GMS 10.5 Tutorial
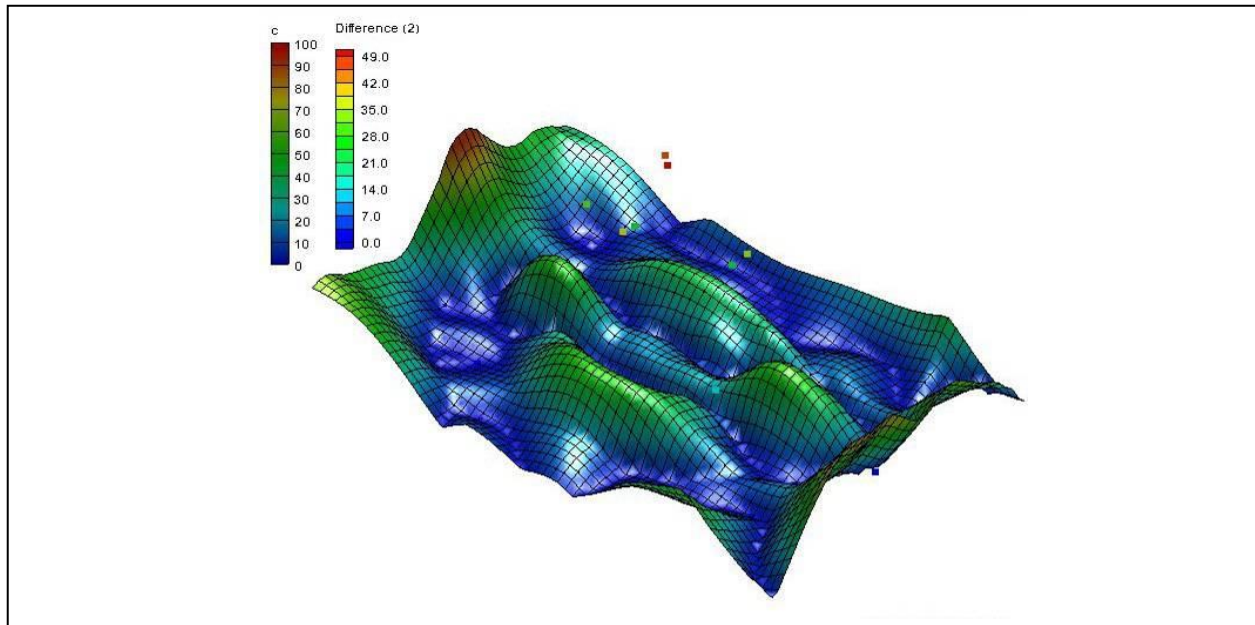
Geostatistics – 2D

Learn the various 2D interpolation methods available in GMS

Objectives

Learn to create scatter point sets from scratch and how to import scatter sets from a file. Investigate the various 2D interpolation methods available in GMS, including linear, inverse distance weighted (IDW), and kriging.

Prerequisite Tutorials

- None

Required Components

- Grid Module
- Geostatistics

Time

- 30–45 minutes

AQUAVEO™

1   Introduction

Two-dimensional geostatistics (interpolation) can be performed in GMS using the 2D Scatter Point module. The module is used to interpolate from sets of 2D scatter points to any of the other object types (meshes, grids, TINs). Several interpolation schemes are supported, including kriging. Geostatistics are useful for setting up input data for analysis codes or for site characterization.

The tools for manipulating scatter point sets and the interpolation schemes supported in GMS are demonstrated and discussed in this tutorial. This includes how to create a scatter point set from scratch, how to import scatter points from a file, creating a bounding grid, interpolating the scatter points using various interpolation methods, and using the data calculator to compare interpolations. The interpolation schemes presented in this tutorial are defined in the *Interpolation* section of the *GMS Online Help*.

1.1   Getting Started

Do the following to get started:

1.   If necessary, launch GMS.

2.   If GMS is already running, select the *File* / **New** command to ensure that the program settings are restored to their default state.

2   Creating a Scatter Point Set

Interpolation in GMS is performed using scatter points. A set of 2D scatter points is defined by a set of *xy* coordinates. A group of scatter points is called a scatter point set. Each scatter point set may also have a dataset associated with it. Datasets are scalar or vector values—one value at each scatter point.

The scatter points represent the geometric objects while the datasets represent the data values at the points. The datasets may represent anything: elevation, heat measurements, concentration of a chemical, etc. The datasets can be interpolated to a TIN, mesh, or grid.

The **Create Scatter Points** ⋮ tool is used to create 2D scatter point sets in GMS by using the following steps:

1. Right-click on an empty space in the Project Explorer and select *New* / **2D Scatter Point Set**.

A new "⬚ set (1)" scatter set will appear.

2. Using the **Create Scatter Points** ⋮ tool, create a number of new scatter points by clicking several times in the Graphics Window.

When creating each point, GMS automatically assigns a dataset value at the point. The settings can be changed so that GMS gives a prompt for the dataset values.

3. Select *Scatter Points* / **Scatter Point Settings…** to bring up the *Preferences* dialog. Notice that "Scatter sets" is selected in the list on the left.

4. In the *2D scatter* section, turn on *Confirm dataset value* and click **OK** to close the *Preferences* dialog.

5. Click anywhere in the Graphics Window to create a new scatter point and bring up the *Scatter Point Dataset Value* dialog.

This dialog will appear each time a new scatter point is created to allow the value to be verified or changed.

6. Click **OK** to exit the *Scatter Point Dataset Value* dialog.

The **Create Scatter Points** tool only enters steady state data, which could also be imported from a file. To create a scatter point set with transient data, it must be imported from a file.

3   Importing a Scatter Point Set

Scatter point sets can be imported from an existing file using the *Text Import Wizard* dialog. This dialog allows importing data into GMS from text files that are in columnar format. The file to be imported here was generated as an Excel spreadsheet and exported from Excel as tab-delimited text. The file contains scatter points which represent locations where the concentration of a contaminant has been estimated using a soil gas survey. The goal is to generate a map of the contaminant plume.

3.1   Open the File

1. Click **New** ⬚ to reset GMS, and select **No** when prompted to save changes.

2.   Click **Open** 📂 to bring up the *Open* dialog.

3.   Select "Text Files (*.txt,*.csv)" from the *Files of type* drop-down.

4.   Browse to the *geos2d* directory and select "plumedat.txt".

5.   Click **Open** to close the *Open* dialog and bring up the *Step 1 of 2* page of the *Text Import Wizard* dialog.

3.2   Import Wizard

The first step of the *Text Import Wizard* dialog specifies how the data is delineated and where in the file the data should begin. For this file, the first row contains column headings. The first column contains the point labels, the second column contains the X values of the points, the third column contains the Y values, and the fourth column contains the dataset values for the points.

1.   Below the *File import options* section, turn on *Heading row* option (red arrow, Figure 1) and click **Next** to go to the *Step 2 of 2* page of the *Text Import Wizard* dialog.
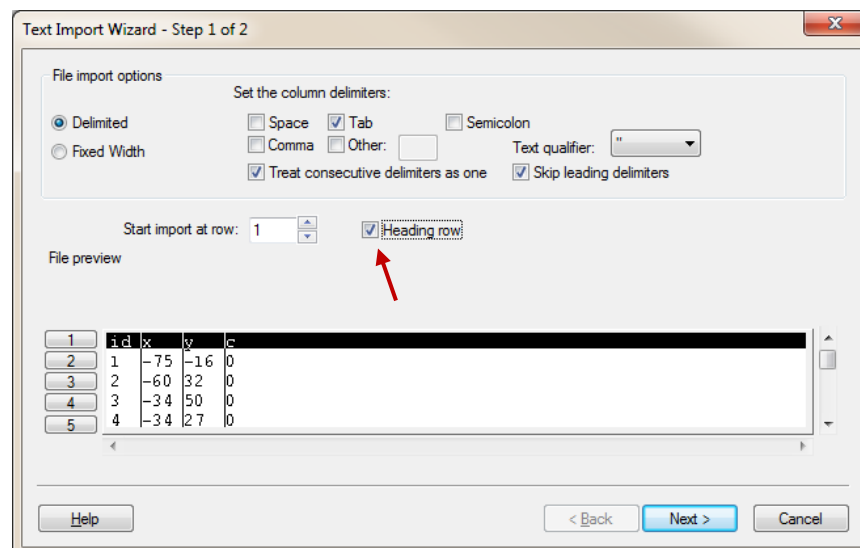

Figure 1   Text Import Wizard with Heading row selected

This page of the *Text Import Wizard* allows specifying the meaning of the data in the file.

2.   Select "2D Scatter Points" from the *GMS data type* drop-down.

The data associated with 2D scatter points is now displayed in the *File preview* section. The next step is to tell GMS what each column in the file means. This is done by selecting the correct data type using the drop-downs in the *Type* row (the first row) of the spreadsheet for each column in the file. The options in the drop-downs change depending on the selected *GMS data type*. Since the use of a heading row was specified in the first step, GMS looked at the headings and automatically found and mapped the X and Y columns by selecting the X and Y selection in the *Type* row of the spreadsheet.

3. Select "Label" from the *Type* row drop-down in the first column.

4. Select "Dataset" from the *Type* row drop-down in the fourth column.

The dataset in this file represents concentrations of a contaminant. The *No data* option near the top of the dialog can be used to specify a key value in the file used to indicate a lack of information.

For example, if no measurement was taken at one of the points, entering "-999" as the *No data* value in the *Text Import Wizard* would tell GMS to ignore these points when doing the interpolation later. The key value can be any value that would not normally be encountered in the dataset. In this case, this option isn't needed.

The dialog should now appear as in Figure 2.

5. Click **Finish** to close the *Text Import Wizard* dialog. A set of points should appear on the screen.
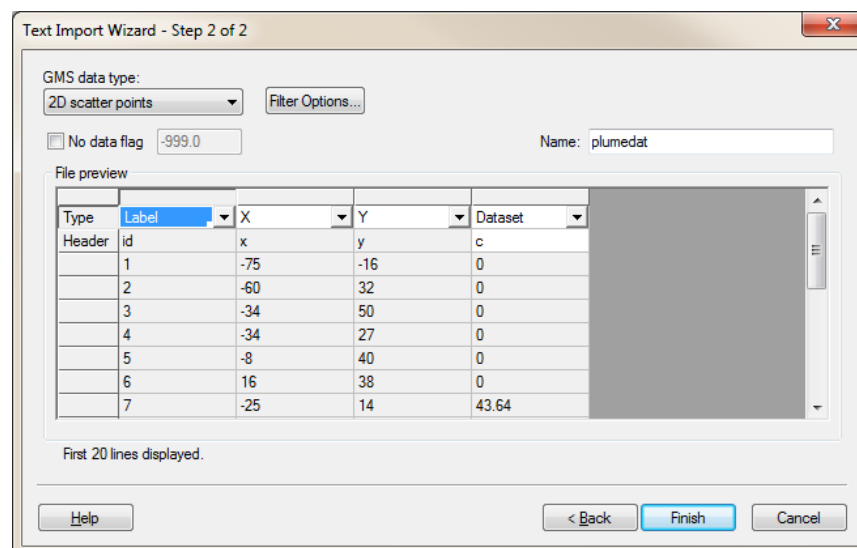

Figure 2   Text Import Wizard dialog, Step 2

3.3   Changing the Display Options

Change the appearance of the scatter points using the *Display Options* dialog:

1. Click **Display Options** 🖥 to bring up the *Display Options* dialog.

2. Select "2D Scatter Data" from the list on the left.

3. On the *2D Scatter Point Set* tab, turn on *Scatter point symbols* and *Contours*.

4. Click **OK** to exit the *Display Options* dialog.

Each of the points should now be displayed as a colored square (Figure 3). The color of the symbol represents the relative concentration of the contaminant at the point. When displaying colored symbols, it is useful to also display a color legend.
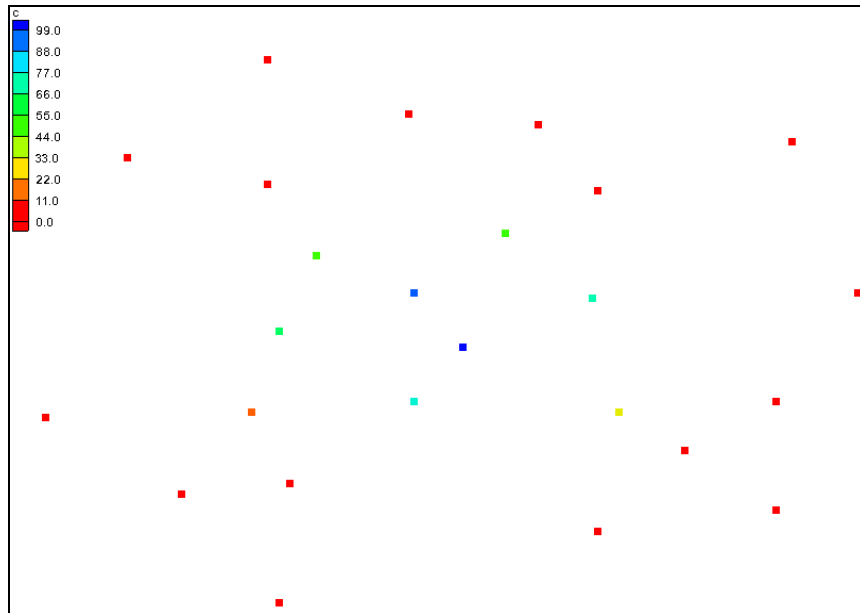
Figure 3   Scatter points with contour colors

5. Click **Display Options** 🖳 to bring up the *Display Options* dialog.

6. Select "▦ 2D Scatter Data" from the list on the left.

7. On the *2D Scatter Point Set* tab, click the **Options…** button to the right of *Contours* to bring up the *Dataset Contour Options – 2D Scatter Points - c* dialog.

8. At the bottom left, turn on *Legend* and click **OK** to close the *Dataset Contour Options – 2D Scatter Points - c* dialog.

9. Click **OK** to exit the *Display Options* dialog.

Notice that the concentrations vary from 0.0 to about 100.0 (see figure to right).
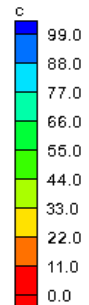
3.4   Creating a Bounding Grid

The goal of this tutorial is to generate a series of contour plots illustrating the plume. It is first necessary to create a grid that bounds the scatter point set. This allows interpolating the concentrations from the scatter points to the grid nodes. The grid is then contoured.

1. Right-click on "▦ plumedat" and select **Bounding 2D Grid…** to bring up the *Create Finite Difference Grid* dialog.

The *x* and *y* dimensions of this grid are already defined. The default values shown in the dialog cause the grid to extend beyond the scatter points by 10% on each side.

2. In the *X-Dimension* section, enter "60" for the *Number of cells*.

3. In the *Y-Dimension* section, enter "40" for the *Number of cells*.

4.   Click **OK** to close the *Create Finite Difference Grid* dialog.

5.   **Frame** 🔍 the project.

A grid should appear on the screen that encompasses only the scatter point set (Figure 4).
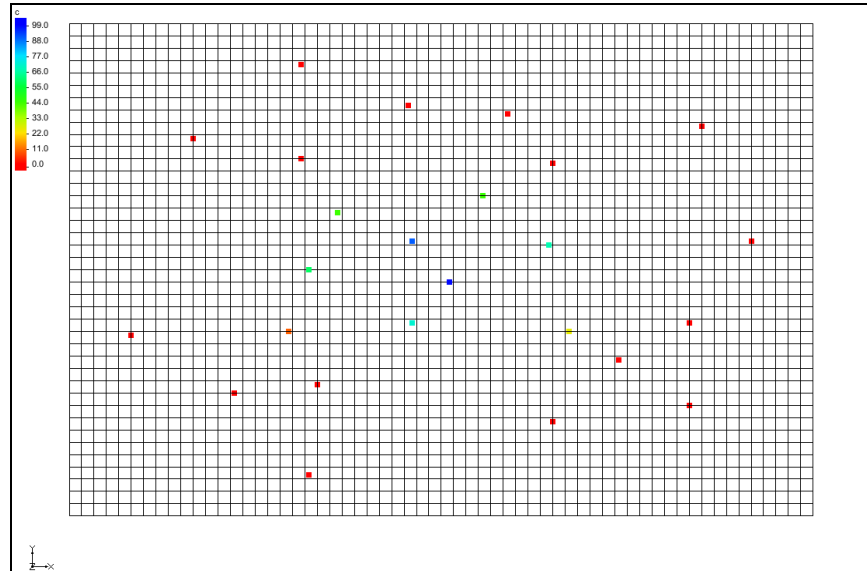

Figure 4   The grid encompassing the scatter points

4   Selecting an Interpolation Scheme

Several interpolation schemes are supported in GMS because each has different strengths
in different situations. Typically, the best approach is to try several schemes and then
determine which scheme is giving the most reasonable results. GMS has been structured in
such a way that multiple schemes can be tested quickly and easily.

4.1   Linear Interpolation

Do the following to try simple linear interpolation:

1.   Select *Interpolation* | **Interpolation Options…** to bring up the *2D Interpolation
     Options* dialog.

2.   In the *Interpolation method* section, select the *Linear* option.

3.   Click **OK** to close the *2D Interpolation Options* dialog.

To interpolate to the grid, do the following:

4.   Select *Interpolation* | **Interpolate → 2D Grid** to bring up the *Interpolate →
     Object* dialog.

5.   Click **OK** to accept the defaults and close the *Interpolate → Object* dialog.

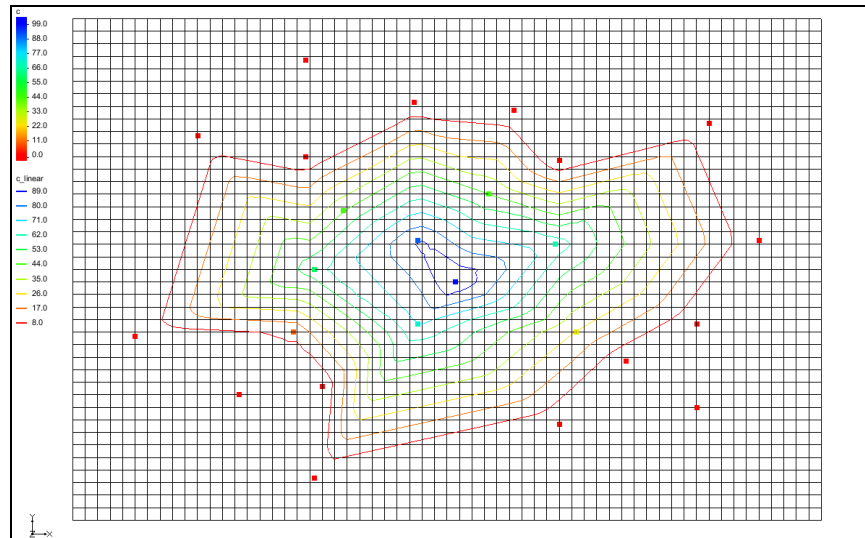A series of contour lines will appear within the grid (Figure 5).

Figure 5   Contour lines after linear interpolation

Viewing the Results

The contours show that the concentrations have been interpolated to the grid.

1. Switch to **Oblique View** ⬦.

Notice that the grid has been deformed to match the contours (Figure 6). Color shading the entire surface can provide an even more effective display of the grid.
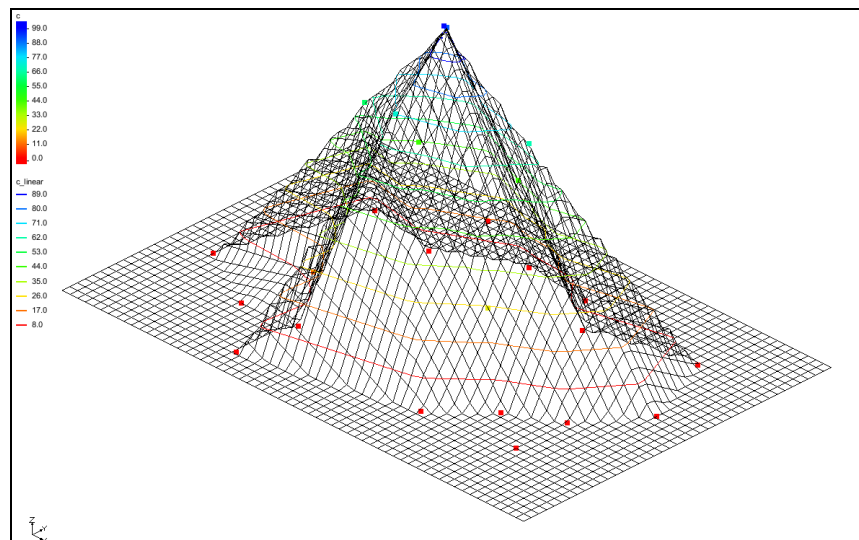

Figure 6   Oblique view

2. Click **Display Options** 🖵 to bring up the *Display Options* dialog.

3. Select "⊞ 2D Grid Data" from the list on the left.

4. On the *2D Grid* tab, in the lower section, turn on *Contours* and click **Options...** to bring up the *Dataset Contour Options – 2D Grid – c_linear* dialog.

5. In the *Contour method* section, select "Color Fill" from the drop-down.

6. Click the **Color Ramp…** button to bring up the *Color Options* dialog.

7. In the *Palette preview* section, click the **Reverse** button.

8. Click **OK** to close the *Color Options* dialog.

9. Click **OK** to close the *Dataset Contour Options – 2D Grid – c_linear* dialog.

10. Click **OK** to exit the *Display Options* dialog.
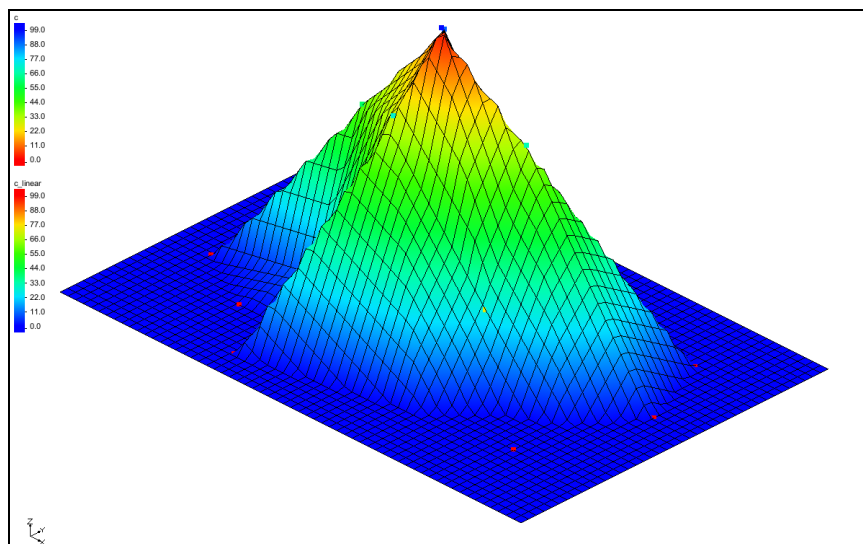
The 2D grid should appear similar to Figure 7.

Figure 7   Linear interpolation with color fill

The outer part of the grid still has a concentration value equal to zero, as indicated by the flat, dark blue contour coloring. When linear interpolation is performed, the scatter points are triangulated to form a temporary TIN. A plane equation is computed for each triangle in the TIN and the coefficients of the plane equation are used to interpolate to points inside the triangle.

Therefore, linear interpolation cannot be performed for grid nodes outside the convex hull of the TIN (the boundary of the TIN). As a result, these nodes are assigned a value of zero. A value of zero is appropriate in this case since the concentrations of the scatter points on the perimeter of the scatter point set are zero.

Viewing the Elliptical Control Function

The scatter points were generated from an elliptical control function. This shows how the different interpolation schemes compare to the original control function. To view the original control function:

1. Right-click on "▦ grid" in the Project Explorer and select **Import Dataset…** to bring up the *Dataset Filename* dialog.

2. Select "All Files (*.*)" from the *Files of Type* drop-down.

3. Browse to the *geos2d* directory and select "ellipse2g.dat".

4. Click **Open** to import the dataset and close the *Dataset Filename* dialog.

5. Select the "▦ control_function" dataset in the Project Explorer to make it active.

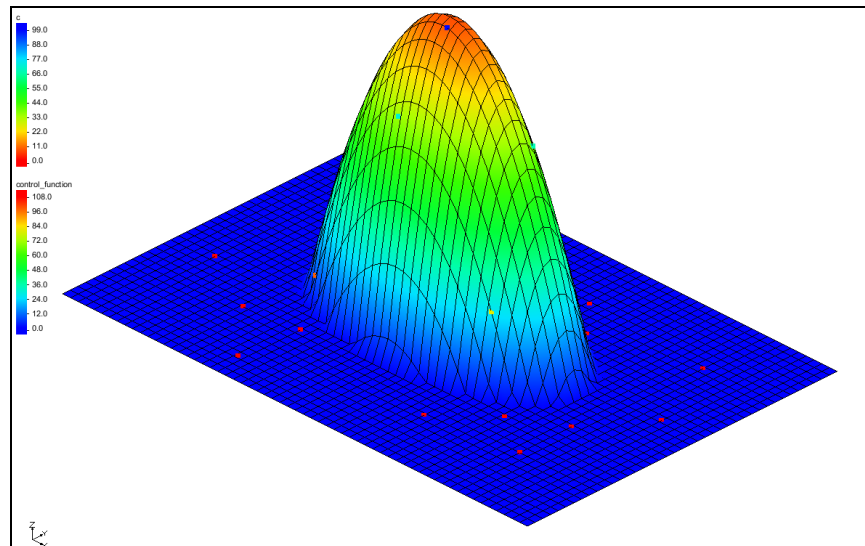The Graphics Window should appear similar to Figure 8.

Figure 8   Interpolation scheme with original control function

4.2   Simple IDW Interpolation

The next scheme to try is a simple form of inverse distance weighted (IDW) interpolation.

1. Select "▦ 2D Scatter Data" in the Project Explorer to make it active.

2. Select *Interpolation* | **Interpolation Options** to bring up the *2D Interpolation Options* dialog.

3. In the *Interpolation method* section, select *Inverse distance weighted*.

4. Click **Options…** to bring up the *2D IDW Interpolation Options* dialog.

5. In the *Nodal function* section, select *Constant (Shepard's method)*.

6. Click **OK** to exit the *2D IDW Interpolation Options* dialog.

7. Click **OK** to exit the *2D Interpolation Options* dialog.

8. Select *Interpolation* | **Interpolate → 2D Grid** to bring up the *Interpolate → Object* dialog.

9. Click **OK** to accept the defaults and close the *Interpolate → Object* dialog.

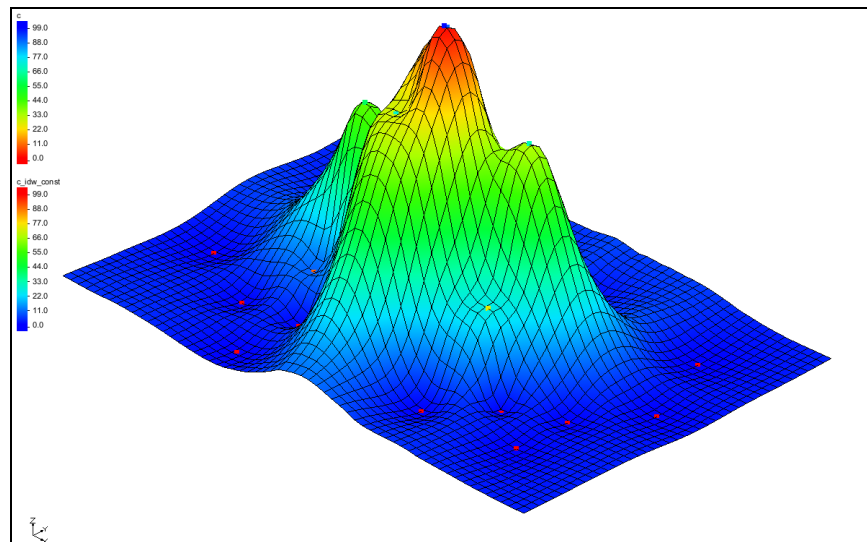The Graphics Window should appear similar to Figure 9.

Figure 9   Inverse distance weighted (IDW) interpolation

The IDW scheme is a simple moving weighted averages scheme. To interpolate a value at a point, a weighted average of the nearby scatter points is used. The weights are an inverse function of distance. The closer a scatter point is to the interpolation point, the greater the weight given to the scatter point.

IDW Interpolation with Gradient Planes

One of the problems with simple IDW interpolation is that the interpolated dataset always tends toward the mean of the dataset in the voids between scatter points. As a result, local minima or maxima in the voids in the scatter point set are not properly inferred. To overcome this problem, a nodal function can be computed at each scatter point.

A nodal function is a plane or quadratic function that is forced to pass through the scatter point and approximate the nearby scatter points in a least squares sense. When the interpolation is performed, rather than computing an average of the dataset values at the scatter point locations, an average is computed of the nodal functions of the nearby scatter point evaluated at the interpolation point. This approach allows local trends to be inferred and often results in a more accurate interpolation.

This is demonstrated through the following steps:

1. Select " 2D Scatter Data" in the Project Explorer to make it active.

2. Select *Interpolation* | **Interpolation Options…** to bring up the *2D Interpolation Options* dialog.

3. In the *Interpolation method* section, click **Options…** to bring up the *2D IDW Interpolation Options* dialog.

4. In the *Nodal function* section, select the *Gradient plane* option.

5. Click **OK** to exit the *2D IDW Interpolation Options* dialog.

6. Click **OK** to exit the *2D Interpolation Options* dialog.

7. Select *Interpolation* | **Interpolate → 2D Grid** to bring up the *Interpolate → Object* dialog.

8. Click **OK** to accept the defaults and close the *Interpolate → Object* dialog.

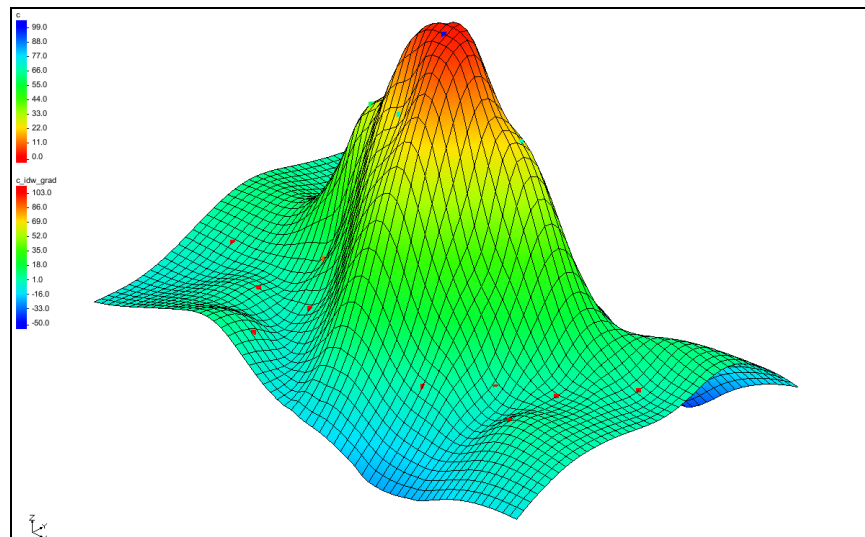The Graphics Window should appear similar to Figure 10.

Figure 10   IDW interpolation with gradient planar nodal functions

Using the Horizontal Anisotropy Option

The interpolation can be stretched in a horizontal direction by using the horizontal anisotropy option in conjunction with a specified azimuth. One reason for using this option might be if it is known that a contaminant plume is elongated in a particular direction due to groundwater flow.

Do the following to change the horizontal anisotropy:

1. Switch to **Plan View** .

2. Select " 2D Scatter Data" in the Project Explorer to make it active.

3. Select *Interpolation* | **Interpolation Options…** to bring up the *2D Interpolation Options* dialog.

4. In the *Anisotropy* section, enter "2.0" for the *Horizontal anisotropy*.

5. Enter "90.0" for the *Azimuth*.

6. Click **OK** to close the *2D Interpolation Options* dialog.

7. Select *Interpolation* | **Interpolate → 2D Grid** to bring up the *Interpolate →
   Object* dialog.

8. Enter "c_idw_grad_h2" for the *Interpolated dataset name.*

9. Click **OK** to close the *Interpolate → Object* dialog.

As can be seen in Figure 11, the data has been stretched horizontally. If desired,
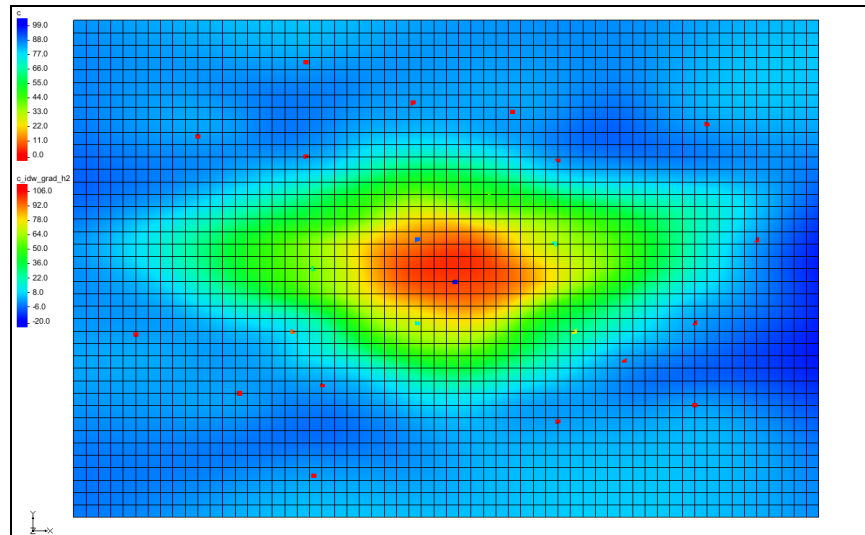experiment with different values for the horizontal anisotropy.

Figure 11   Horizontal anisotropy

IDW Interpolation with Quadratic Nodal Functions

The nodal functions used in IDW interpolation can also be quadratic functions which are
constrained to pass through the scatter point and approximate the neighboring scatter
points in a least-squares fashion. The averaging or blending of the quadratic functions
during the interpolation stage often results in a very smooth surface.

1. Switch to **Oblique View** .

2. Select *Interpolation* | **Interpolation Options…** to bring up the *2D Interpolation
   Options* dialog.

3. In the *Interpolation method* section, click the **Options…** button to bring up the
   2D IDW Interpolation Options dialog.

4. In the *Nodal function* section, select the *Quadratic* option.

5. Click **OK** to exit the *2D IDW Interpolation Options* dialog.

6. Click **OK** to exit the *2D Interpolation Options* dialog.

7. Select *Interpolation* | **Interpolate → 2D Grid** to bring up the *Interpolate →
   Object* dialog.

8.   Click **OK** to accept the defaults and close the *Interpolate → Object* dialog.

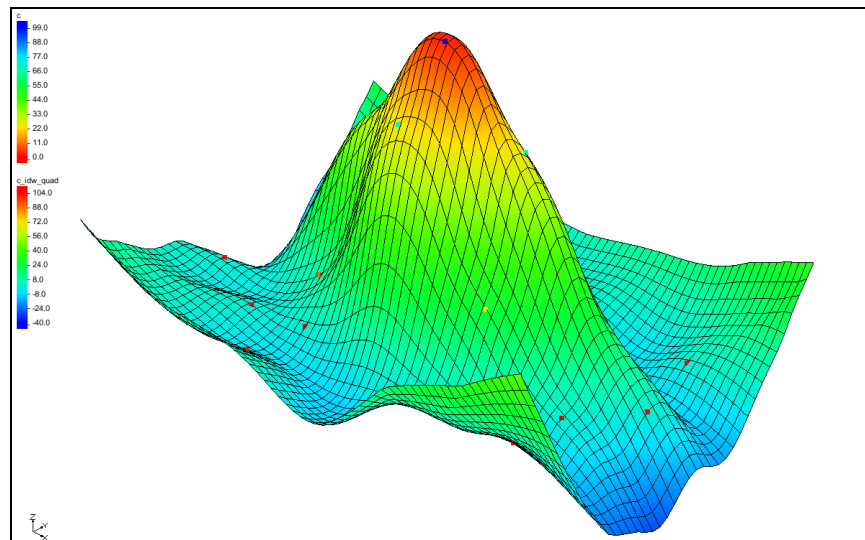The Graphics Window should appear similar to Figure 12.

Figure 12   Quadratic nodal functions

Truncation

Notice that the minimum value listed in the color legend is a negative number. This is impossible since there is no such thing as a negative concentration. By inferring trends, the nodal functions can sometimes project the plume values beyond zero and into the negative range. This type of error can be easily fixed using truncation.

1.   Select *Interpolation | **Interpolation Options…*** to bring up the *2D Interpolation Options* dialog.

2.   In the lower right section, turn on *Truncate values*.

3.   Select *Truncate to specified range,* then enter "0.0" for *Min* and "150.0" for *Max*.

4.   Click **OK** to close the *2D Interpolation Options* dialog.

The concentrations should not go below zero, but it is okay to allow the interpolation scheme to infer a maximum concentration greater than the maximum measured value.

5.   Select *Interpolation | **Interpolate → 2D Grid*** to bring up the *Interpolate → Object* dialog.

6.   Enter "c_idw_quad_trunc" for the *Interpolated dataset name*.

7.   Click **OK** to close the *Interpolate → Object* dialog.

Notice that many of the concentrations around the perimeter of the map are now zero (Figure 13).
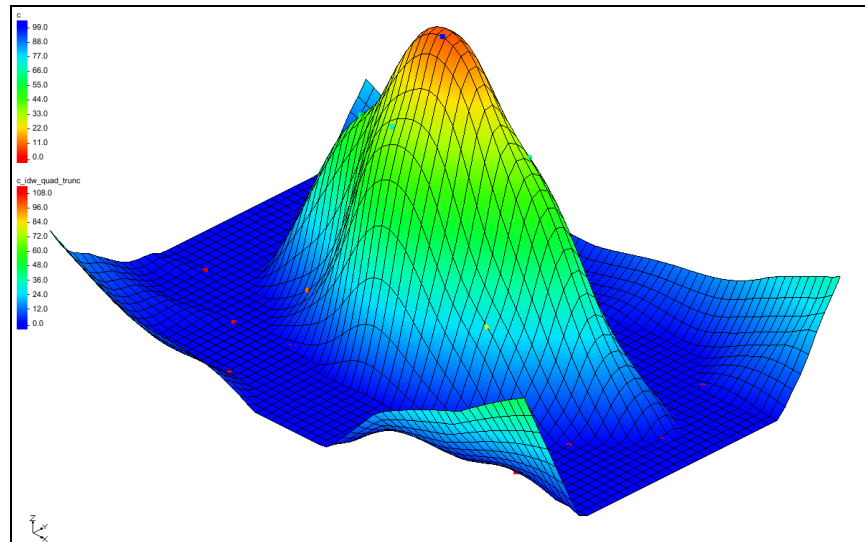
Figure 13   Truncated quadratic nodal functions

4.3   Kriging

The last interpolation scheme to test is kriging. Kriging is based on the assumption that points near each other have a certain degree of spatial correlation, but points that are widely separated are statistically independent. Kriging is a set of linear regression routines that minimizes estimation variance from a predefined covariance model.

1. Select *Interpolation |* **Interpolation Options…** to bring up the *2D Interpolation Options* dialog.

2. In the *Interpolation method* section, select *Kriging*.

3. Click the **Options…** button to bring up the *Kriging Options* dialog.

Creating the Experimental Variogram

There are a large number of options to be specified in the *Kriging Options* dialog. While a variogram must always be defined, the defaults shown are adequate in most cases.

1. Click **Edit Variograms…** to bring up the *Variogram Editor* dialog.

2. In the *Experimental variogram* section, click the **New** button to bring up the *Experimental Variogram* dialog .

3. Click **OK** to accept the defaults and close the *Experimental Variogram* dialog.

A curve should appear in the upper window of the *Variogram Editor* (Figure 14). This curve is called an experimental variogram. The experimental variogram is found by calculating the variance in dataset values of each scatter point in the set with respect to each of the other points and plotting the variances versus the distance between the points.
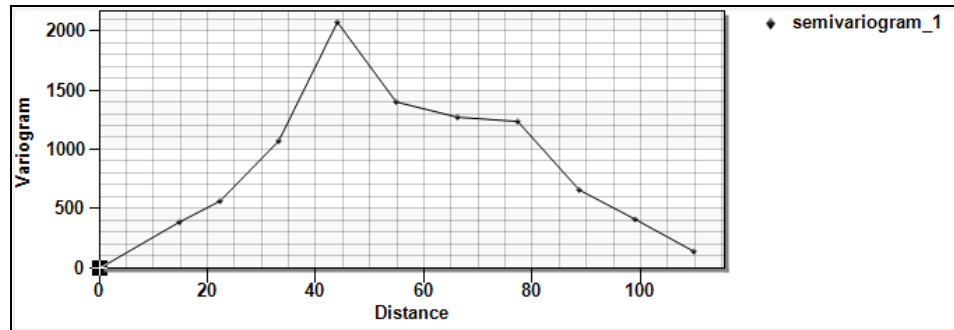
Figure 14   The experimental variogram curve in the Variogram Editor dialog

As can be seen in the plot of the experimental variogram, the shape of the variogram indicates that the variance is smaller at short separation distances. In other words, points that are close together have similar data values. The variance in the data values with many datasets becomes somewhat random after a certain level of separation. The variogram also oscillates around a value corresponding to the average variance. However, many of the points have zero values with concentration data and this tends to pull the experimental variogram back down.

Creating the Model Variogram

Once the experimental variogram is computed, the next step is to define a model variogram. A model variogram is a simple mathematical function that models the trend in the experimental variogram. The model variogram is used in the kriging computations.

1. In the *Nested structure* section of the *Variogram Editor* dialog, click the **New** button.

A red model variogram will appear in the variogram.

2. Select "Gaussian" from the *Model function* drop-down.

The red variogram will change slightly.

3. Enter "25.0" in the field to the right of the *Nugget* slider.

4. Enter "1965.0" in the field to the right of the *Contribution* slider.

5. Enter "63.0" in the field to the right of the *Range* slider.

At this point, a reasonable fit should exist between the model variogram and the left half of the initial variogram. The right half is difficult to fit because of the zero values used in this case.

6. Click **OK** to exit the *Variogram Editor* dialog.

7. Click **OK** to exit the *Kriging Options* dialog.

8. Click **OK** to exit the *2D Interpolation Options* dialog.

9. Select *Interpolation* / **Interpolate → 2D Grid** to bring up the *Interpolate → Object* dialog.

10. Click **OK** to accept the defaults and close the *Interpolate → Object* dialog.

As seen in Figure 15, notice that this interpolation scheme results in an upward curvature towards the outside edges of the grid. This can be corrected by adding more scatter points with a concentration of "0.0" in these outlying areas.
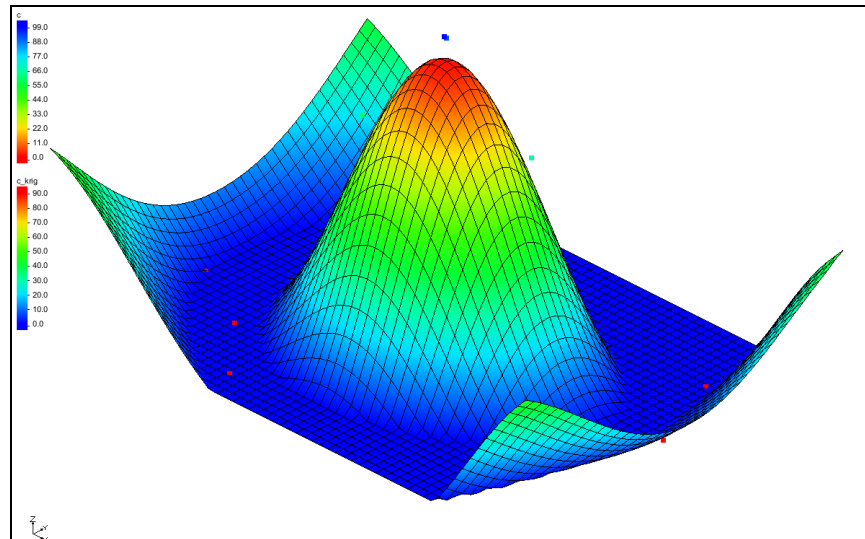

Figure 15   Model variogram based on kriging interpolation

5   Switching Datasets

Now that the grid has been interpolated using several different interpolation schemes, review the results by replotting some of the interpolated datasets. To do this, switch back to one of the previous datasets using the Project Explorer.

1. Switch to **Plan View**.

2. Select the "c_linear" dataset under " grid" in the " 2D Grid Data" folder in the Project Explorer.

3. Use the *Up* and *Down* keys to switch between the datasets. Notice the differences in the contours for each dataset.

6   Using the Data Calculator

The **Data Calculator** tool can be useful when comparing two datasets generated by interpolation. Use the *Data Calculator* to compute the difference between the kriging and IDW quadratic datasets by doing the following:

1. Click **Data Calculator** ⊞ to bring up the *Data Calculator* dialog.

The available datasets are listed in the *Datasets* section, where they are numbered from "d1" to "d10". Datasets are referenced in the mathematical expression using the "d" numbers. The "c_krig" dataset should be numbered "d10" and the "c_idw_quad" dataset should be labeled "d7".

2. In the *Expression* field, enter "abs(d7-d10)".

3. In the *Result* field, enter "Difference".

4. Click **Compute** to perform the calculation and create a new "Difference" dataset in the Project Explorer. It will also appear in the *Datasets* section as entry "d11".

This computes the difference between the two datasets. Any new dataset computed using the *Data Calculator* is automatically designated the active dataset. With that calculation completed, it is helpful to view some basic statistics related to the new dataset.

5. Click **Dataset Info…** to bring up the *Dataset Info* dialog.

This dialog displays basic statistics related to the selected dataset such as minimum, maximum, and mean data values. Feel free to review the statistics.

6. Click **OK** to exit the *Dataset Info* dialog.

7. Click **Done** to exit the *Data Calculator* dialog.

8. Switch to **Oblique View** ◈.

The contour plot now displayed represents the "⊞ Difference" dataset that was just computed (Figure 16).
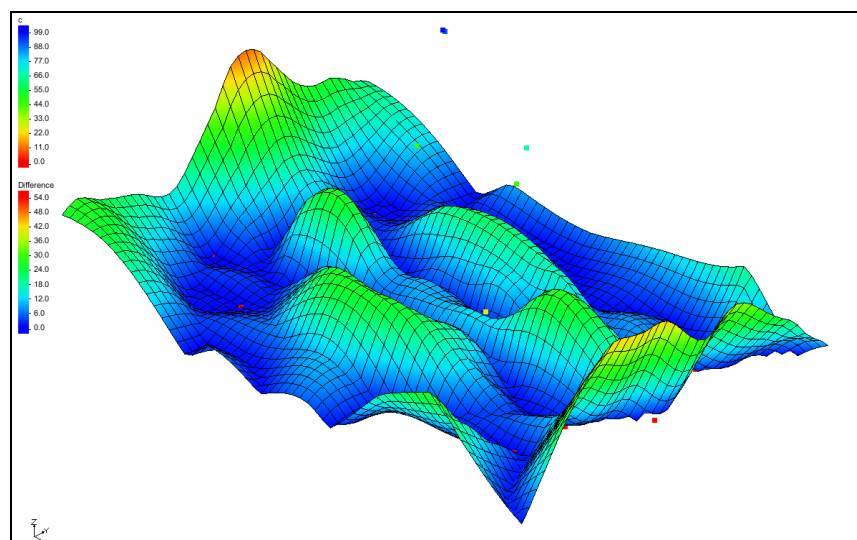

Figure 16   Difference contour in oblique view

7   Conclusion

This concludes the "Geostatistics – 2D" tutorial. Key concepts demonstrated and discussed in this tutorial include the following:

- Interpolating from a sparse set of points to a different set of points (usually more dense and surface-defining points) by using 2D scatter points.

- 2D scatter points can be created manually or by importing them from a file.

- Several interpolation algorithms are available in GMS.

- Linear algorithms do not interpolate beyond the convex hull of the scatter points.

- Anisotropy and truncation can be used to help control the interpolation process.