Examples  ➤ Interpolating Continuous Elevation Data From Sparse Point Data Using the Krige Method

Dialog Box  ➤ Using the operation dialog box interface
➤ Using the dialog box interface to create or edit scripts

The Krige operation is a data interpolation operation. Given a data set of discrete points (e.g., sample data rather than whole population data), the Krige operation is used to find a function which best represents the surface so that unknown data points can be predicted based on known data points. In other words, the Krige operation will predict the whole population from the sample set, thereby creating a continuous data set from a sparse data set. It is important to have knowledge of the behaviour and smoothness of the data before choosing to use this operation.

Sparse Data Map
The Sparse Data Map drop-down list is used to specify a map layer of sparse point data (<1000 points) with a background of “VOID” cells with values to be estimated.
Output Precision
The Output Precision field and drop-down list is used to specify the precision and data units of the values to be interpolated. Units can be None, mm, cm, m, and km. The units of measure specified will be the data units of the output map layer. If a decimal point is used, the data of the resulting map layer will be of type floating point. For example, for floating point results to be within 1cm, specify 0.01m or 1.0cm. For fixed point results to be within 1cm, specify 1cm.

Mask Map
Use the Mask Map drop-down list to specify a map layer that restricts the interpolation to the cells in the mask map layer that have a non-void value. All cells that have the value “VOID” in the mask map will not be included in the interpolation.

Data Model
Choose the Spherical equation to estimate the variogram for a complexly varying data surface or the Exponential equation to estimate the variogram for a smoothly varying data surface.

Range
The Range field is used to define how close a known point must be to the point to be predicted to be included in the construction of the surface. The Range value defines the Neighbourhood of the target cell. You must know the nature of the data being mapped to correctly specify a Range value; however, in general, the larger the Range value, the smoother results but the greater the dependence on values far away from the target cell. The default value for the Range is the larger of either the number of rows or the number of columns. The default Range value is the greater of the number of rows or number of columns.

Syntax
- Syntax and type conventions
- Using the Script window interface
- Using the dialog box interface to create or edit scripts

Kringe map [In area map]
    [To vertical distance]
    [Spherical | Exponential]
    [Range value];

Kringe map
The Krige statement specifies a map layer of up to 1000 sparse data points on a background of cells with the value “VOID” for which values are to be interpolated.
In Area map
The In modifier is used to specify a masking map layer that will restrict interpolation to specific cells. For example, it could be used to restrict the interpolation to non-urban areas only. Non-VOID cells in the masking map layer identify those cells that are to be interpolated. VOID cells in the masking map layer will result in corresponding VOID cells in the output map layer.

To Vertical distance
The To modifier is used to specify the data type and precision of the output map layer. If the parameter contains a decimal point, the resulting map layer will be of type floating point; otherwise, the resulting map layer will be of type fixed point. For example, for floating point results to be within 1cm, specify 0.01m or 1.0cm. For fixed point results to be within 1cm, specify 1cm. The units of measure specified will be the data units of the output map layer.

Spherical | Exponential
Choose the Spherical equation to estimate the variogram for a complexly varying data surface or the Exponential equation to estimate the variogram for a smoothly varying data surface.

Range
The Range modifier is used to define how close a known point must be to the point to be predicted to be included in the construction of the surface. The Range value defines the Neighbourhood of the target cell. You must know the nature of the data being mapped to correctly specify a Range value; however, in general, the larger the Range value, the smoother results but the greater the dependence on values far away from the target cell. The default value for the Range is the larger of either the number of rows or the number of columns. The default Range value is the greater of the number of rows or number of columns.

Details
Recommendation
We strongly recommend that you have an understanding of the Kriging method of data interpolation and its inherent problems before using the results of this operation for critical applications. You should be aware that data interpolation leads to the potential for error since the interpolated values may not reflect the true values.

Also, the choice of interpolation method is dependent on the nature of the data. We recommend that you have a good understanding of the variability, trend, and spatial correlation of the phenomena you are mapping.
Use the Krige operation only if you have a high degree of confidence in the exactness of the original data. For data of low confidence, use the Interpolate operation followed by the Scan Average operation to smooth the results. Use the Krige operation when there is a finite range at which a point no longer has any influence over the values of the unknown points; when the influence of any point is not a function of distance but rather a function of the complexity/smoothness of the surface. The degree of smoothness of the data is important to the choice. This implies that you have knowledge of the behaviour of the surface being modelled.

If you do not have knowledge of the nature of the data to be estimated and the strengths and weaknesses of the Krige method, we recommend that you obtain professional training in the field related to the data you are analyzing and that you either consult a statistician familiar with the Krige method or obtain training in spatial statistical analysis before using the results of the Krige operation in critical applications.


**Processing Speed**

Speed of processing is also a consideration when choosing which interpolation method to use. Some researchers suggest that the choice of an interpolation method should be made on the basis of ease of computation since there is very little difference among the results of the various methods, and all produce equally suspect results. Others argue that Kriging provides a marginal advantage that, when multiplied against dollars of return based on the accuracy of the predictions, offsets the disadvantages of the slower processing speed.

**Kriging and Data Interpolation**

Kriging is a regression based interpolation method that is used to predict unknown values from irregularly spaced known values. It was originally developed for mapping in the fields of Geology and Geophysics, mining, and photogrammetry.

Kriging takes into account the interdependency of samples that are close to each other while allowing for a certain independence of the sample points. It avoids the building of a surface based on trends with introduced randomness. Kriging is based on the structural characteristics and behaviour of spatially located data. Samples taken closer together are expected to be more alike than samples taken farther apart because points that are close
together tend to be strongly correlated whereas, points that are far apart tend
not to be correlated.

The weights applied to the known values are obtained from a system of
linear equations in which the coefficients are the values of variograms or
covariance functions. The functions calculate the correlation between
known points or known and unknown points. To obtain the function, the
variance error must be minimized. The variogram yields the size of the zone
of influence, the isotropic nature of the variable, and the continuity of the
variable through space.

**What Do I Need?**
Use the **Kriging** operation on map layers of sparse data points (<1000) for
which you wish to create a continuous data set. Often continuously varying
data such as soil type, population density, or topography is collected as a set
of discontinuous, unevenly spaced data points. The non-data cells in the
input map layer should be assigned the value “VOID”. For contour data or
incomplete data sets, such as remote sensing images with data gaps, use the
**Interpolate** operation to interpolate unknown values.

**Troubleshooting Error Messages**
Here are some of the most common error messages for the **Kriging** operation
with suggestions on what to do if you see them:

**Error, the original map data units are not compatible with
the units specified in the “To” modifier.**
Either the data units in the input map layer Information window is
missing or is not compatible with the units specified by the **To**
modifier, or the units specified by the **To** modifier is missing (set to
no units). Either change the input map layer units in the Information
window to be compatible with the units of the **To** modifier or change
the units of the **To** modifier to be compatible with the input map
layer data units.

**Error, negative or zero distance/resolution specified.**
The **To** modifier must specify a value greater than zero.

**Error there are too many points in the input map for the
kriging algorithm (maximum =1000).**
Use the **Random** operation to reduce the number of data points. The
number of known points in the input map layer exceeds the
maximum number allowed. The ability of the algorithm to produce
results in a reasonable period of time is questionable. Either reduce
the number points using the **Random** operation, or use the **Interpolate** operation instead.

**Sorry, the map has exceeded the maximum number of zones allowed (65 000).**

The data range and level of precision is too high (exceeds the maximum number of zones). Try reducing the precision specified by the **To** modifier.