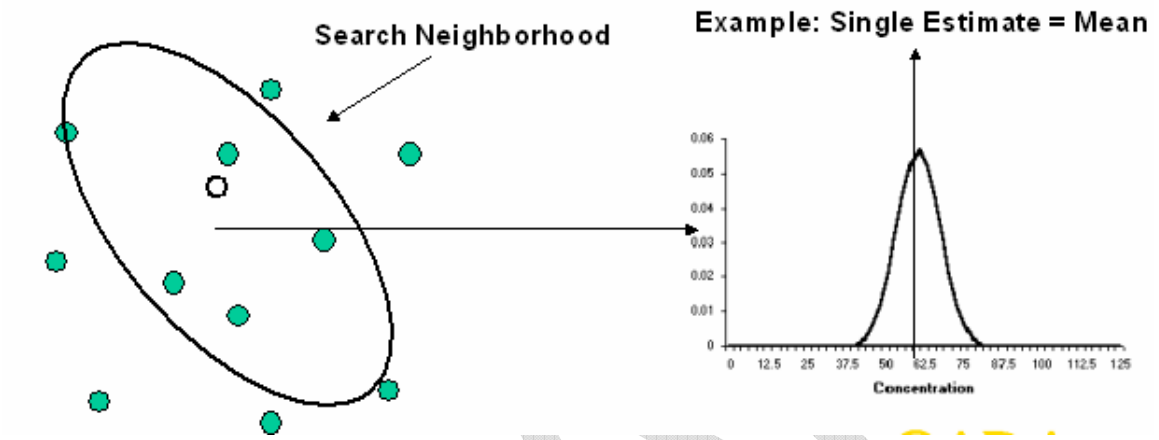
Chapter 30: Advanced Geospatial Methods Part I: Overview and Correlation Modeling

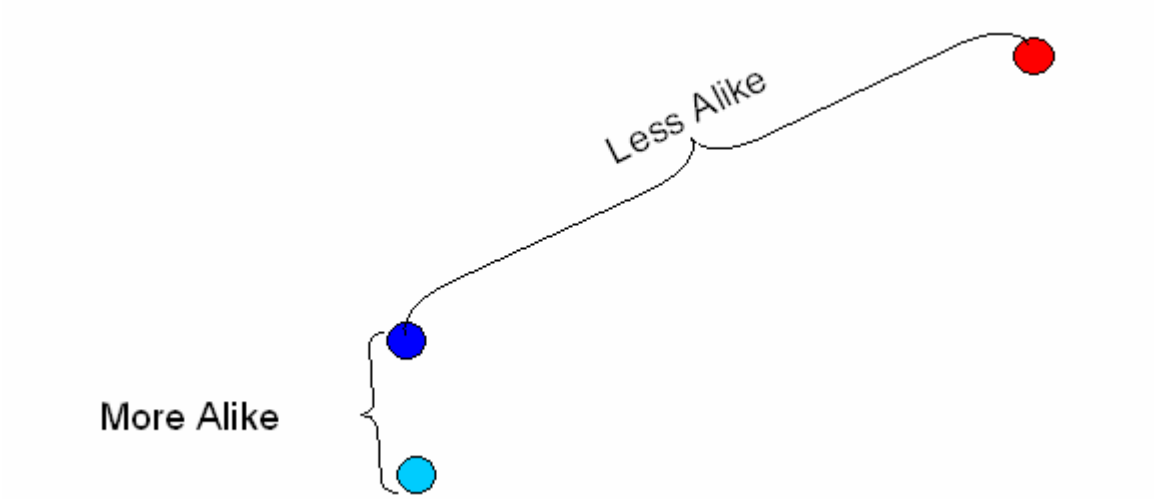
With basic spatial analysis tools, each interpolant produced a single estimate for each unsampled point. With a geostatistical approach, a distribution of possible values is constructed and used as a model for the actual unsampled value. Both an estimate and a model of uncertainty can be now obtained. From this distribution of points, a central moment, such as the mean or media, is chosen as a single estimate for contouring purposes.



SADA provides three kriging (geostatistical) models: Ordinary, Indicator kriging, and cokriging. Ordinary kriging assumes a normal or lognormal distribution for the data. Indicator kriging is a non parametric approach that does not assume any distribution. Cokriging allows you to include other types of data in the analysis that may not be direct measurements of your contaminant of interest. This permits the inclusion of various kinds of field detection measurements, geophysics, and so forth.

Like the methods discussed in Basic Geospatial Methods, these methods are based on a weighted combination of nearby samples. However, the development and expression of these weights is more complex. It may be helpful at first to think of kriging as an advanced form of the inverse distance method. Recall that the inverse distance method weights sampled values by their distance from the unsampled location. Kriging approaches the problem in much the same way. However, rather than distance (d), the weights are based on the amount of spatial correlation or spatial covariance that samples exhibit at varying distances.

If data are spatially correlated, then on average, sample points that are close to each other are more alike than sample points further away. (More complex spatial correlations exist but this type is the most common).

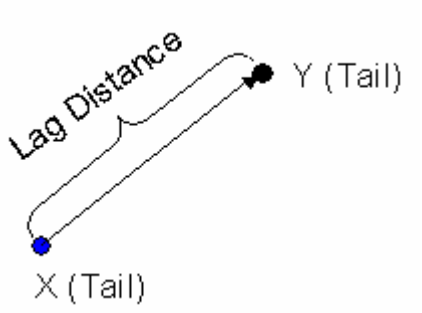


Concepts in Spatial Correlation

The degree to which data are more or less "alike" for any given distance can be calculated. SADA uses the semi-variogram method, which returns a measure of variance for any given distance of separation. This measure is defined as half of the average squared difference between values separated by distance h. The term h is referred to as the lag or lag distance.

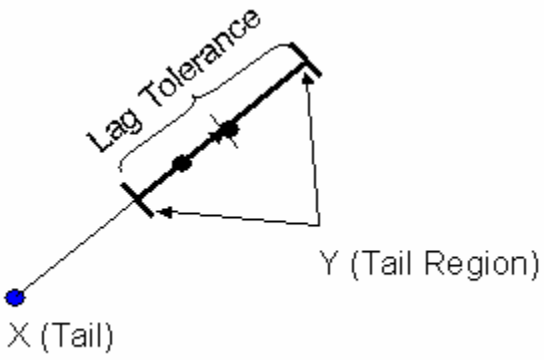
$$\gamma(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} (x_i - y_i)^2$$

where N(h) is the number of pairs separated by distance h, x_i is the starting sample point (tail), and y_i is the ending sample point (head).

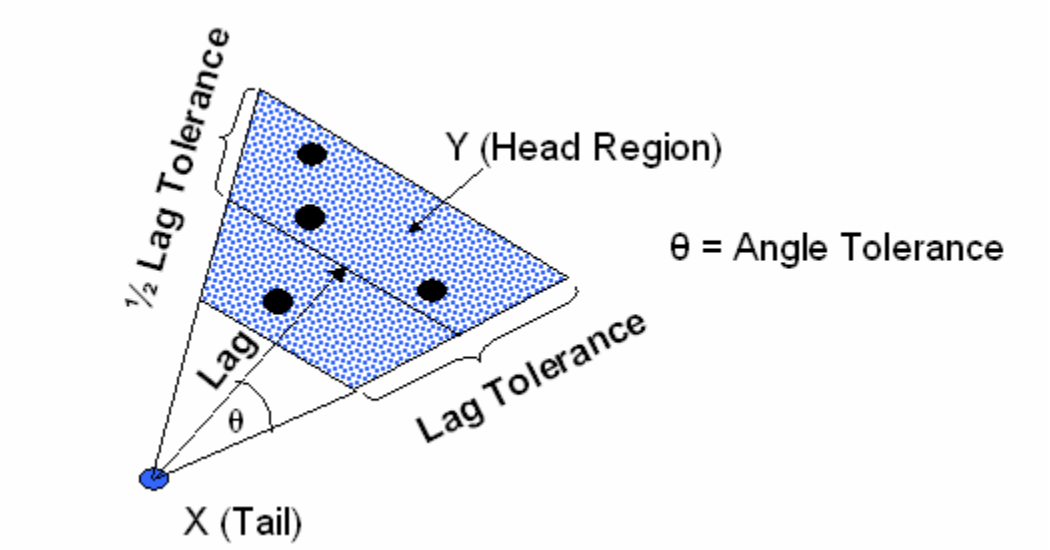


Rarely in practice, will you ever have any sample points separated by exactly a lag distance h. Therefore, a lag tolerance centered about the lag distance will permit a capture of more data points in the calculation of γ(h). In the figure below, all data points on the dark thick line area will be used.

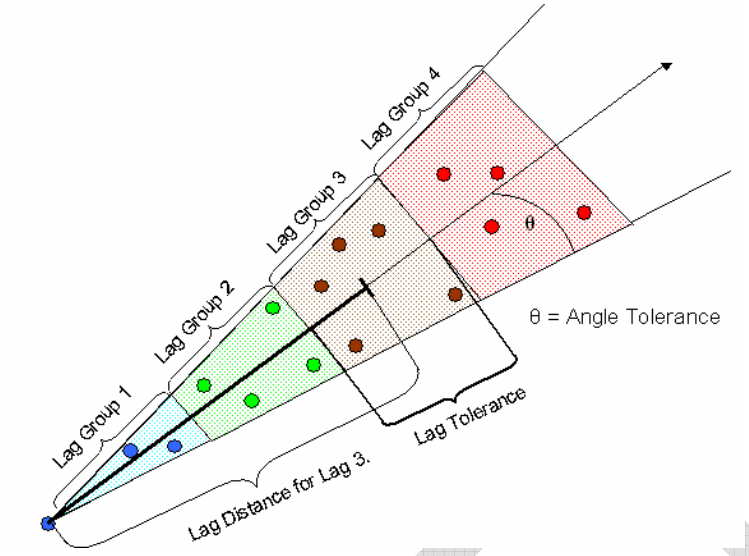
So if we are interested in the variance of all data points separated by 10 feet and we permit a lag tolerance of 2 feet. We will actually be calculating the variance of all pairs of data between 9 and 11 feet apart.



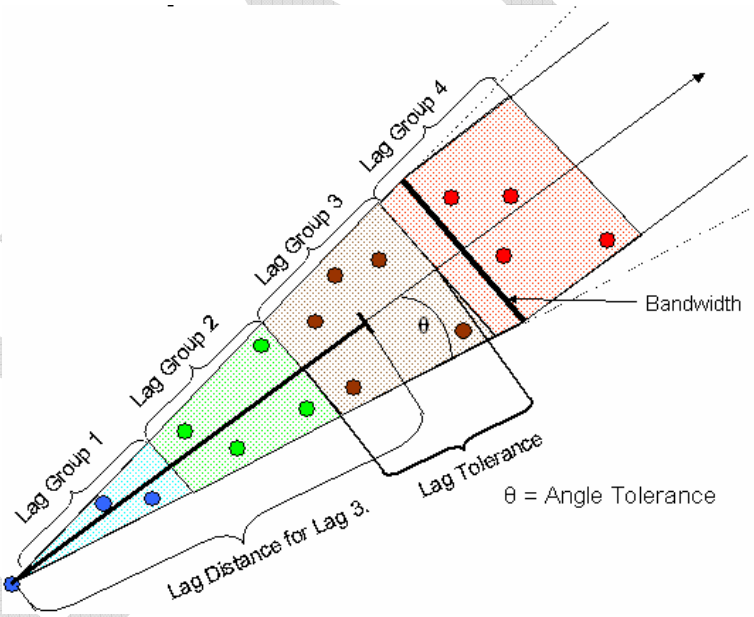
Although assigning a lag tolerance helps, most cases will never have enough samples separated by a lag - tol/2 to lag + tol/2 along a straight line to calculate the semivariogram value. Therefore, an angle tolerance, θ, is also introduced to expand the region and to include more points in the calculation of the semivariogram value for the specified lag distance. In the figure below, all data points within the blue shaded area will be used.



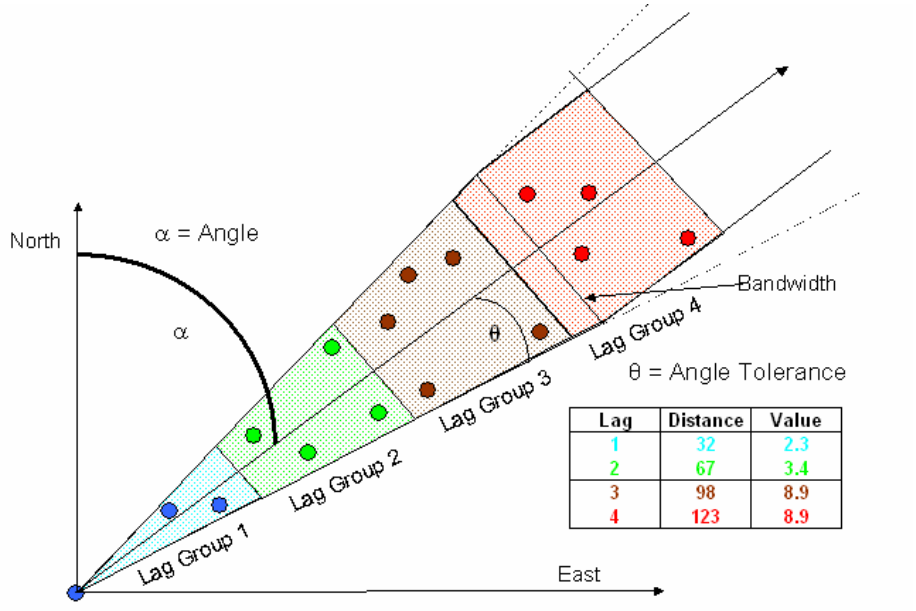
If we repeat this operation for a number of lag distances, we would generate a cone shaped object expanding outward from the point of interest. This cone would be partitioned by lag groups centered about our lag distances.



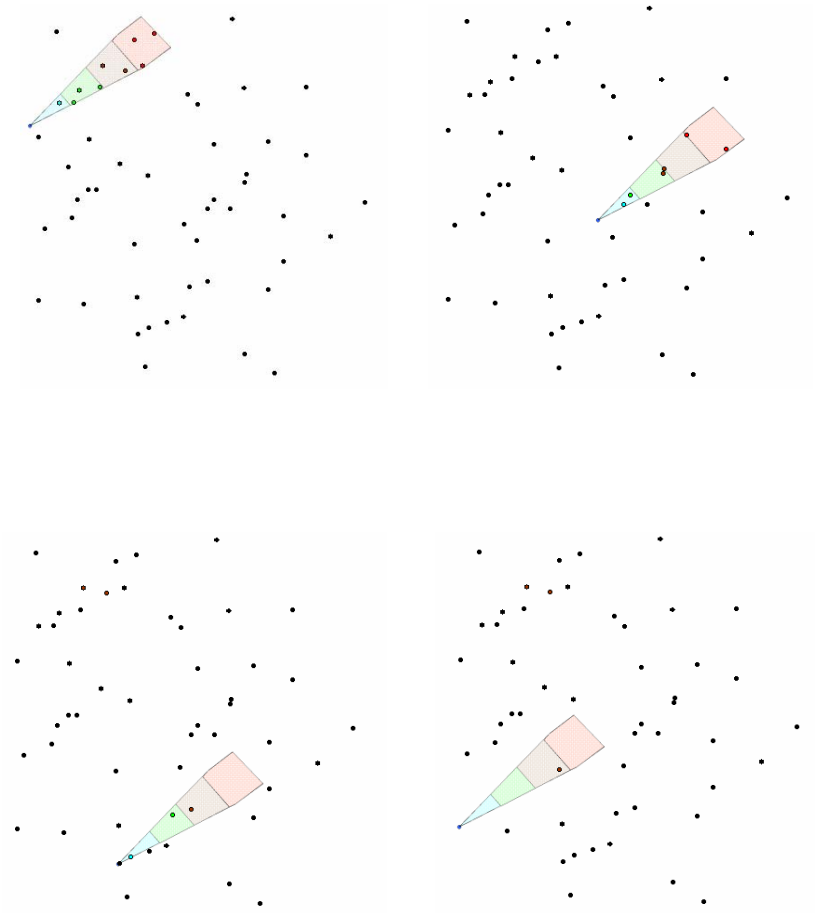
As the cone stretches farther out, it opens up increasingly wide, capturing more and more data points as it moves away. In practice, geostatisticians will often apply a constraint called the bandwidth. This bandwidth limits the expansion of the cone to a certain width. If you do not wish to constrain the cone's expansion, specify a very large bandwidth.



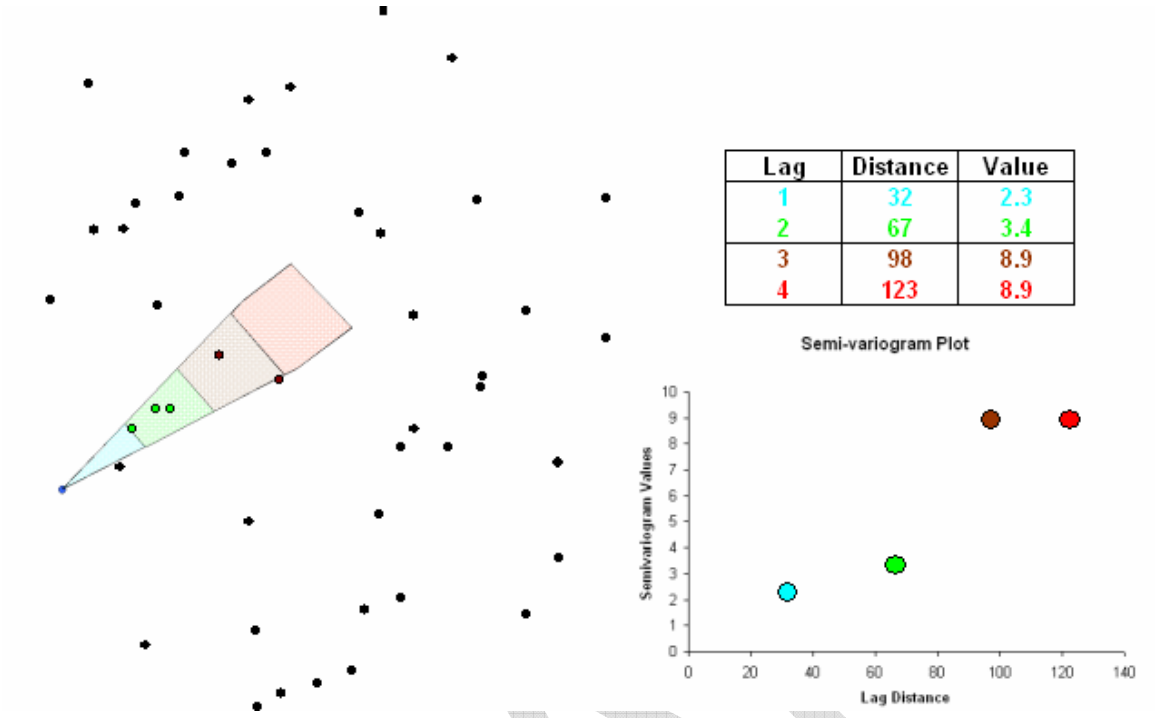
Our final parameter is the angle. The angle specifies in what direction you will be calculating the semi-variogram values. This is sometimes referred to as the angle of anisotropy. Now we are constraining our semi-variogram values to a certain direction.



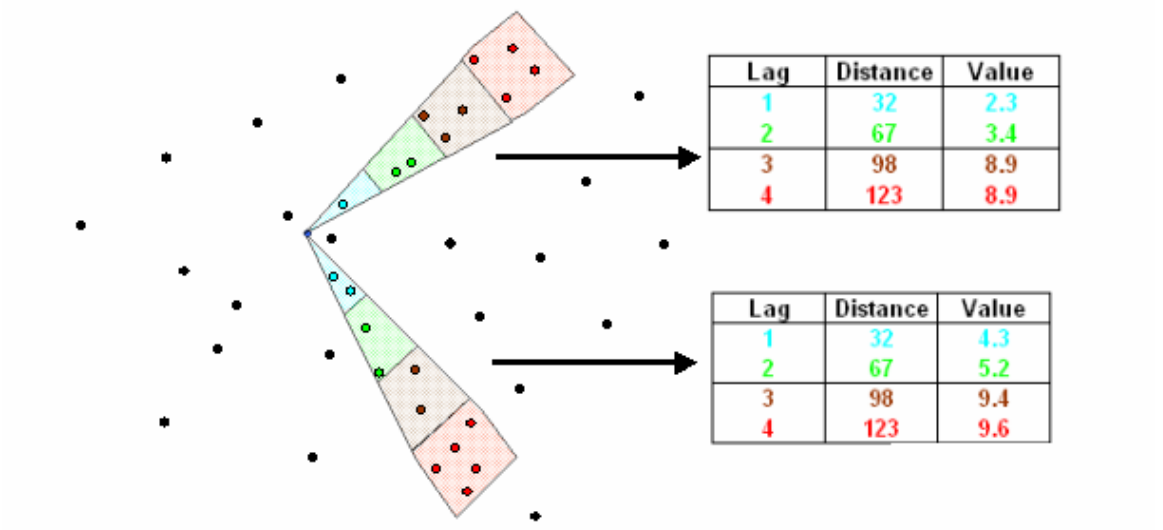
The semi-variogram cone is relocated at every sample location to build up the total set of pairs for each lag bin.



The semi-variogram values are then plotted.

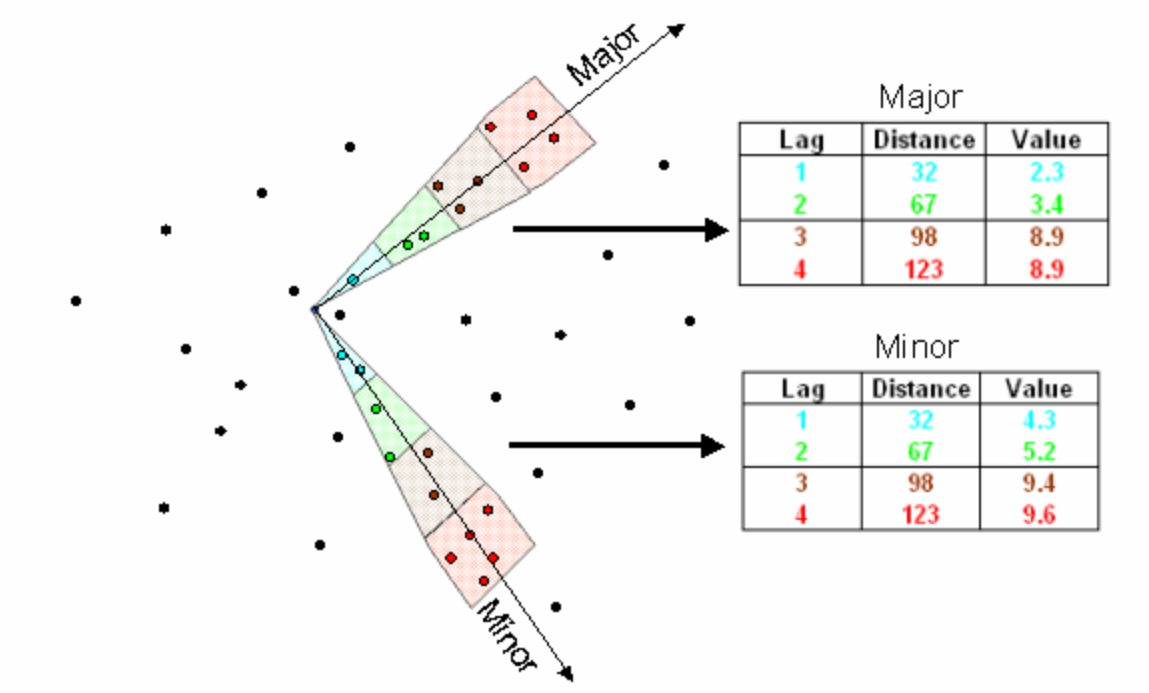


Note though, by specifying an angle α, we are excluding all those data points located outside of the cone from α - θ degrees to α + θ degrees. In other words, we are exploring how data are correlated in a particular direction. If we find that data are more correlated in one direction than another, the data are said to be anisotropic. This means that data in the direction α are more alike than in other directions.



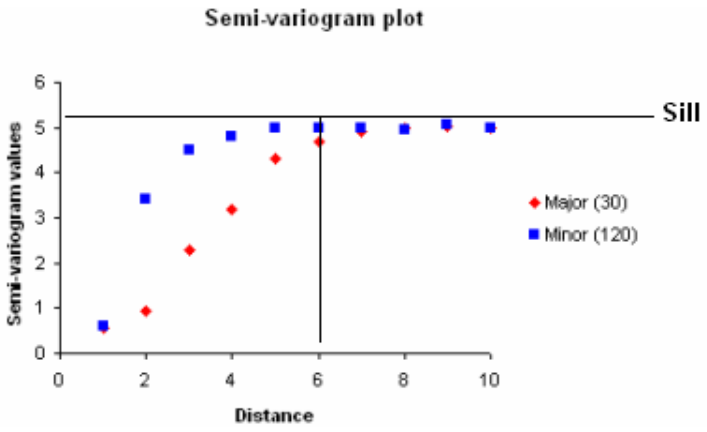
In fact, if anisotropic conditions exist, the direction of highest correlation is considered the major direction of anisotropy. The perpendicular direction is referred to as the minor direction

of anisotropy. The major direction of correlation will exhibit semi-variogram values that increase at a slower rate than any other direction.



Theoretically, the semi-variogram values will continue to rise until they reach the sill value. The sill is the point at which the data are now far enough apart to be independent. The sill value should be roughly equivalent to the variance of the data set. A semi-variogram plot is useful in detecting the sill value and location.

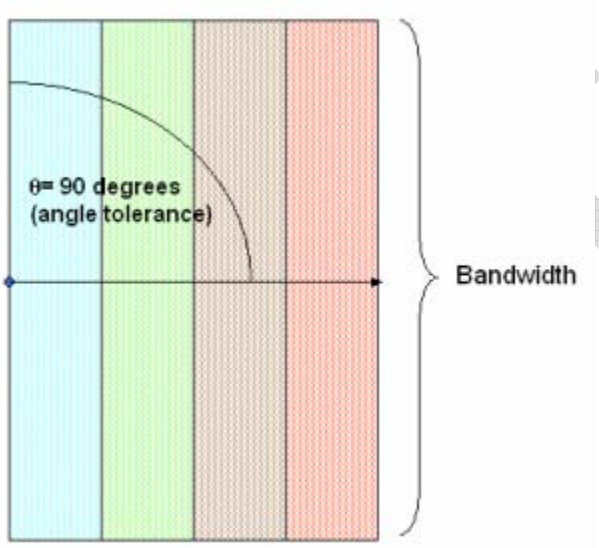
Lag	Major (30)	Minor (120)
1	0.56	0.6
2	0.95	3.4
3	2.3	4.5
4	3.2	4.8
5	4.3	5
6	4.7	5
7	4.9	5
8	5	4.95
9	5.01	5.05
10	4.99	5



In the above example, we see a major direction at 30 degrees and the corresponding minor direction at 120 degrees. A sill value of approximately 5 is detected around 6 feet of separation.

Isotropic Variograms

In order to calculate an isotropic or omni-directional variogram, simply set the angle tolerance to 90 degrees and make the bandwidth significantly larger than the site. This will force the cone to consider the entire spectrum of data points.



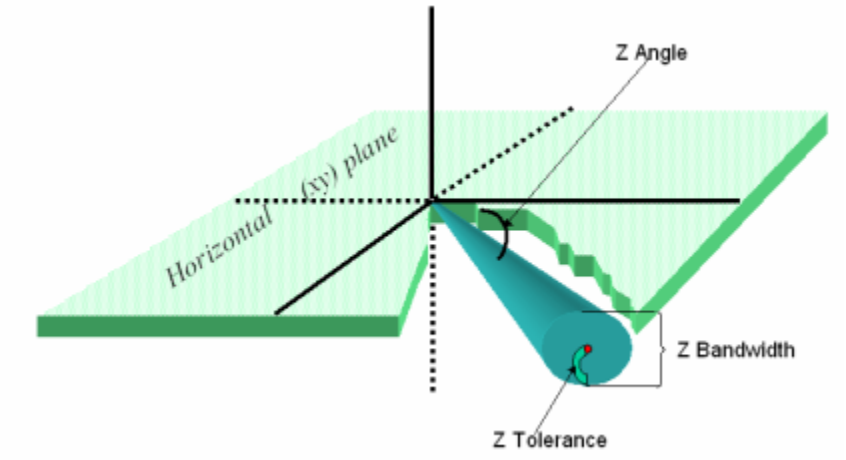
Three Dimensional Variography

Three-dimensional semi-variogram calculation is the same approach as in the two-dimensional case. In addition to the previously defined parameters, a z angle (dip), z tolerance, and z bandwidth must be specified.

Z Angle (Dip) – The angle below the horizontal plane that the cone should dip.

Z Tolerance – The tolerance on this dip angle.

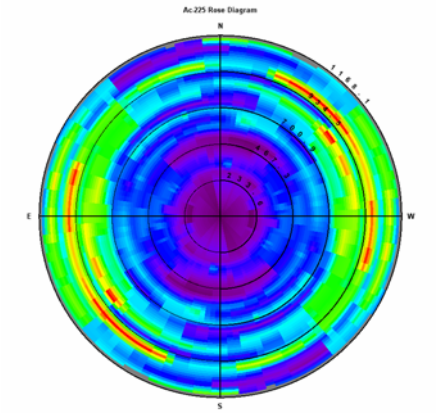
Z Bandwidth – The maximum distance the vertical component of the cone is permitted to go.



Variogram Maps (Rose Diagrams)

Rather than viewing only on angle at a time, users can view semivariogram values in all directions at once. Here the semi-variogram is calculated with an increasing, incremental horizontal angle. Instead of plotting the results on an xy plot, they are plotted on a disk where the color correponds to the value of the semi-variogram value. The center of the disk is the zero distance point. For three dimensional data, you can specify a range of z dip angles and SADA will compute this disk for incremental changes in the dip as well.

The benefit of rose map is that Directions with low “trough like” values are in present they represent the major direction of anisotropy. In the following image, there is a trough of low values running N-S. The use of colors makes interpretation of variogram result much easier.

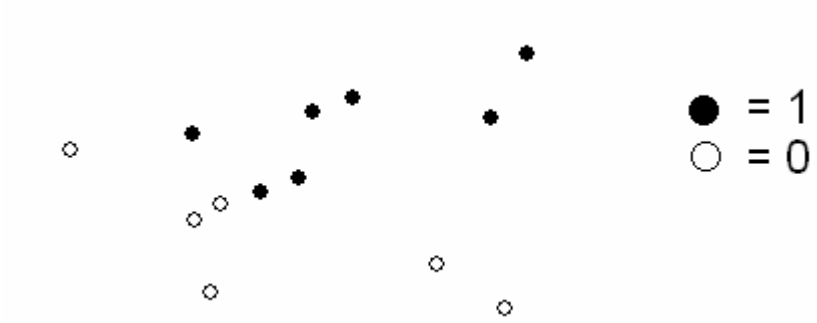


Indicator Variograms

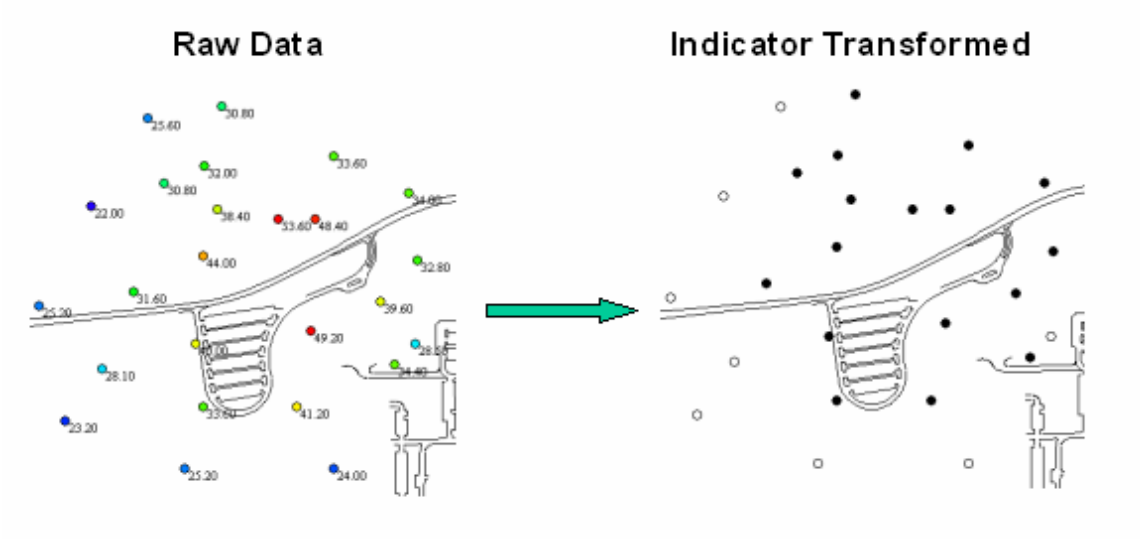
To this point, all variograms we have calculated have been on the actual data points. This type of variography is necessary for ordinary kriging, cokriging, and sequential gaussian simulation. Indicator kriging and sequential indicator simulation require variography be calculated on the indicator transformed values. When performing an indicator transform, you need a threshold value t. All sample values greater than t become equal to 1, and all values less than t are 0. We can write the indicator transform of the data as follows

$$I(u_i) = \begin{cases} 0 \; if \; u_i > t \\ 1 \; if \; u_i \leq t \end{cases} \quad for \; all \; i$$

If we plot the transformed data values, they would look something like this.



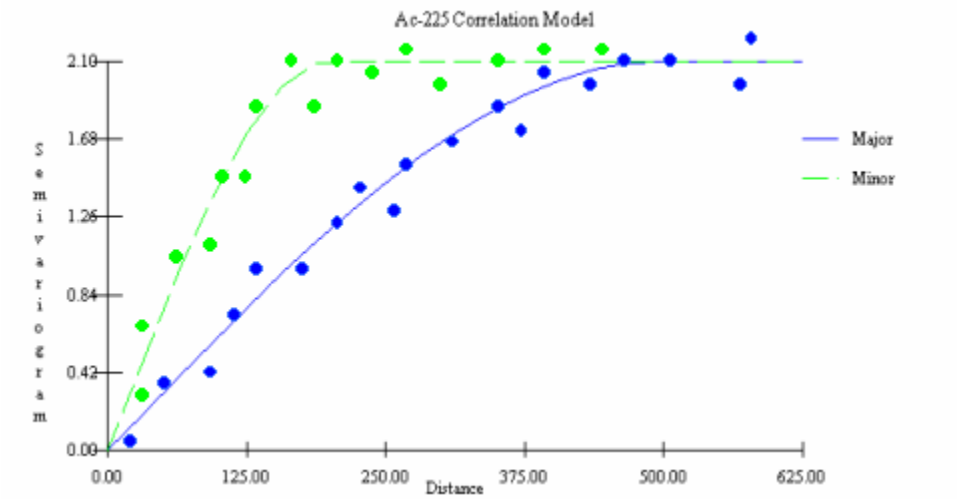
The following example shows Arsenic data that has been indicator transformed at a threshold of 30 mg/kg.



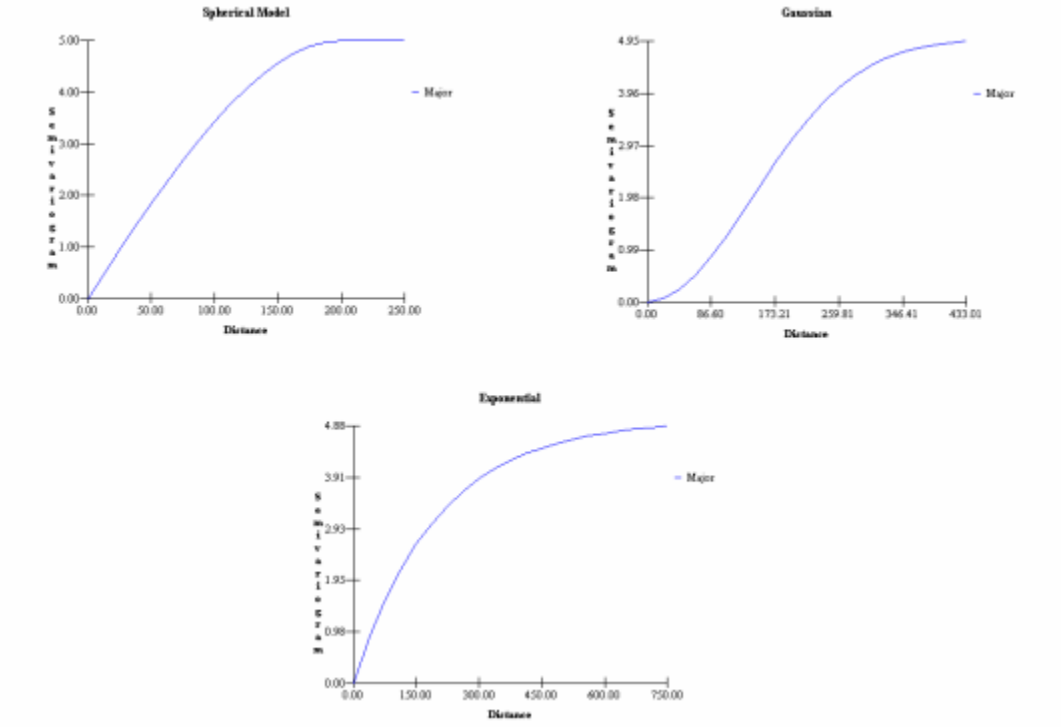
Calculations of the variogram of the indicator transformed data then proceed exactly as in the case of the untransformed data.

Spatial Variogram Modeling

When kriging or simulation models are used, they will require that the semi-variogram values are known for any distance h. At this point, we only have these values for a few discrete distances. Therefore, we need to fit a model to the data so that a semi-variogram value is available for every distance.



SADA provides 3 standard correlation models that provide a great deal of flexibility in semi-variogram data: Spherical, Exponential, and Gaussian.



These correlation models actually are elliptical or ellipsoidal in 2d and 3d space. Typical model plots only show model behavior along the major and minor axis. Recall that these two axes are determined through the user of the rose variogram. While correlation models are defined along all angles between between the major and minor axis, it is sufficient for you to fit the model along the major and minor axis .The specification of these models in space amounts to the specification of an ellipsoidal geometry overwhich the model is extended. the All three models require the same set of parameters and most of these parameters were encountered while specifying a search neighborhood. For a review of how to construct a search window please see the Basic Geospatial Methods chapter.

Ellipsoidal (Domain) Geometries

Major Range – distance to sill or correlation length along the major anisotropic axis.

Minor Range – distance to sill or correlation length along the minor anisotropic axis.

Angle – the horizontal angle of anisotropy

Z Angle(3d) – the angle of anisotropy in the Z plane (equal to the Dip parameter in experimental variography).

Z Range(3d) – a value describing how anisotropy behaves in the z minor direction, relative to major axis.

Rotation(3d) – how the anisotropic ellipsoid is rotated about its major axis.

Model Properties

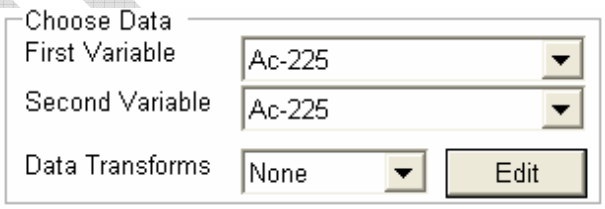
Contribution – The model's contribution to the sill (maximal model value)

Nugget – where the model should cross the y axis (white noise)

SADA also has a varigram fit function to assist the user with assigning these values.

If you have not already done so, open SpatialCorrelation.sda in SADA. This file contains 2D soil data (Ac-225) and 3d groundwater data (Chlordane). We'll use these two data sets to start getting familiar with semi-variogram calculations.

Make sure you have selected Soil and Ac-225. Then n the interview drop-list select Model Spatial Correlation. We will use this file to demonstrate how to assess and model spatial correlation for all forms of kriging and simulation available in SADA. Each method requires different transforms of the data and/or multiple structures. Click on the step "Correlation Modeling". The parameter window associated with this step appears involved and to some degree it is. Let's start by looking at the top of the parameter window at the information block called Choose Data.



Here you see a First and Second Variable. Recall that in a semi-variogram construction there is a head and tail variable. This is where you specify these two variables. At the moment, they are both set to Ac-225. This means that we are only interested in examining the correlation structure of Ac-225 measurements. If you select either drop list arrow, you'll notice that Ac-225 is the only option. It is also possible to specify another variable an examine the spatial correlation between Ac-225 and a "helper" data set such as field survey data set for example. So called helper data or secondary data is necessary for cokriging and cosimulation techniques as is the correlation structure between helper and primary data (eg Ac-225). In the section below on correlation modeling for cokriging, we'll show you how to add "helper" data and change the second variable to this helper information.

The Data Transforms parameter contains three methods for transforming the data: None, Normal Score, and Unit Transform. None means that raw data values will be used only.

When performing a sequential gaussian simulation (SGS), the data must be normally distributed. Unfortunately most data sets are not normally distributed. However, a normal score transformation can convert any distribution into a normal distribution. The simulations are run in this space and then back transformed into real space. The following explanation of the normal score transform is based on Deutsch and Journel (1992).

Each data value will be enumerated by z_i (z_1, z_2, …, z_n). The cumulative probability associated with each of these will be c_i. So c_5 = prob($Z < z_5$) and so forth. The normal score transform of z_i will be noted as y_i which is then calculated as

$$y_i = G^{-1}\left(\frac{c_i + c_{i+1}}{2}\right)$$

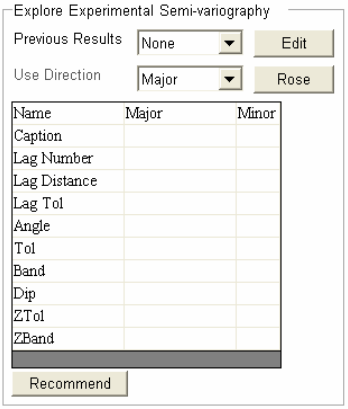
Where G(y) is the standard normal cumulative distribution function. Then $y_c = G^{-1}(C)$ is the corresponding standard normal c-quantile. SADA uses the *nscore* program developed by Deutsch and Journel (1992, p. 211) to perform the normal score transform.

The Unit Transform option simply rescales the data so that there is a mean of zero and a variance of 1.

$$y_i = \frac{z_i - \bar{z}}{\sigma^2}$$

This drop list can also contain any indicator cutoff thresholds you may defined. The indicator thresholds are defined using the Edit button. We'll visit this shortly when we talk about indicator estimation and simulation.

The information block below the "Choose Data" block is called "Explore Experimental Semi-Variography". At the top is a drop down list of previous semi-variogram evaluations we may have done and saved. The edit button just to the right of this drop-list allows you to manage your saved results. At the moment, there is nothing in this list box to choose from. Later we'll show you how to save your results for both semi-variogram plots and rose maps alike.



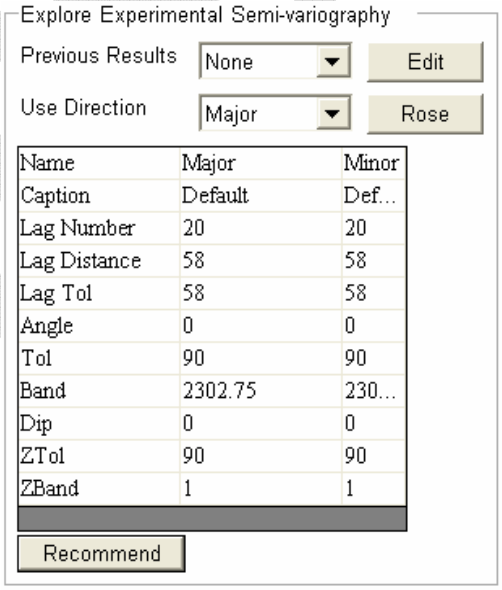
Next is the "Use Direction" option. This drop lists allows you to choose from Not Used, Major, Minor, or Both. The options in this list box are a bit misleading at first glance. They do not refer directly to the major or minor direction of anisotropy, but rather to the columns in the parameter table just below them. In that parameter table, there is a major and a minor column. Typically investigators will parameterize the search window for the major direction in the major column. Likewise, the minor column holds parameters typically associated with the minor direction. However, you do not have to restrict yourself to this idea. In fact, when you begin you will likely not know the major or minor direction. It is better to think of this as the ability to parameterize at most two search directions at once. The "Use Direction" drop list allows you to tell SADA which column of parameters you want to use at the moment. Sometimes you don't wish to use either because you only want to plot the model and not the point values (see below).

Next to the "Use Direction" drop list is the Rose button. This will generate a rose map for you using one of the columns. So you cannot generate a rose map for Not used or Both. You must select either the major or minor column to base the rose map upon.

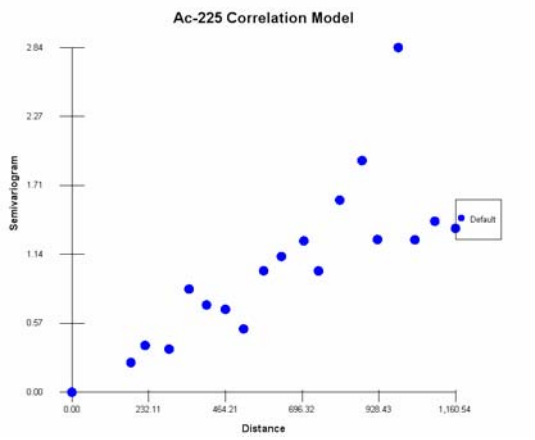
Practicing A Bit

For practice, let's step through a couple example semi-variogram calculations. The first step is to explore semi-variography values in the search for any anistropy that may exists. Most of your time will likely be spent on this step. The first thing we need to do is populate the block of information called Explore Experimental Semi-variography. At the moment its blank and perhaps even a bit intimidating. If you have a good handle on how the data are patterned across your site, you can make some good educated guess as to what the the parameters might be. From that point you can use actual variography results to further refine your parameter selection.

Another way to begin, especially if you are relatively new to this, is to use the Recommend button. The recommend button use some rough rules of thumb to get you started. You should under no circumstances believe that these recommendations are optimal in any sense. Rather they serve as a good starting point. Push the Recommend button for exploring experimental variography now and say Yes to the confirmation message box that appears. You should have something like this:



At the bottom of the parameters window there is a “Show Me” button. This will produce a semi-variogram plot of using the selected columns(major and/or minor). Go ahead and push this button now. SADA produces a standard semi-variogram plot.

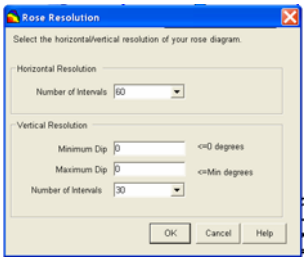


From this plot, you can see how data are varying as a function of the their separation distance. This is the central concept in geostatistical modeling and will drive the selection of model parameters. Now this is an isotropic or omni-directional sem-variogram. By omni-directional, we mean that the cone geometry that we discussed earlier is so wide that for any given point it includes the entire site. Some data are isotropic and so an omni-directional variogram is completely appropriate.

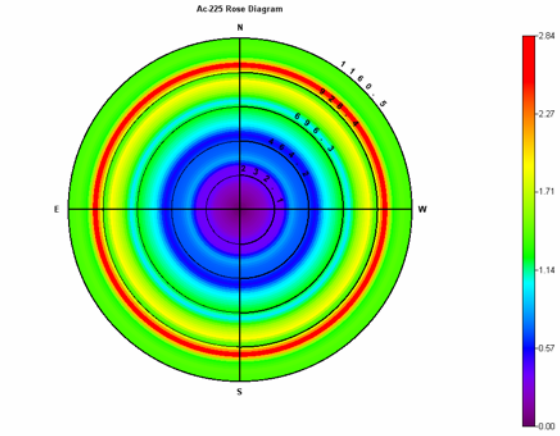
Let’s modify the parameters to see if there is a dramatic change when we narrow the focus of the semi-variogram cone and look in a particular direction. To do that we will need to change only two parameters: Tol and Band. For 3d applications you would also need to change the Ztol and Zband. Tol is the angular tolerance that controls how wide our cone is. Right now its 90 degrees which means regardless of the horizontal angle (Angle), widen the cone so much that it sees to both the right and left 90 degrees. In other words, everything (recall that “behind” each point is exactly the same result as in “front” of each result). Change the Tol to 45 and the bandwidth to 1000. Press Show Me again.

You can see that there is a difference but it is difficult to determine what changed. The distribution of points changed and the y-axis also changed. A better way to evaluate directional variograms is the rose map. The rose will incrementally increase the Angle parameter from 0 to 180, calculate the semi-variogram values, and perform a linear interpolation between points. This will make visualization much easier. One way to approach this is to set your tolerance back to 90 degrees (do this now). This will perform an omni-directional rose map that will look like a wagon wheel. Using this as a base, we can slowly tighten the spread of the variography cone and see how the rose map changes. Sites with anisotropic data will move away from a wagon wheel look and will start showing troughs of low semi-variogram values in the direction of anisotropy.

We will use the Major column as our set of parameters so make sure you have Major selected for “Use Direction”. Press the Rose button.

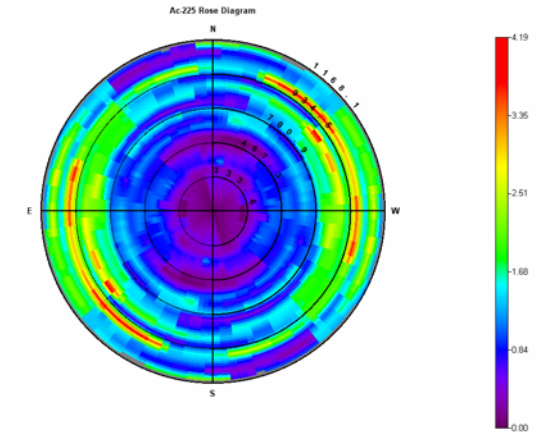


This window is asking how fine of an angular resolution do you wish both in the horizontal angle and vertical angle. The default is 60 intervals between 0 and 180 degrees or about ever 3 degress the semi-variogram is recalculated (note that 180-360 is the mirror result of 0-180). We are not currently dealing with 3d data but if we were, we could specify a minimum and maximum angle to define a range of reasonable vertical angles. This range is then divided into a number of intervals as well. We'll keep the default values and press OK. An isotropic or omni-directional rose diagram is produced.

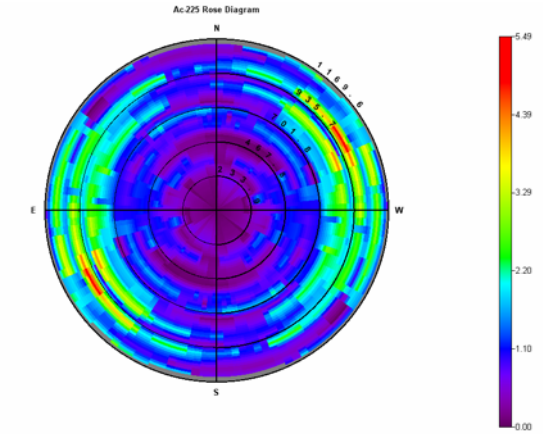


If the data are truly isotropic (direction doesn't matter), then this wheel like pattern will continue to hold even as we vary parameters under the major column. If anisotropy exists, we will see a trough of low values open up in the direction of anistropy. So the values are not as important at this point as the relative behavior of the rose map: wheel or trough?

To demonstrate let's begin by changing the Tol parameter under the Major column to 45 (degrees). This will narrow the field of view for the model as it sweeps around the site. With this narrowed view, if anistropy exits, it will show up in the scan. If not, a "wheel" like results will persist even though the values may be different. Press the Rose button again and accept the interval defaults.



A trough of low values appears to be opening up as the wagen wheel geometry begins to fade. Let's narrow the angular tolerance even further to 30. Press Rose and accept the defaults.



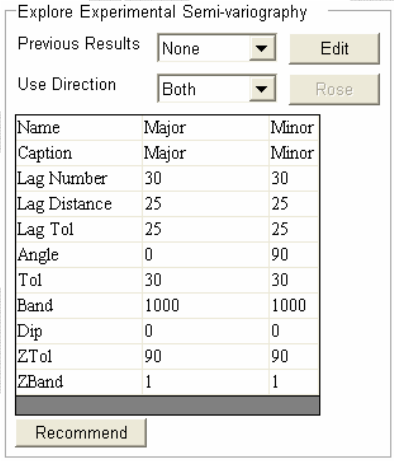
By now a fairly pronounced trough has opened up along the N-S transect indicating anisotropic conditions in that direction. This trend is exactly what we would expect even by just visually inspecting the data map itself (note that the data map showed a trend running N-S). Therefore, the Angle of 0 degrees is entirely appropriate as the direction of anisotropy. We will keep the tolerance at 30 degrees and the bandwidth at 1000. At the bottom of the parameter press the Show Me button and this will show you the exact semi-variogram values for a N-S direction of anisotropy (assuming you have Angle =0 and tolerance = 30).

Let's increase the number of lags and shorten the lag distance and tolerance. This will give us more points to visualize. Change the Lag Number from 20 to 30 and shorten the lag distance and tolerance parameters to 25.

Name	Major
Caption	Default
Lag Number	30
Lag Distance	25
Lag Tol	25
Angle	0
Tol	30
Band	1000
Dip	0
ZTol	90
ZBand	1

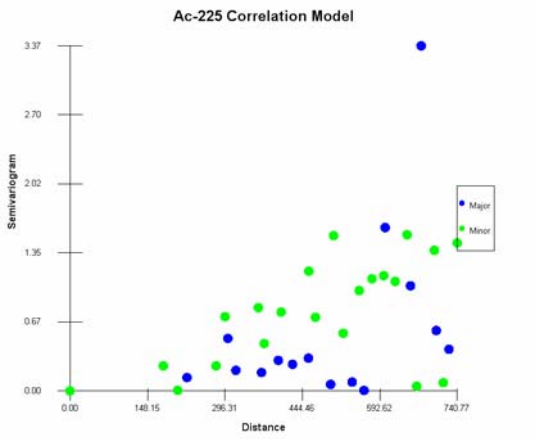
In addition, let's parameterize the minor direction. This is necessary only if anisotropy exists. If it does not exist, you would use only the major direction and keep the Tol parameter at 90.

Enter exactly the same values for the Minor column as for the major column with the following exception. The Angle parameter should be 90. This is because the minor and major axes are perpendicular to each other. Therefore, if the major direction of anisotropy is 0 then the minor must be 90. If the major direction were 45, then the minor direction would be 135 and so forth. In the caption row, change value to Major and Minor for the major and minor colum respectively. One last thing, you will need to change your direction from Major to Both. Your semi-variogram parameters should look like the following.



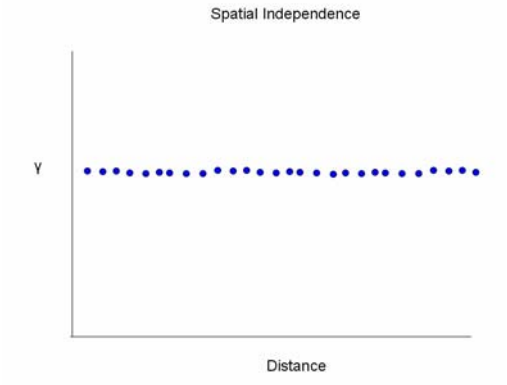
Name	Major	Minor
Caption	Major	Minor
Lag Number	30	30
Lag Distance	25	25
Lag Tol	25	25
Angle	0	90
Tol	30	30
Band	1000	1000
Dip	0	0
ZTol	90	90
ZBand	1	1

Now at the bottom of the parameters window press the Show Me button. The following result appears.



Save your file. The green dots are semi-variogram values calculated along the E-W transect (90degrees). The blue dots are the N-S semi-variogram values. Notice how the green dots are almost always greater than the blue dots. This is because data in the E-W direction are more variable over short distances than the N-S data. The next task is to model these values. We will get to that shortly. First we will discuss semi-variography requirements for each geostatistical model. For each case, you will approach the problem in a manner similar to what we've just described. Of course, situations vary a lot and rarely will you get well behaved semi-variography results for small environmental data sets. You may need to use judgment and a great deal of patience in your evaluation.

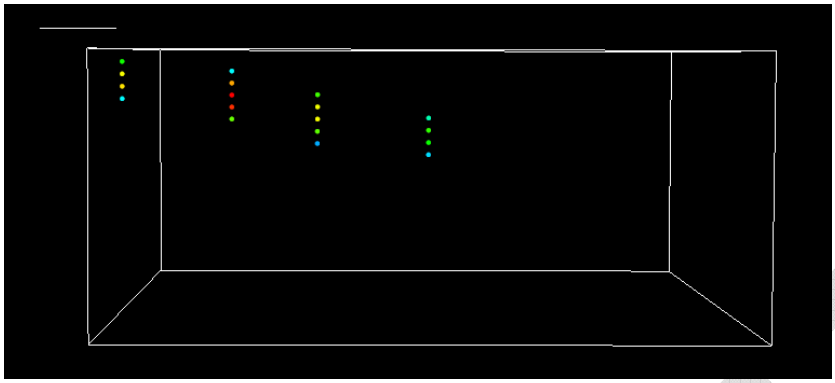
If you cannot get a well behaved semi-variogram, you still need to be careful about assuming there is no spatial auto-correlation. In the figure above, both the green and blue dots are bouncing around quite a bit. There is however, still a trend of increasing variability with distance. If ignore this and assume no correlation among the data, you are saying this is the behavior of the semi-variogram values.



3D Variography

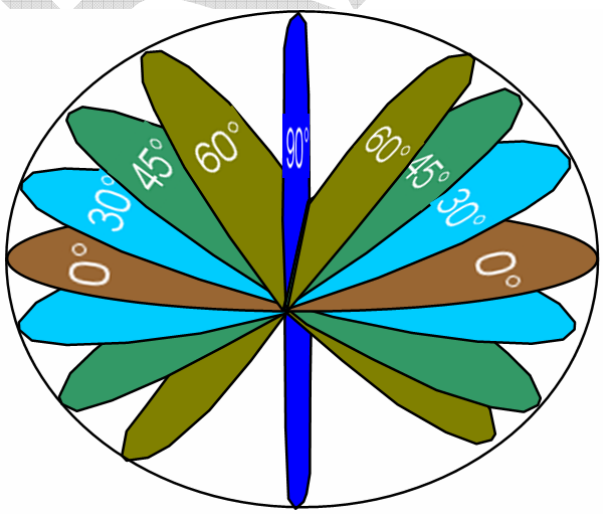
In fact, this type of result is evidence of the lack of spatial correlation. We will demonstrate how to derive and peruse a 3d dimensional variograph evaluation in SADA. Swithch from soil to groundwater and we'l look at Chlordane. The site has been setup with a vertical layer design of 10 foot layers from 0 to 52 feet deep (check the setup the site step). If you use the vertical layer drop list on the main tool bar you can see that a trend of high values is running not only from north to south but dropping vertically as well. The following image was taken from the 3d Viewer (see chapter on mastering the 3d viewer) where low value points have

been hidden to show the trend in the upper values. The shot is due East so your looking at the transect of high N-S values that are dropping vertically as the move south.



There is clearly a correlation structure for high valued samples. We want to capture that information in our correlation assessment. Note that the type of correlation structure may vary across the range of sample values. For example, if you remove the high values and look at only low valued samples, you can see there isn't much direction to their behavior. In fact, the structure looks very omni-directional with very few difference in values outside the plume. Both of these behaviors can be capture using indicator kriging or simulations. We'll show how to do this shortly. Right now, we'll focus just on how to get multiple rose maps that can not only show you horizontal variation but vertical variation as well.

If you could see a three dimensional variography result, it would look like a sphere comprised of many, many points that vary in color and density.Like a tight cluster of stars in an circular or ellipoidal shape. SADA will allow you to see cross sections of this cluster one at a time. If you could see the cross sections in 3d space, each Rose transect would have the following geometry.



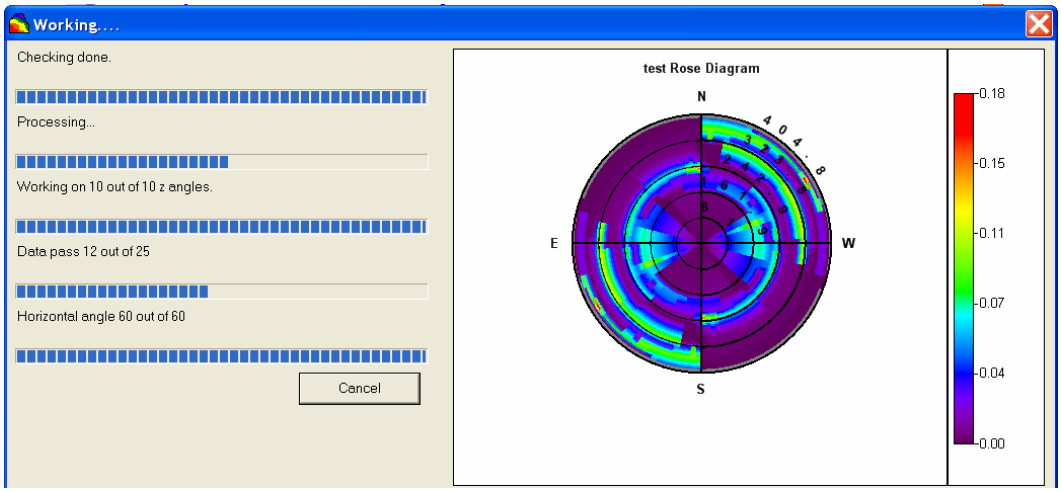
Of course each transect would not be a single color but painted with a rose map. Be aware that when you are doing 3d variography, the parameters of horizontal angle, horizontal angle tolerance, and even the lag and lag distances now pertain to the direction of the cross section. For 2d data sets, this is in fact horizontal. But in 3d its horizontal relative to the dip angle.

SADA will generate all these transects for you and allow you to scroll through them.

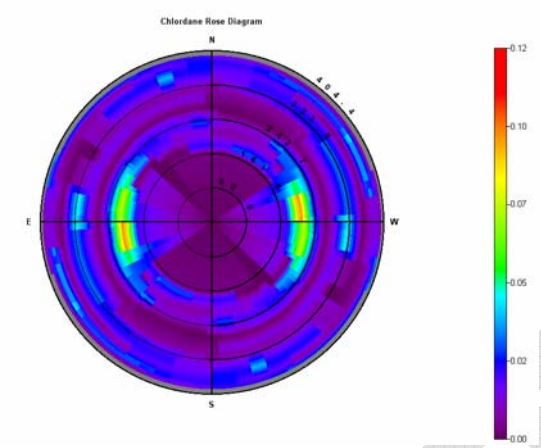
For the Chlordane data set, use the following parameters. The data has a smaller spatial range than the Ac-225 data, so we'll need to modify the default parameters some. Make sure you have Chlordane selected for both the first and second variable, no transform, major direction only and that the parameter values are the same as in the image below.

Now press the Rose button. Accept the default horizontal parameters. But for the vertical select a minimum dip of 0 and a maximum dip of -5°. Select the number of intervals to be 10. So SADA will calculate 10 transects. One at 0°, another at 0.55°, another at 1.11° and so forth. Recall that we want to use a small maximum interval of 5° because with such a vertically flat site, values much more than this will cause ellipsoidal geometries that jump out of the site to quickly.

You will notice now a larger progress window appear. Three dimensional variographys can take considerably more time and SADA is showing you the progress as calculations are being done. SADA will look at a single transect and calculate the variography data a little at a time so that you can see progress unfolding. For each new transect, the rose image will start out crude and become finer as the total number of samples is used.

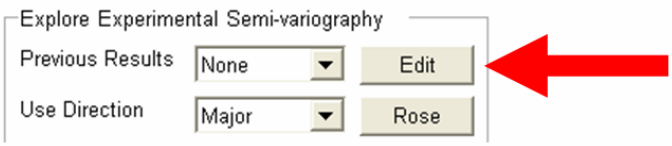


When SADA is finished you will be presented with the first rose transect.

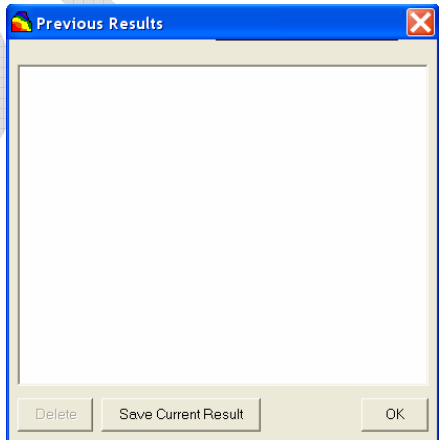


Notice that the depth layer droplist is now a transect selector.

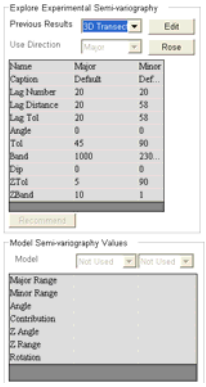
You can now switch between the transects and observe variography results along each vertical dip transect. Before you begin exploring this feature, lets save these results. Variography generation can take some time and you don't want to have to repeat the calculations each time you want to see the results. The Rose button runs the calculations from scratch every time you press it. To avoid this lets save the resuls by pressing the edit button next to the "previous results" drop list.



You will be presented with the Rose/Correlation Result manager. This manager will contain a list of all the results you have previously saved. You can delete and save new results here. The list is empty as no previous results have been saved at the moment. Press the "Save Current Result" button.



Enter "3D Transects"and press OK. Press OK to the Rose manager as well. In the Correlation Modeling parameter window select 3d Transects from the previous results drop list.



Notice that the parameters we used in this calculation are entered into the major/minor column but everything is disabled. This makes sense because you cannot modify parameters of a stored result. If you want to make parameter changes and rerun the model, select None in the "previous results" drop list again.

As you change the Z dip in the main tool bar, you will notice that for each dip value, there is a relatively constant N-S trough. This is what we would expect with the presence of the high value samples. The difference between the transects is not markedly different although the trough intensifies as you get around -1 to -2 degrees. This is also expected.

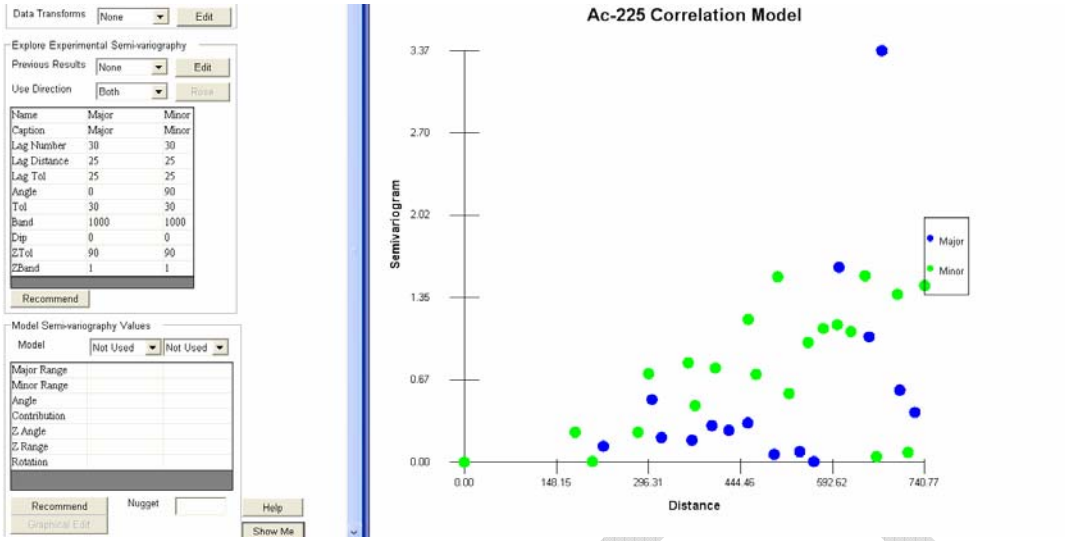
However, based on the visualiation of the data in 3d, there is likely two different correlation structures going on here. Low values are uniformaly scattered in an almost omin-directional structure. The high values are dipping along a N-S transect. To really capture this we will need to use an indicator based formulation. This is discussed in the section on correlation modeling in preparing for indicator kriging below.

These two applications are really only an introduction to correlation analysis. Data sets will widely vary and some patience will be required to sufficiently capture patterns observed in the data.

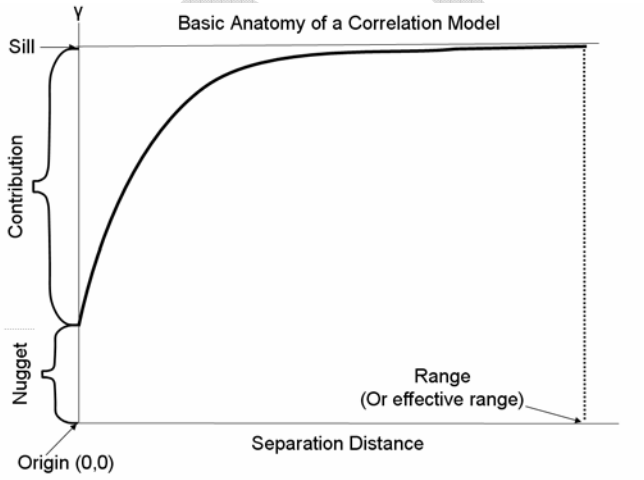
Correlation Modeling in SADA

When kriging or simulation models are used, they will require that the semi-variogram values are known for any distance h. At this point, we only have these values for a few discrete distances. Therefore, we need to fit a model to the data so that a semi-variogram value is available for every distance. Let's do this for the Ac-225 model.

Switch back to Ac-225, select Model Spatial Correlation as the interview and click on the Correlation modeling step. Make sure you parameters are as follows and press the Show Me button in the lower right of the parameter window.



We need to fit a model through these points. There are three possible correlation models in SADA. Each one has a particular shape and models can be combined to create additional shapes. Each model has the same basic anatomy defined by sill, nugget, contribution, and range parameters. Please refer to the following picture as we talk about each parameter.



Sill

The sill is the highest value that the model will reach or asymptotically approach. The sill is equal to the nugget plus the contribution.

Nugget (n_0)

The nugget parameter describes how the model behaves as it approaches the origin of the graph. For example, a nugget of zero means the the curve of the model passes through (0,0). A nugget of one means that the model passes through (0,1). The interpretation of the nugget is that for some reason as data become closer and closer (asymptotically closer), their variation does not approach zero (Goovaerts, 1997). This is a useful feature that is used to account for things like sampling errors and/or sudden changes in an attribute (think of a gold nugget buried amid sand).

Contribution (c)

This defines how high model values increase relative to the nugget value.

Range (a)

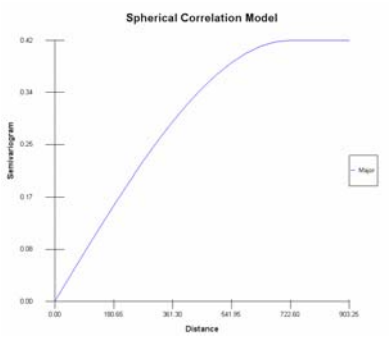
This is the point at which the model reaches its sill value. Some models never reach their sill value, but rather only asymptotically approach it. In this case, the range is sometimes referred to as the effective range or that range where it is essentially the same as the sill.

SADA provides three models. Based on Deutsch (1995), we have,

Spherical

This model rises from the nugget value to its sill (n_0 + c) value at range (a) and then flattens out immediately.

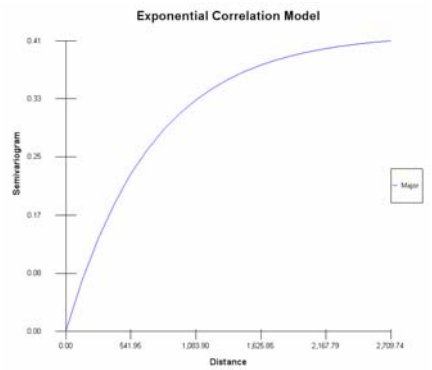
$$g(h) = nugget + c \cdot \begin{cases} 1.5 \cdot \frac{h}{a} - 0.5 \cdot \left(\frac{h}{a}\right)^2 & \text{if } h \le a \\ 1 & \text{otherwise} \end{cases}$$



Exponential

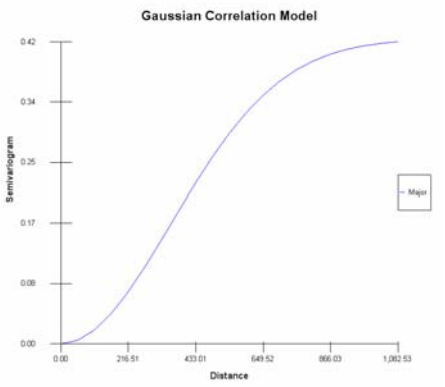
This model rises from its nugget value and asymptotically approaches the sill at an effective range of 3 times a.

$$g(h) = nugget + c \cdot \left[1 - \exp(-\frac{h}{a})\right]$$



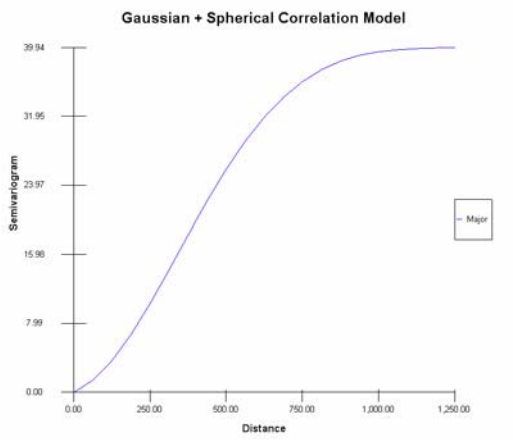
Gaussian

This model rises from its sill and forms an "s" like shape as it asymptotially approachs its sill at an effective range of $3\sqrt{a}$. This model is known to cause some instabilities in the kriging solutions. It should be used only with great caution.



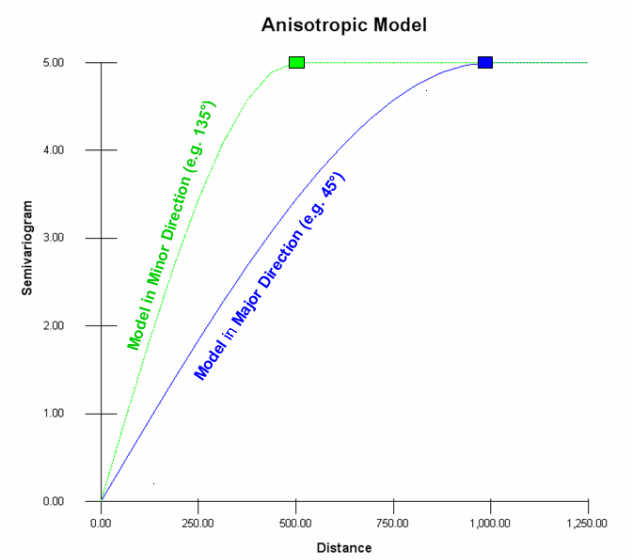
$$g(h) = nugget + c \cdot \left[1 - \exp(-\frac{h^2}{a^2}) \right]$$

It is possible to nest these models to create different shapes. For example, the following figure combines a Gaussian model with range 500 and contribution of 30 with a Spherical model of range 1000 and contribution of 10. A nugget of zero is used.

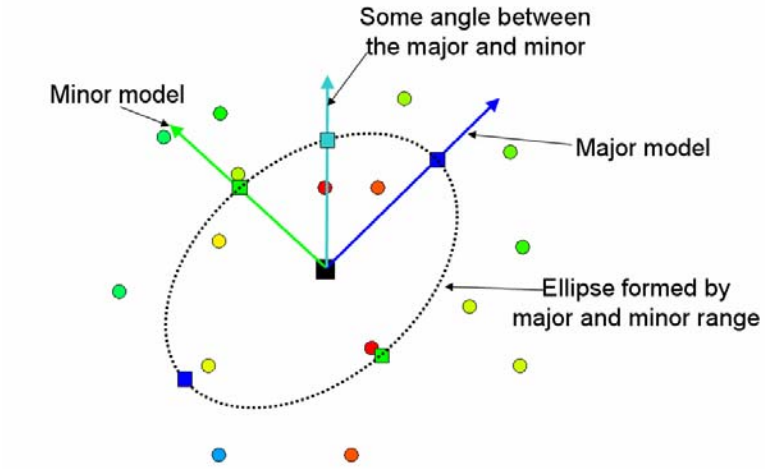


Notice how the shape begins with a slight "s" hook near the origin and then resembles a spherical more at the larger distances. Typically, a single model is usually sufficient.

Now if there is anisotropy, then one must also specify a minor direction model. The only parameters that distinguish a major model from a minor model is the major range and minor range. The two models must be of the same type (spherical, exponential, or gaussian) and have the same sill, nugget, and contribution. In the following image, a spherical model is used with a major range of 1000 and a minor range of 500.

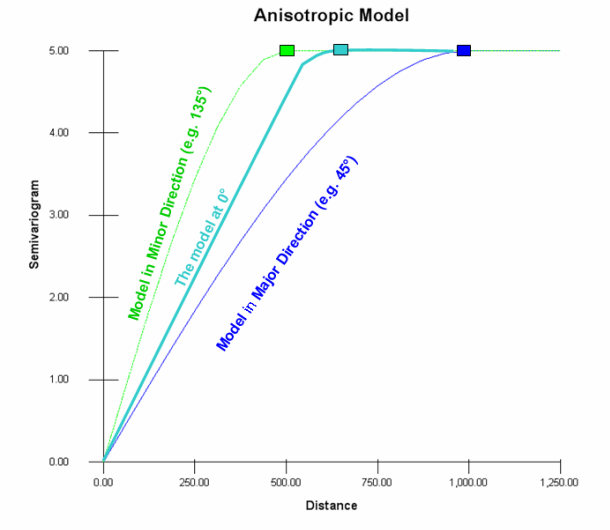


You will often see semi-variogram models shown in this manner: a minor direction model and a major direction model. These models are used to decide what weights will be assigned to data points in the neighborhood of the node we wish to estimate. So clearly, if a data point lies exactly in the major direction (relative to the node) then the major model will be used. If a data points lies exactly in the minor direction (relative to the node) then the minor model will be used. But what about all the directions in between? Most data in fact will lie somewhere off one of these directions. In the figure above, the range/sill point of both models is denoted by rectangles. Let's show where these would be found relative to some arbitrary node we wish to estimate or simulate.

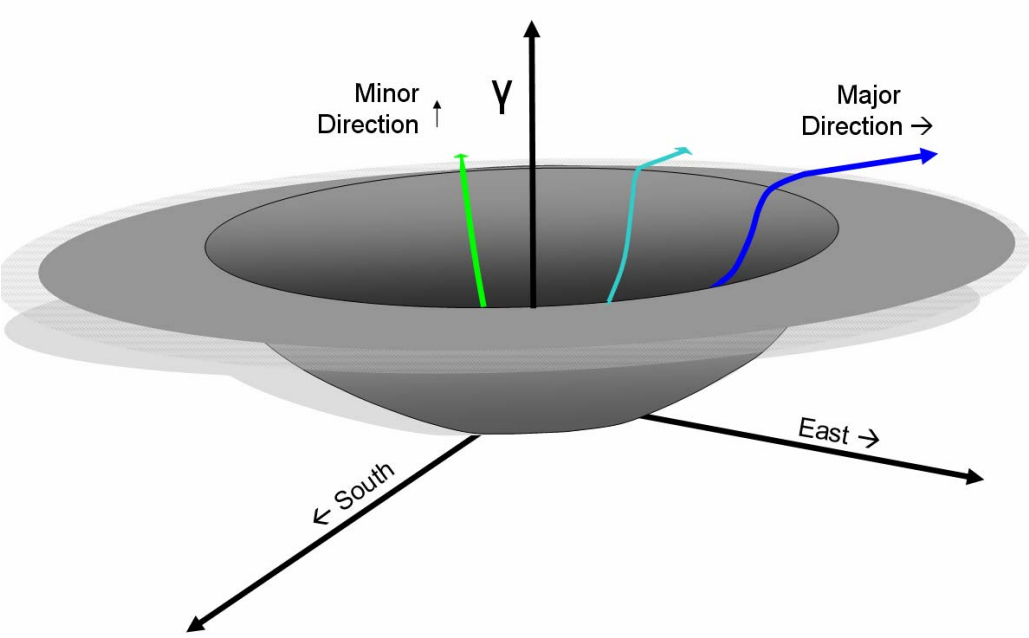


Suppose you were to stand a z-axis up on its end at the center of the node we wish to estimate and then draw the major correlation model along the major direction and draw the minor model along the minor direction. The image above is what you might see if you were to look straight down from above the site. The major and minor models would just look like straight lines because you're seeing them from the above. Now the major range and minor range rectangles seen in the former image are now drawn in their proper places once again in the graph above. Since correlation models are mirror or reciprocal in nature, you can also draw them on both sides of the node. Notice how these can form the vertices of an ellipse. This ellipse will pass exactly through the range/sill point on the major model and exactly through the

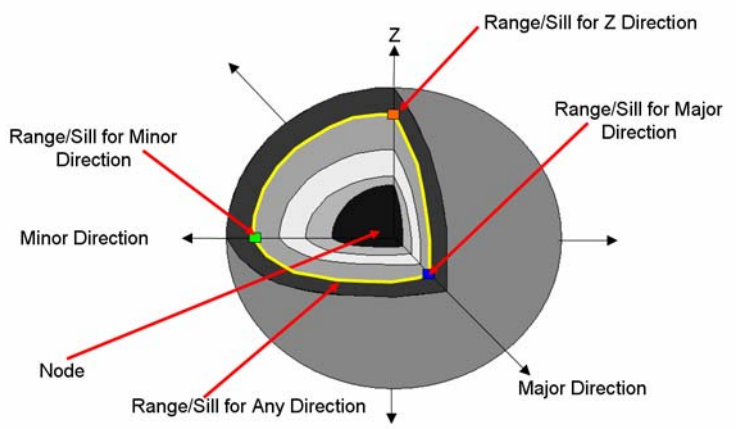
range/sill point on the minor model. It will also exactly pass through range/sill point of any angle found between them. In the image above, the light blue line is some angle (e.g. 0°) between the minor and major angles. This is how any angle between the major and minor direction is handled. In fact, one could also plot the model along this angle in a standard graph along with the major and minor direction.



The following image is a 3d representation of this same model viewed from an off angle. If you could see every single model between the major and minor direction it would form a 3d object not unlike a bowl. The interior of the bowl would follow the shape of the model. So it would appear gaussian, exponential, or spherical. Outside the range, the surface would flatten out and continue indefinetly. In the following image, the surface is colored in gray, but the three specific directions are shown in the image.



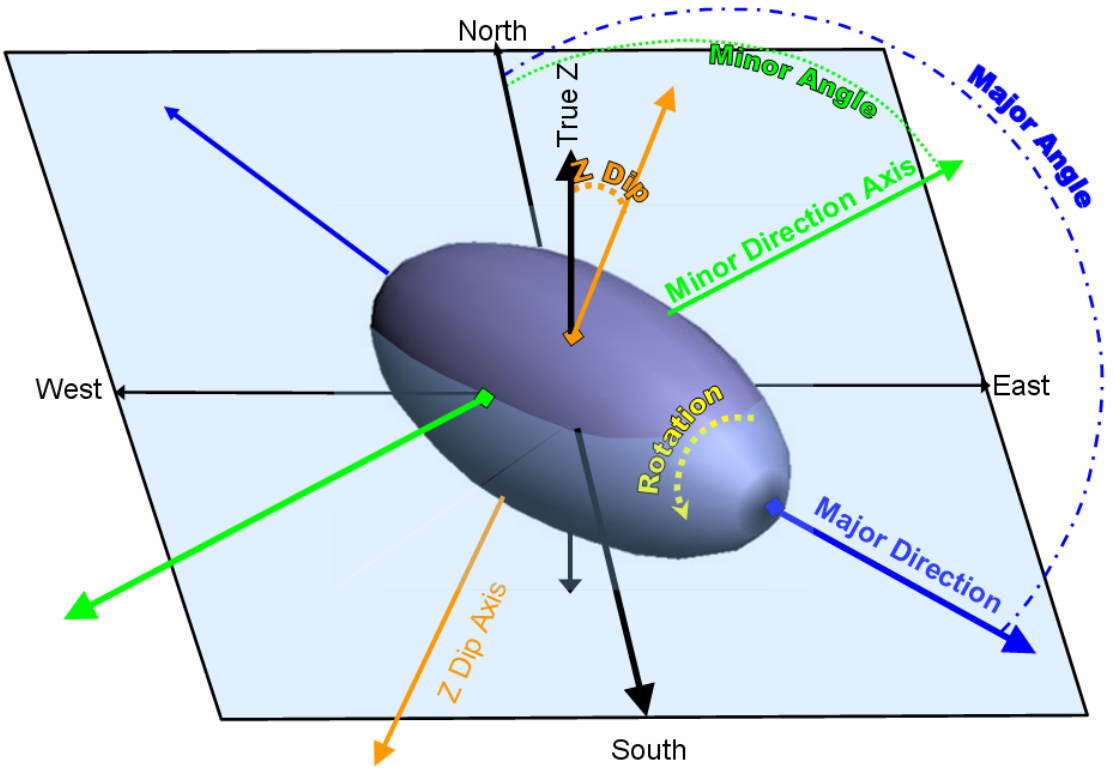
Correlation modeling in 3d follows the same type of thinking. It is a little more difficult to visualize. Since we are talking true 3d, we can't use the z-axis for our semi-variogram axis anymore. This will have to be the true z-axis (depth below surface). Instead, the model values will have to be represented by color gradations. In 3d, we have a correlation volume rather than a planar geometry as seen above. It woud look like a sphere or ellipsoid with the unestimated node at its center. If we could cut away a section of the sphere we might recognize some familiar items.



So at the center is the unestimated node. The blue and green range/sill rectangles for the major and minor directions are shown just inside the outer hull. They are now joined by the z range/sill shown in orange. The ellipse range/sill that we showed earlier is now a range/sill isosurface. In the image above, it's the yellow peel or line you see in the transect. It's a bit complicated, but spend some time with it and you'll get it.

It can get a bit more complicated in 3d. There is also a z angle and a rotation parameter as well. Z angle can be important. Rotation has less use in practice but there are applications

where it may be useful. In the previous image we saw a depiction of an isotropic 3d correlation model. We know its isotropic by looking at it because it's a sphere. To show the effect of these new parameters in 3d, we'll need to use an ellipsoid. We'll remove the cutaway to simplify the picture and see what the ellipsoid will look like when we apply z angle and rotation.



*Base Ellipsoid shape taken from Wikipedia Ellipsoid Topic. Picture freely available for use under GNU. (http://commons.wikimedia.org/wiki/File:Ellipsoid_3d.jpg)

So in 3d space the correlation model geometry is a ellipsoid. We have elected to trim the ellipsoid back to the range/sill surface for in all directions. The range/sill location for the major, minor, and z directions are noted again by rectangles of corresponding color. They are by definition on the surface of the ellipsoid if we are triming th ellipsoid just to the model rane in each direction.

The ellipsoid is tipped forward into the horizontal plane through the node center (noted by light blue area) by the Zdip parameter (in negative degrees). The rotation angle is how this ellipsoidal shape is rotated along is major axis. Just as in search neighborhoods, great care must be made in specifying the z angle. This will normally have very small values due to the "flatness" of most environmental sites (See previous discussionon search neighborhoods). The z dip should be motivated by evaluating semi-variogram values

Correlation modeling in SADA

In SADA these parameters are found in in the information block at the bottom called "Model Semi-variography Values". Again a number of parameters are there and they are empty. There are three columns. The first column just describes the parameters that are needed by

the models. These are the parameters that define an ellipsoidal geometry for the correlation model. As mentioned earlier, there are only 3 possible models in SADA: spherical, exponential, and Gaussian.

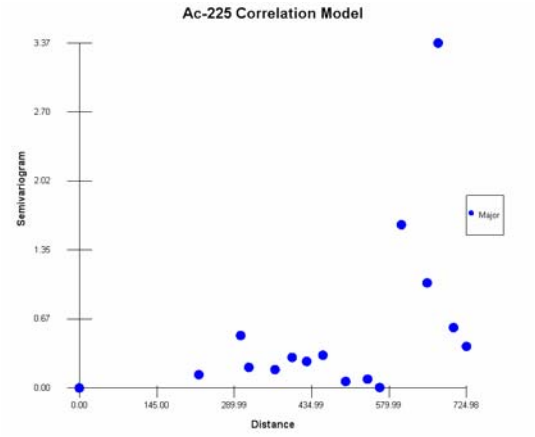
If you click on the first Model drop-list, you will see these three options. If you click on the second Model drop-list, you will see these three options repeated. You have access to these two models because SADA allows you to use nested models. As previously mentioned, a nested model is really just the sum of to individual models. This allows you to form different shapes that might fit your data better. Most of the time, one model is sufficient.

Because of the alignment of each model under the Major and Minor columns in the variography block above it, it is easy to make the mistake that the first model models the major column and the second model models the minor column. This is NOT the case. Both the major and minor direction are modeled with one model only. This model can be a nested model (by using both model columns) but still there is only one model.

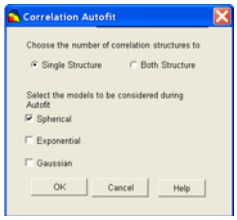
So how do you take one model and fit it through two sets of points? Recall that the semi-variography shown above is presenting the variography calculations along only two directions: 0 deg (North-South) and 90 deg (E-W). Each of the three models will have a major and a minor range parameter. This is the only way they can be distinguished from each other. It is a limiting factor but mathematically necessary for reasons outside the scope of this guide. The best way to understand how this is done, is to work through an example.

From a practical view, the plot of semi-variogram values for both the major (blue) and minor (green) directions is a messy result. It can be difficult to determine how to choose a model that will fit through these points. Methods such as minimizing the least squares is not recommended (ref?) for a variety of reasons discussed shortly. As a result, many analysts fit the model by hand. SADA does provide a minimize by least square option although it is recommended only as a start. To make the process more tractable, let's begin first by dealing with the major (blue) values first.

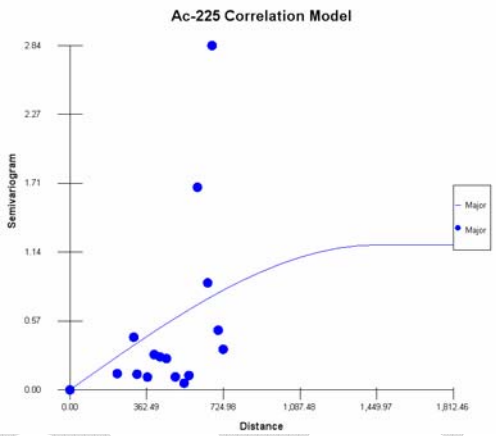
In the information block "Explore Experimental Variography", for the use direction drop list, select Major and press the Show Me button at the bottom of the parameters window. You should get the following result.



We will now use the recommend button to attemp to auto-fit a line through these points. In the semi-variography values block, push the recommend button. The autofit window will appear.



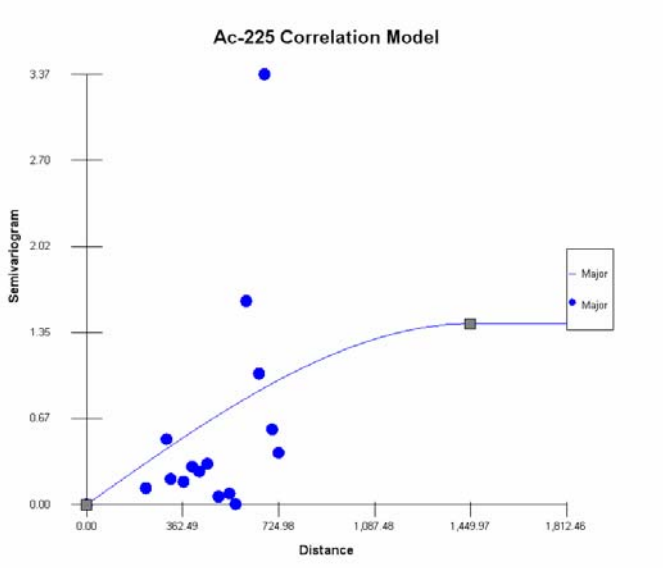
For this example, we'll choose to use only a single structure (not two nested models) and we'll only look at the spherical model. If you select more than one, SADA will try and determine using a least squares comparison which one is best. We'll being by using the spherical model. Press Ok and SADA will present a least squares fit line through the data.



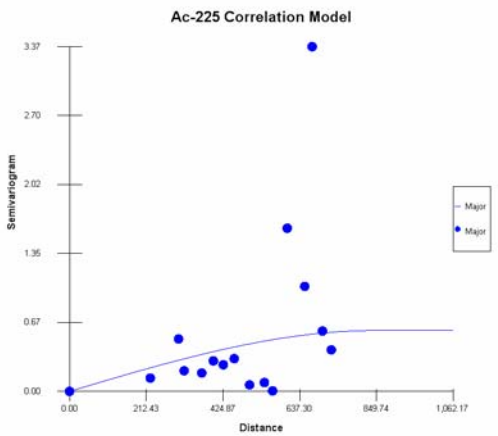
This is not necessarily the optimal answer for two reasons. First, outliers can negatively affect the total fit as the model stretches to accommodate them. Secondly, from a geostatistical perspective, it is better to concentrate on fitting your model at close ranges than worrying about the total fit. Those data points that are closest to any estimation node will weigh more heavily in the estimation and therefore should be more accurately fitted.

You can play around with the model parameters by manually entering different parameter values. Now note that the standard xy type graphs for correlation models do not show the impact of angle choices, or z range. These would only be seen in a 3 dimensional rendering as we showed earlier. You'll need to examine the choice for angles and z range by using rose maps.

Another way to is to graphically edit the model. You can only do this when you have a single (unnested model). With the correlation model up, press the graphical edit button. Two gray rectangles will appear. One at the origin, and one near or at the range/sill point.



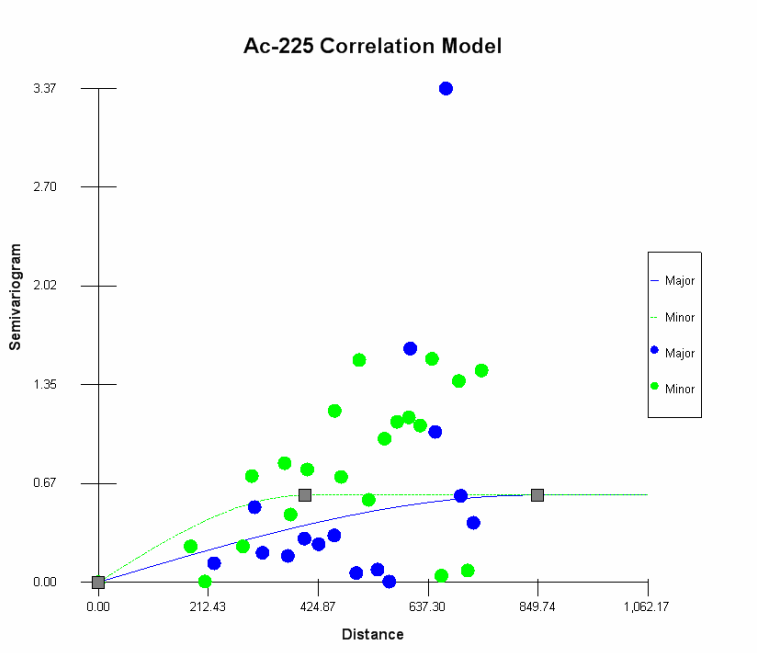
These edit boxes are moved by left mouse clicking and while holding the button down, slide your mouse around. You can move the left most gray rectangle only on the y axis. This is the nugget parameter. You can move the right most gray rectangle up and down or left and right. However, it cannot move below the nugget as this is nonsensical. SADA will simply stop moving the rectangle if you try and move below the nugget. It is possible to edit both the major and minor range at once but it can be more tricky. Spend some time moving these points around and when you're done press the Graphical edit button again. The following image shows the a correlation model modified in such a way.



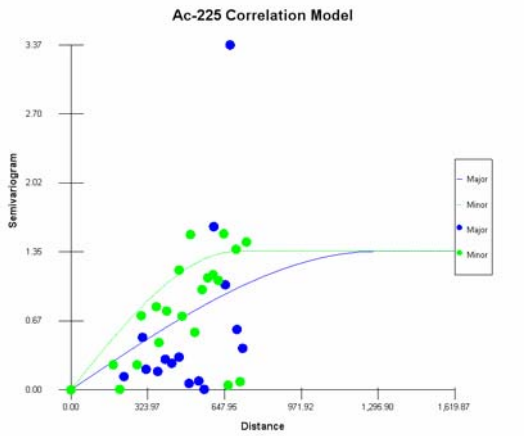
If you do not have anisotropic conditions, you are done. If you do, you can now turn on the green dots and attempt to fit these as well. In the expermintal variography block switch the Use Direction drop list to both. Press the Show me button again and the green dots will reappear.

If you press the graphical edit button again (don't do it), you'll still get only two graphical edit boxes. In reality, there are three graphical editboxes: one for the nugget, one for the major range, and one for the minor range. The problem is that because your major and minor ranges are currently the same, these two edit boxes are sitting on top of each other. Let's fix this first.

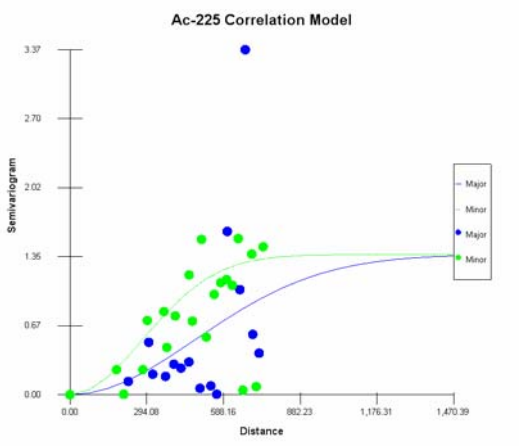
For your spherical model, change the minor range to roughly half the major range and press the Show Me button again. You should now see three buttons.



It is clear that in our image, we may have to increase the sill for both models in order to accommodate the minor direction (green dots) fit. As you work with all three edit points, you will notice that you also cannot make the major range smaller than the minor range and vice versa. If the minor range runs into the major range during editing, they will simply collapse into a single edit box again. You'll have to manually enter a new minor range into the parameter window to separate them. Go ahead and play with the editing now. The following image shows the results of such a manual edit.



Now play with some fitting the exponential model and the gaussian model. You can practice using the autofit routine and graphical edits as well. The following image shows that a gaussian may fit the data best. Unfortunately, gaussian is notorious for causing instabilities in the computations.



Each geostatistical algorithm will require a slightly different variation on the correlation model. We will now talk about correlation needs under each approach.

Correlation Modeling for Ordinary Kriging

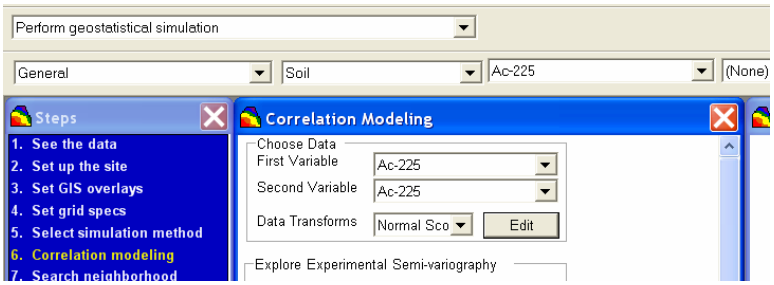
Ordinary kriging requires a single spatial correlation model for the untransformed data. The correlation modeling step is also available under the Interpolate My Data interview. If ordinary kriging is selected as the interpolant, SADA will restrict the variable options to only the contaminant currently selected and the transform option will be restricted to none. Switch to back to Soil and Ac-225. Select Interpolate my Data as the interview. In the Interpolation Methods Step choose ordinary kriging. Click on the Correlation Model parameters step again. Notice that the parameters we entered earlier are still there. This is because we were working with untransformed data which is exactly what ordinary kriging needs. Also click on the data transforms drop list and notice that the only option is none. Because we are doing ordinary kriging, SADA restricts our choice to this option only.

If you worked through the previous section on practicing with correlation modeling in SADA you have done the work you need to do in preparation for ordinary kriging of Ac-225. If you are actually interested in doing ordinary kriging now, you can proceed to the section below on ordinary kriging.

Correlation Modeling For Sequential Gaussian Simulation

Sequential Gaussian simulation (SGS) creates equi-probable realizations of contamination across the site. We will not discuss SGS at this point other than to say it requires a correlation model for the normal score transformed data. The normal score transform was discussed earlier in this chapter and converts any data set into a normally distributed result. SGS requires normally distributed data and since very environmental data sets are normally distributed this is an important step.

With soil/Ac-225 selected in the file SpatialCorrelation.sda, start by switching the intervew from interpolate my data to “Perform Geostatistical Simulation”. Click on the step “Select simulation method” and choose Sequential Gaussian Simulation. Click on the correlation modeling step.



To calculate the semi-variography for normal score transformed data is very easy. In the "Choose Data" block, select Normal Score for the Data Transform. If you are in the geostatistical simulation interview with sequentialy gaussian simulation selected, then this is the only option. You will now perform the correlation assessment activities as you normally would. SADA will do the transformations of the data behind the scenes for you.

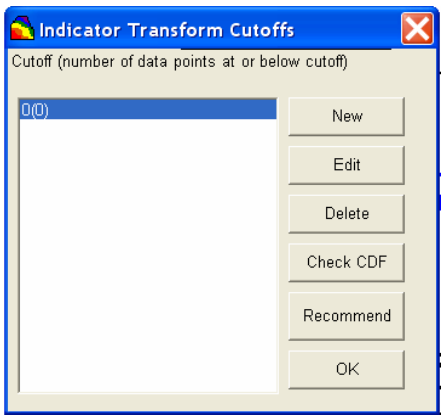
Correlation Modeling for Indicator Kriging/Indicator Simulation

When defining the point wise probability distributions (the distribution of possible concentration values at an unestimated node), ordinary kriging and sequentially gaussian simulation assume a normal distribution model. If your data is not normally distributed then you may consider using indicator kriging or sequentialy indicator simulation (keep in mind the normal score transform feature converts any distribution into a normal distribution permitting you to use sequentialy gaussian simulation-see above).

Rather than assume any distribution, the probability distribution or more accurately, the cumulative distribution function is generated numerically at each point. Recall that a cumulative distribution function specifies the probability that a response will be below any given threshold. Indicator methods accept a set of thresholds (1pCi/g, 3pCi/g, … etc) and determine the probability that the concentration is less than each of these thresholds individually. Therefore, you will need to specify the set of thresholds, and provide a correlation model for each.

The choice of thresholds may be motivated by different objectives. It is usually important to get a good representation of the range of possible values. This can be done by establishing indicator thresholds along certain percentile ranges (e.g. 10^{th}, 20^{th},….90^{th}). Another important threshold may be the decision criteria under consideration. If you have a decision criteria in mind, and don't specify it as a indicator threshold, then the code will be forced to estimate the CDF at your criteria by interpolating between the thresholds just above and just below. SADA can help with these.

If you have not already opened SpatialCorrelation.sda do so now. Select Soil and Ac-225 and select Model spatial correlation as the interview. Click on the step Correlation Modeling. In the Chooose Data parameter block, press the Edit button.

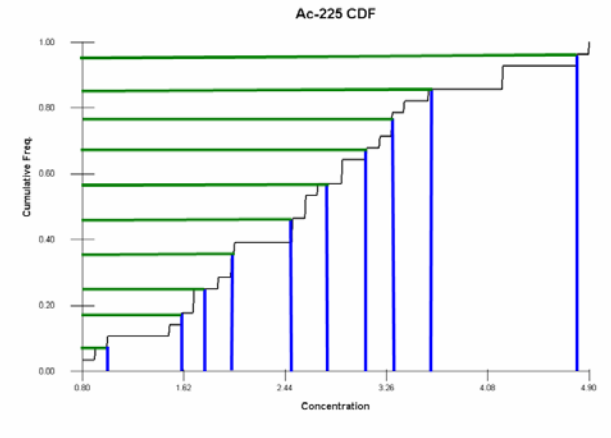


This is the indicator transform cutoff manager. Sometimes the work “cutoff” will be used but it is synonymous in this context with threshold. The first thing we’ll point out is the Recommend button. This button will analyze your data and produce ten indicator thresholds that roughly correspond to the 10th, 20th, etc percentiles of your data set. Press this button now and SADA will warn that any previous thresholds will be eliminated. Say Yes.

SADA selected the following thresholds.

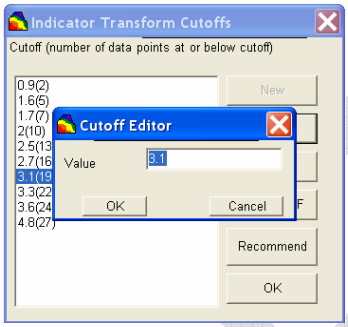
0.9(2)
1.6(5)
1.7(7)
2(10)
2.5(13)
2.7(16)
3.1(19)
3.3(22)
3.6(24)
4.8(27)

The first number is the indicator threshold. Thse second number in parenthesis is the number of data falling below this value. So you can see that this number increases with increasing threshold values. If you plot these results on a cumulative distribution function you will see a fairly evenly spacing among CDFvalues and irregular spacing among concentration values. This is normal as the goal is to have about 10% between each CDF value and concentration selections are adjusted to accommodate.



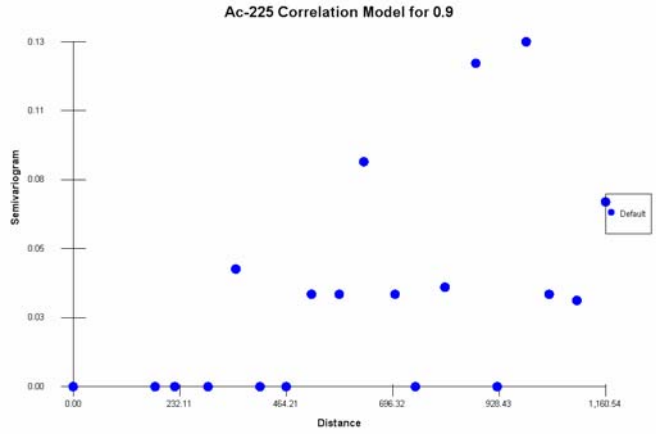
Suppose now, that we have a particular decision threshold in mind, for example, 3.0 pCi/g. You are currently only permitted 10 threshold values. This means that you will need to give up one of the 10 currently selected. There are a couple of ways to make this choice. First, you simply find a value that is very close and adjust it slightly. If the closes choice isn't so close, you may need to slightly a number of thresholds one direction or another. The imporant thing is to really callibrate the model right around your decision threshold. So lower threshold values may not be as important and one may be dropped.

Let's adjust the 3.1 cutoff to 3. In the list box, select the 3.1 cutoff and press the Edit button.



Enter 3.0 and press Ok.

You may notice when you being assessing correlation structures for each of these thresholds, that those thresholds in the very low and very upper end may produce erratice variography results or at least, variography values that are not well suited for some of the models. Consider the lower end of threshold values. When the indicator transforms are applied at the 0.9 threshold, the 2 out of 28 samples will be valued at 0. While 26 out of 28 samples will be valued at 1. This can mean that for many variography calculations all data will simply be one and the variography may be zero or near zero. When a 1 is encountered, the semi-variogram suddenly bounce higher. Consider the following semi-variogram calculation for 0.9 pCi/g.

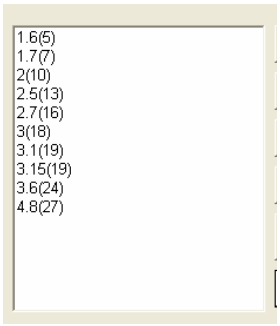


A number of the results fall on the zero value line. This is because during the semi-variogram calculations, those two samples that were less than 0.9pCi/g where not encountered. If you're decision criteria is a very low value, you may have difficulty using indicator kriging to model your contaminant and corresponding decision. You may need to consider a normal score transform and use ordinary kriging instead.

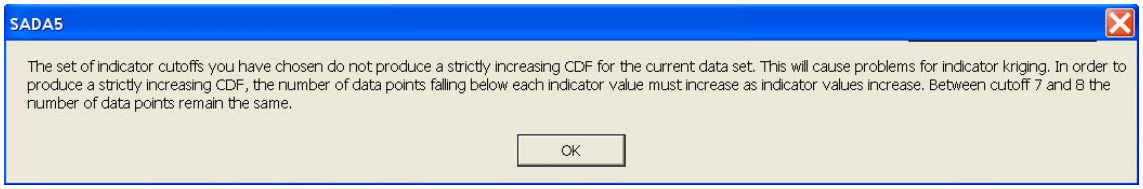
As a demonstration, let's eliminate the 0.9pCi/g threshold hold due to the semi-variogram problems and also because it is considerably lower than our decision rule. Select 0.9pCi/g and press the Delete button. Answer yes to the confirmation question.

We can add a new threshold by simply pressing the New button. Let's add another cutoff near the 3.0 pCi/g. Press New and enter 3.1 again. Press OK.

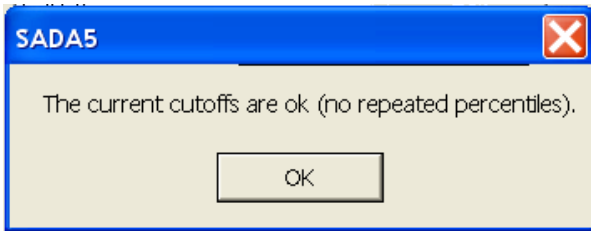
When performing indicator kriging or simulation your global CDF must be monotonic increasing. This simply means that every threshold value you select must have more data values below it than the previous threshold value did. So for example, the following selections violate this rule.



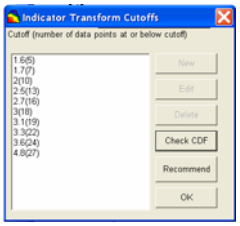
The problem here is that the threshold 3.1 and 3.15 both have 19 values that are less. SADA can quickly check this for you by pressing the Check CDF button. If you have a problem such as the one above, SADA will warn you.



In this case, you will need to manually adjust the thresholds so at each parenthetical number (the CDF) is non-repeating. If there are now problems, you get an OK message.



Make sure you have 9 cutoff threshold values that look like this and Press OK.



Back in the Parameter window, in the Choose Data parameter block, select the drop-list button next to Data Transforms. Notice that all your cutoff criteria are now loaded in this list. Now the only time you'll see this loaded in this list will be when you have the interview Model Correlation Selected or anytime you have Indicator kriging or Sequential Indicator Simulation selected. Otherwise, SADA will always restrict the data transforms to those that make sense in the current context. For example, if you had selected Ordinary Kriging under the Interpolate my data interview, all these cutoffs would no longer appear. It doesn't mean they are been lost. SADA is just hiding them in that context since they are not relevant to Ordinary kriging.

Select 1.6 and press the Recommend under the Explore Semi-variography block and the Recommend under the Modeling Semi-variography values block (choose spherical). Press Show Me in the lower right hand corner of the parameter window.

Repeat this process for 1.7, 2, 2.5 and so forth. What you are doing is establishing a correlation model for each of these indicator thresholds. This is required by indicator kriging and sequential indicator simulation alike. You will learn more about these interpolators/simulators in the next chapter. Save your file.

Correlation Modeling for Cokriging

Cokriging allows you to support your geospatial model with other data types that are correlated with the contaminant you are trying to model. For example, gamma count data might be a well support the estimation of Cs-137 in the soil. Depending on the flavor of cokriging you choose, you may need to estimate the correlation model for combinations of your contaminant and your auxilliary data. Suppose you are interested in modeling Contaminant "A: and wish to supplement with Field Detection Device "B". You may be required to estimate the spatial correlation between A and B. In this case, the "head" in the variogram discussion will be an A measurement and the "tail" will be a B measurement. This allows you to see how Contaminant A and Field Device B spatially vary as the separation distance increases. If you have also geology result C, the number of correlation models can get out of hand. In a fully specified cokriging modeling you would need the following correlation models: A/A, A/B, A/C., B/B, B/C, and C/C. There are some simplifying options and when you select cokriging as your interpolant, you will be asked to choose among some easier options. Suppose for now we only have Contaminant A and Field Device B.

Unrestricted coregionalization model (very hard to implement)

In this choice, you must model the correlation model for A, for B, and for A/B. Here you are free to model each variogram separately. However, there is no guarantee that the cokriging system has always a solution and a warning message is issued

Linear model of coregionalization (hard to implement)

Here both A, B, and A/B are modeled using the same correlation model, the same range, and the same anisotropy parameters. The sill is permitted to vary but a constraint called the Cauchy-Schwarz (CS) inequality must be met for each model.

$$\left|Sill_{1,2}\right| \leq \sqrt{Sill_1 x\ Sill_2}$$

Suppose we had the following models for Contaminant A with Contaminant B.

$\gamma_A(\mathbf{h})$ = 25 + 20 × Exp(range=2km)

$\gamma_B(\mathbf{h})$ = 9 + 29 × Exp(range=2km)

$\gamma_{A,B}(\mathbf{h})$ = 2+ 18 × Exp(range=2km)

We would be ok, because

$|2| \leq \sqrt{25x9}$ (nugget model) and $|18| \leq \sqrt{29x20}$ (exponential model) (Goovaerts, 1997).

Intrinsic Model of Coregionalization (easy to implement)

This model is more constraining than the linear model of coregionalization in that one more condition is imposed: the contribution of each basic model must be the same across all variogram models. For example, in the following intrinsic model the contributions of each basic model, while satisfying the CS inequalities, represent 50% contribution of the sill:

γ_A(h) = 20 + 20 ´ Exp(range=2km)

γ_B(h) = 9 + 9 ´ Exp(range=2km)

$\gamma_{A,B}$(h) = 5+ 5 ´ Exp(range=2km)

With $|5| \leq \sqrt{20x9}$

Another way to think about the intrinsic model is as a set of variogram models that are all proportional to the same function.

The model above can be expressed as a rescaling of the following model:

0.5 + 0.5 × Exp(range=2km).

The rescaling constant would be 40(γ_A(h)), 18(γ_B(h)), and 10($\gamma_{A,B}$(h)) (Goovaerts, 1997).

Markov Model Of Coregionalization (very easy to implement)

The Markov Model (MM) is the most straightforward model of coregionalization: only one direct variogram needs to be modeled and the other variograms are derived through a proportional relationship. Two Markov models are available: MM1 and MM2.

The Markov Model (MM) is the most straightforward model of coregionalization: only one direct variogram needs to be modeled and the other variograms are derived through a proportional relationship.

Markov Model (I) states that the cross variogram is proportional to the variogram of the primary variable:

$$\gamma_{12}(\mathbf{h}) = \sqrt{\frac{C_{22}(0)}{C_{11}(0)}}\rho_{12}(0)\gamma_{11}(\mathbf{h})$$

Where $C_{11}(0)$ and $C_{22}(0)$ is the variance of the primary and secondary variables respectively, while $\rho_{12}(0)$ is their correlation coefficient. This model is used in the framework of collocated cokriging which does not require knowledge of the variogram of the secondary variable since only one secondary datum is used for interpolation.

Markov-Bayes

This is a special application of the Markov model to a map of local prior probabilities. Rather than ordinary cokriging, an indicator cokriging approach is used, whereby hard data are first converted to 0s or 1s depending on whether they exceed a specified criterion. The cokriging method is then applied to these 0s and 1s along with the prior probability map. This results in an updated probability map that contains the influences of both the hard and soft data. In SADA, one must first create a prior probability map. (This is done by creating a user defined map and then choosing the interview Update My Probability Map.)

The rescaling for the Markov-Bayes model is slightly different than a traditional Markov approximation. Let Y represent the soft prior probability map data (see user defined maps(link)). Let I represent the sample data you wish to interpolate, where values have been transformed to zero if the measured value is less than or equal to the decision criteria, and 1 otherwise. In the correlation modeling step, you develop the correlation model, γ_I, for the indicator transformed data set. Then borrowing from Goovaerts (Geostatistics for Natural Resources, 1997) , we have that

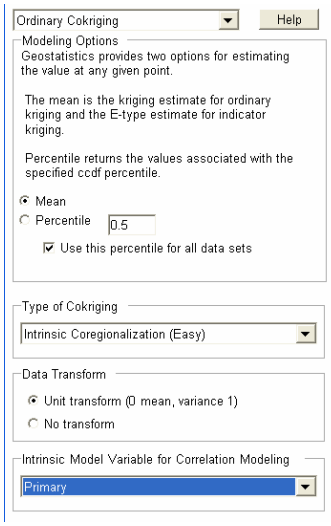
$$\gamma_Y(0) = |B|\gamma_I(0)$$
$$\gamma_Y(h) = B^2\gamma_I(h) \quad \forall h > 0$$

Where the coefficient B is defined as the difference between two conditional expectations:

$$m^1 = E[Y(u) \mid I(u) = 1] \in [0,1]$$
$$m^0 = E[Y(u) \mid I(u) = 0] \in [0,1]$$
$$B = m^1 - m^0$$

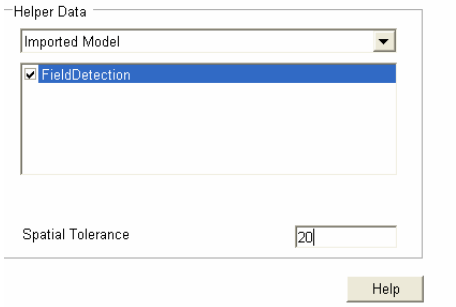
Try it Yourself

While much of this seems fairly intimidating, much of the effort has been worked out for you in the interface. For example, lets show how easy it is to do an intrinsic model of coregionalization. Later on you'll revisit this example but we'll introduce it now. Open up the file SpatialCorrelation.sda. Select Soil and Ac-225. In the Interview select Interpolate My Data. In the Step Interpolation methods select Cokriging and make sure that you have Intrinisic Coregionalization (Easy) selected. We'll worry about what the rest of this means later.



With intrinsic coregionalization, you only have to model one correlation structure. You can choose to model either the Ac-225 data or the Field Detection Results. Ac-225 is our primary because that is the contaminant we've selected.

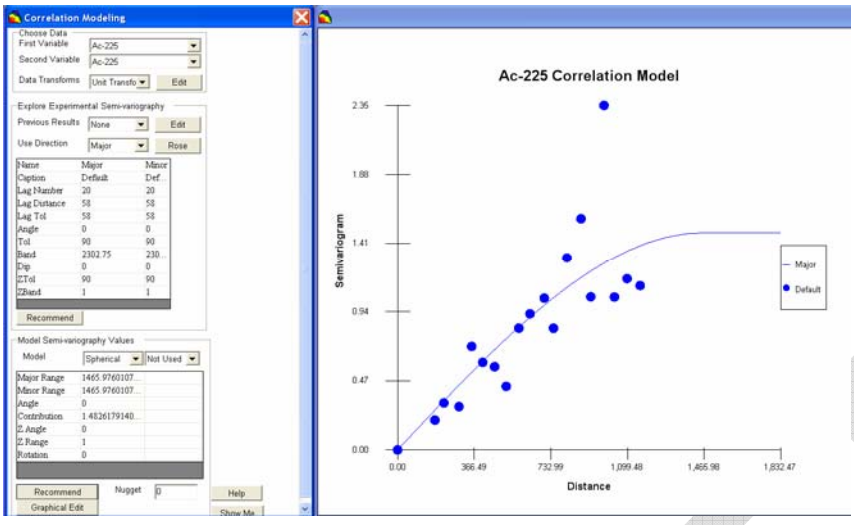
Click on the step "Choose helper data" and make sure the parameter window looks like this.



This is where we tell SADA what data set we want to "help" us better model our Ac-225 data. Make double sure the spatial tolerance is 20.

Click on the the Correlation Modeling Step. You can see that the only option is to perform correlation modeling on Ac-225. This is because we choose to only model our primary (Ac-225) variography. Notice that Unit transform is the only option under Data Transforms. This is also because of our previous choice in the interpolation methods step. Don't worry about this right now. It will be covered again later. The point here is that all you have to do is the same thing we did in the previous example and SADA will take care of callibrating the other correlation strucutres.

In the interest of practice, press both Recommend buttons on the correlation modeling form (variogram first) and model second (choose spherical). You should get a reasonable correlation model on the first try.



Keep in mind that in practice, you would need to look at anisotropy and the rose diagram before moving forward with actual interpolation.

Summary

This chapter presented the central concept of correlation structures, how to measure them and model them. This step is likely the most time consuming step you will encounter in the geospatial modeling processes. The next chapter introduces the different geospatial models and how correlation models developed in this chapter are used.

References

Deutsch, C. V. and A. G. Journel (1992). GSLIB: Geostatistical Software Library and User's Guide. New York, Oxford University Press.

Goovaerts, P. (1997). Geostatistics for Natural Resource Evaluation. New York, Oxford University Press.

