	Raster Dataset

[Data types for geoprocessing tool parameters](#)

Script example

```
# Kriging_sample.py
# Description:
# Interpolate a series of point features onto a rectangular raster using kriging.
# Requirements: None
# Author: ESRI
# Date: Sept 6, 2005

# Import system modules
import arcgisscripting

# Create the Geoprocessor object
gp = arcgisscripting.create()

try:
    #Set the input feature dataset
```

```

inputFeatureDataset = "C:/data/pts.shp"

#Set the output raster name
outputRaster = "C:/data/final_1"

#Don't write out the variance raster
Output_variance_of_prediction_raster = ""

#Set the output extent
gp.Extent = "0 0 6 6"

#Define the semivariogram
Model = "Spherical "           #(Remember the space at the end)
LagSize = "0.4496 "           #(Remember the space at the end)
MajorRange = "2.6185 "        #(Remember the space at the end)
PartialSill = "542.65 "       #(Remember the space at the end)
Nugget = "0"
Semivariogram = Model + LagSize + MajorRange + PartialSill + Nugget

# Set the attribute field
attributeName = "z_value"

#Check out Spatial Analyst extension license
gp.CheckOutExtension("Spatial")
#Process: Kriging...
gp.Kriging_sa(inputFeatureDataset, attributeName, outputRaster, Semivariogram, _
              "1", "VARIABLE 12", Output_variance_of_prediction_raster)
except:
  # If an error occurred while running a tool, then print the messages
  print gp.GetMessages()

```

Map Algebra syntax

Kriging(<point_feature | point_file>, {spot_item}, {barrier_cover | barrier_file}, {BOTH | GRAPH | GRID}, {output_variance}, {method}, {SAMPLE, {num_points}, {max_radius}}, {cellsize}, {xmin, ymin, xmax, ymax})

Kriging(<point_feature | point_file>, {spot_item}, {barrier_cover | barrier_file}, {BOTH | GRAPH | GRID}, {output_variance}, {method}, {RADIUS, {radius}, {min_points}}, {cellsize}, {xmin, ymin, xmax, ymax})

Parameter	Explanation
<point_feature point_file>	The source of the data containing points with z-values to be converted to a raster surface. <ul style="list-style-type: none"> point_feature — A point coverage or shapefile with z-values stored in an item. point_file — An ASCII GENERATE file in the POINT format.
{spot_item}	Name of a numeric item in the <in_feature>.PAT or a related table containing the point values. The default item name is SPOT. Use the # symbol as a placeholder when the point locations are input using a <point_file>.
{barrier_feature barrier_file}	The source of barrier information used for limiting the line search for input sample points. <ul style="list-style-type: none"> barrier_feature — A coverage or shapefile containing arcs. barrier_file — An ASCII file of lines in GENERATE format.
{BOTH GRAPH GRID}	Keyword specifying whether a raster, an INFO datafile containing semivariance values suitable for graphing, or both, will be created. <ul style="list-style-type: none"> BOTH — Both an INFO datafile containing semivariance data as well as the specified out_grid and optional {output_variance} raster will be created. The INFO datafile will be named out_grid.SVG. If out_grid.SVG already exists, the information will be appended to it; otherwise, the new file will be created. This is the default. GRAPH — Only an INFO datafile containing semivariance data for plotting the semivariogram will be created. The INFO datafile will be named out_grid.SVG. No {output_variance} raster will be created. If out_grid.SVG already exists, the information will be appended to it; otherwise, the new file will be created. GRID — Only the out_grid and optional {output_variance} raster will be created.
{output_variance}	Name of the optional out_grid containing predicted semivariance values for each mesh point in <out_grid>. If the GRAPH option is specified, an {output_variance} raster will not be created.
{method}	Keyword specifying the type of mathematical function used to model the semivariance. An out_grid.SVG file can also be used. <ul style="list-style-type: none"> SPHERICAL — Spherical semivariogram model. This is the default. CIRCULAR — Circular semivariogram model. EXPONENTIAL — Exponential semivariogram model. GAUSSIAN — Gaussian or normal distribution semivariogram model. LINEAR — Linear semivariogram model with a sill. UNIVERSAL1 — Universal kriging with linear drift. The GRAPH and BOTH options

are not available with the UNIVERSAL1 method.

- UNIVERSAL2 — Universal kriging with quadratic drift. The GRAPH and BOTH options are not available with the UNIVERSAL2 method.

An out_grid.SVG file allows you to make several kriging runs with a variety of methods and cell sizes, then use that SVG file as the semivariogram model.

SAMPLE	Keyword indicating the method used to select neighboring input sample points for interpolating the value of each cell in the output raster.
{num_points}	An integer value specifying the number of the nearest input sample points to be used to perform interpolation. A commonly used value is 12 points.
{max_radius}	Specifies the distance, in map units, within which to limit the search for the nearest input sample points. If the number of points for the SAMPLE option cannot be satisfied within the {max_radius}, a smaller number of points will be used.
RADIUS	Keyword indicating that a search radius will be used to select neighboring points for interpolating the value for each cell in the output raster.
{radius}	The distance, in map units, specifying that all input sample points within the specified radius will be used to perform interpolation. The default radius is five times the cell size of the output raster.
{min_points}	An integer value defining the minimum number of points to be used for interpolation. If the required number of points is not found within the specified radius, the search radius will be increased until the specified minimum number of points is met. The default is zero.
{cellsize}	The width or height of a cell in map units. Cell size defaults to the current environment setting. If no {cellsize} is set, the cell size on the Cell Size tab on the ArcGIS Spatial Analyst Options dialog box must be set to establish the current environment setting.
{xmin, ymin, xmax, ymax}	The dimensions, in map units, of the window in which to perform the interpolation. Defaults to the current environment settings. If no environment is set, the default is the boundary of the input extended outward 0.5 times the cell size.

Map Algebra example

```
kriging(cover)
kriging(file, #, #, graph, #, circular)
kriging(cover, item, #, #, out_var, linear, sample, 9)
kriging(cover, #, #, #, #, gaussian, #, #, #, 36)
kriging(file, #, #, grid, #, universal2, radius, 400)
```

ArcObjects syntax

See [Krige](#) and [Variogram](#)