

Spatial Analysis and Decision Assistance (SADA) Version 4

Presented by
George E. Powers
Office of Nuclear Regulatory Research
Nuclear Regulatory Commission

Environmental Assessment Methods in SADA
Denver, CO
May 25, 2005

SADA General Information

Windows--based freeware designed to integrate scientific models with decision and cost analysis frameworks in a seamless, easy to use environment.

- Visualization/GIS
- Statistical Analysis
- Geospatial Interpolation
- Geospatial Uncertainty Analysis
- Human Health Risk Assessment
- Ecological Risk Assessment
- Custom Analysis
- MARSSIM Module
- Area of Concern Frameworks
- Cost Benefit Analysis
- Sampling Designs
- Export to Arcview/Earthvision

SADA has been supported by both the DOE, EPA, and the NRC. SADA Version 3.0 had about 11000 downloads. Version 4.0 has had 800+ since December, 2004.

SADATM

Spatial Analysis and Decision Assistance

SADA General Information (cont.)

Free stand-alone package for Windows 98, 98SE, NT SP4 or higher, 2000, ME, and XP.

Contact information, updates, documentation, and downloads are available online at <http://www.tiem.utk.edu/~sada/>

A SADA user's group, email, annual conferences, and on-site training.

A substantial help file is included.

Conduct “black and white box” testing internally as well as an external beta release period.

Publish verification document on the website.

SADATM

Spatial Analysis and Decision Assistance

Data Formats

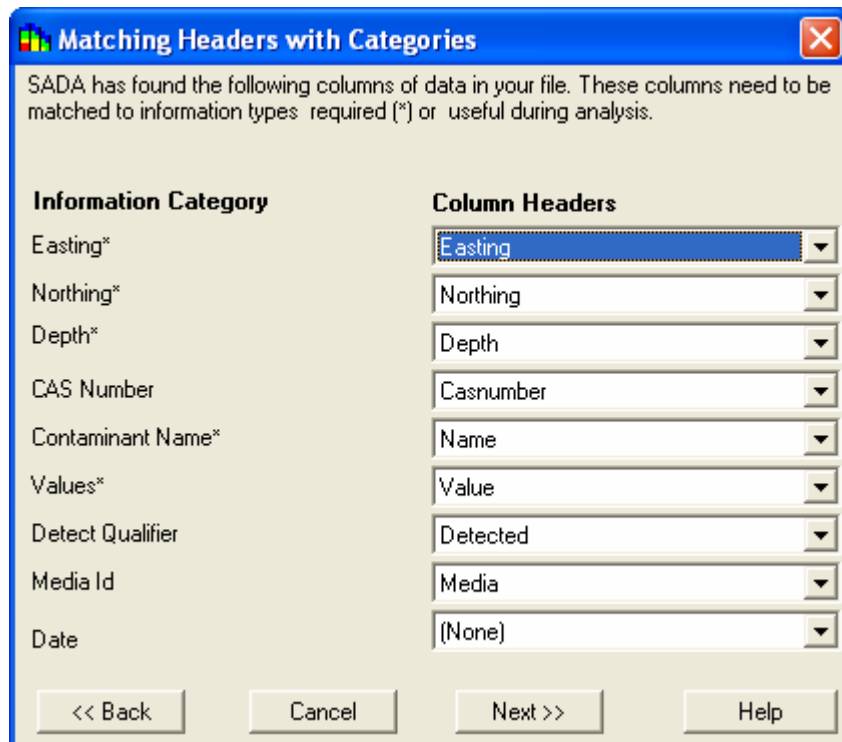
- SADA can accept data in two formats: comma delimited files (csv) and Microsoft Access.
- SADA requires the presence of certain fields in the data set.
 - Easting
 - Northing
 - Depth
 - Value
 - Name
- SADA can use other forms of information as well
 - Media
 - Detection
 - Date
 - CAS Number
- Any other form of meta data can be imported as well. User can plot and retrieve this meta data during an analysis.
- SADA recognizes soil, sediment, surfacewater, groundwater, air, biota, and background, and the “basic” media type. Basic is assigned to data that have no media type.



Importing Data Into SADA

The next step is to match the columns of information in the ascii data file to information categories that are required or may be useful in SADA.

Required information categories are followed by an (*) and must be assigned to a column in the ascii data file. A category is not assigned if the (none) option is selected in the drop down box. The Depth category is required only when data exist at varying depths. If the Detect Qualifier is not assigned, the data are assumed to be all detects.



Matching Headers with Categories

SADA has found the following columns of data in your file. These columns need to be matched to information types: required (*) or useful during analysis.

Information Category	Column Headers
Easting*	Easting
Northing*	Northing
Depth*	Depth
CAS Number	Casnumber
Contaminant Name*	Name
Values*	Value
Detect Qualifier	Detected
Media Id	Media
Date	(None)

<< Back Cancel Next >> Help

If Media ID, which denotes the type of media the contaminants are sampled in (e.g. soil or groundwater) is not assigned, SADA adds an artificial media column titled 'Basic' and the human health risk and/or ecological risk modules cannot be setup later.

After the columns have been set, press **Next>>**. SADA begins the conversion process and presents the data as it will be imported into the Data Editor.

SADATM

Spatial Analysis and Decision Assistance

Data Editor

The Data Editor provides a chance to identify errors in the data set and correct them during the import process. It may also be accessed from the **Tools Menu** at any time later for data corrections or additions.

SADA highlights cells with red if they contain an unacceptable value. To determine the exact error, place the mouse over the red cell and the yellow text box near the top explains the problem. Once there are no red cells, the process continues.

ID	Name	Casnumber	Easting	Northing	Depth	Value	Detected	Media
1	Ac-225	14265851	27596.25	21900	0	1.99657	1	SD
2	Ac-225	14265851	28310.25	21900	0	1.63026	1	SW
3	Ac-225	14265851	28935	21900	0	0.86914	1	SW
4	Ac-225	14265851	27685.5	22200	0	2.053298	1	SW
5	Ac-225	14265851	28131.75	22200	0	4.185278	1	SD
6	Ac-225	14265851	29202.75	22500	0	1.49788	1	SD
7	Ac-225	14265851	27150	23160	0	1.70351	1	SD
8	Ac-225	14265851	27685.5	22920	0	2.306226	1	SD
9	Ac-225	14265851	28042.5	23100	0	4.965262	1	GW
10	Ac-225	14265851	28221	23100	0	4.232573	1	GW
11	Ac-225	14265851	28667.25	23220	0	2.951485	1	GW
12	Ac-225	14265851	29113.5	22980	0	0.99677	1	SD
13	Ac-225	14265851	27417.75	23580	0	1.92566	1	SD
14	Ac-225	14265851	27774.75	23640	0	2.939644	1	SD
15	Ac-225	14265851	28310.25	23400	0	3.121663	1	SD
16	Ac-225	14265851	28935	23460	0	0.83135	1	SD
17	Ac-225	14265851	28200	22560	0	4.8	1	SD
18	Ac-225	14265851	NA	22500	0	3.3	1	SD
19	Ac-225	14265851	27200	22380	0	2.03	1	SD
20	Barium	7440393	27596.25	21900	0	42.77435607	1	SD
21	Barium	7440393	28310.25	21900	0	35.12087239	1	GW
22	Barium	7440393	28935	21900	0	18.48691343	1	SD
23	Barium	7440393	27685.5	22200	0	43.55465395	1	SD

It is recommended that the **Automatic Error Checking** box remain checked so SADA looks for mistakes as you type. When the user is entering or pasting large amounts of data and does not wish the process to be slowed, however, it may be preferable to uncheck the **Automatic Error Checking** box and check errors later with the **Check Errors** button.

SADATM

Spatial Analysis and Decision Assistance

The New SADA Look: Scalable Interfacing

The image shows the SADA software interface with several components labeled by arrows:

- Analysis Box**: Points to the top menu bar (File, Graphics, Setup, Reports, Statistics, Tools, Help).
- Data Type Box**: Points to the 'Soil' dropdown menu.
- Data Name Box**: Points to the 'Anthracene' dropdown menu.
- Labels Box**: Points to the '(None)' dropdown menu.
- Layers Box**: Points to the 'Z = 0' dropdown menu.
- Interviews**: Points to the 'Plot my data' button.
- Steps Window**: Points to the 'Steps' panel on the left, which contains a list of steps (1. Choose your data, 2. Set vertical layers, 3. Set GIS layers, 4. Set polygons, 5. Show the results, 6. Import sampled data, 7. Format picture, 8. Auto-document, 9. Add to results gallery) and navigation buttons (<< Back, Next >>).
- Parameters Window**: Points to the 'Data Query' panel, which includes options for 'Data Query' (All, Interval), 'Duplicate Data' (Use all values, Use only detected values, Use most recent value, Use most recent detected value), and 'Non Detects' (Use zero, Use half the detection limit, Use the full detection limit).
- Results Window**: Points to the 'Anthracene Sample Locations (Z = 0)' plot, which is a scatter plot showing sample locations with a color scale from 0.80 to 5.50.

The main window displays the 'Anthracene Sample Locations (Z = 0)' plot, which is a scatter plot showing sample locations with a color scale from 0.80 to 5.50. The plot is titled 'Anthracene Sample Locations' and shows a distribution of points across a grid of Easting (X-axis) and Northing (Y-axis) coordinates. The color scale on the right indicates the concentration of Anthracene, ranging from 0.80 (blue) to 5.50 (red).

The New SADA Look

The image shows the SADA software interface with several components labeled by arrows:

- Analysis Box**: Points to the top menu bar (File, Graphics, Setup, Reports, Statistics, Tools, Help).
- Data Type Box**: Points to the 'Soil' dropdown menu.
- Data Name Box**: Points to the 'Anthracene' dropdown menu.
- Labels Box**: Points to the '(None)' dropdown menu.
- Layers Box**: Points to the 'Z = 0' dropdown menu.
- Interviews**: Points to the 'Plot my data' button.
- Steps Window**: Points to the 'Steps' panel on the left, which lists a sequence of tasks from 'Choose your data' to 'Add to results gallery'.
- Parameters Window**: Points to the 'Data Query' panel, which includes options for 'Duplicate' and 'Non Detects'.
- Results Window**: Points to the 'Anthracene Sample Locations (Z = 0)' plot, which is a scatter plot of sample locations with a color-coded legend on the right.

The 'Data Query' panel shows a list of parameters: 'Ecological', 'General', 'Human Health', and 'NRC Screening Values'. The 'General' parameter is currently selected.

The 'Results Window' displays a scatter plot of sample locations. The x-axis is labeled 'Easting' and ranges from 27,026.00 to 29,202.7. The y-axis is labeled 'Northing' and ranges from 21,900.00 to 22,944.00. A color scale on the right indicates values from 0.80 to 5.50.

The New SADA Look

The image shows the SADA software interface with several components labeled by arrows:

- Analysis Box**: Points to the top menu bar (File, Graphics, Setup, Reports, Statistics, Tools, Help).
- Data Type Box**: Points to the 'Soil' dropdown in the 'General' tab.
- Data Name Box**: Points to the 'Anthracene' dropdown.
- Labels Box**: Points to the '(None)' dropdown.
- Layers Box**: Points to the 'Z = 0' dropdown.
- Interviews**: Points to the 'Plot my data' button.
- Steps Window**: A vertical panel on the left with a list of steps (1. Choose your data, 2. Set vertical layers, 3. Set GIS layers, 4. Set polygons, 5. Show the results, 6. Import sampled data, 7. Format picture, 8. Auto-document, 9. Add to results gallery) and navigation buttons (<< Back, Next >>).
- Parameters Window**: A central panel titled 'Data Query' with options for 'Data Query' (All, Interval), 'Duplicate Data' (Use all values, Use only detected values, Use most recent value, Use most recent detected value), and 'Non Detects' (Use zero, Use half the detection limit, Use the full detection limit).
- Results Window**: A scatter plot titled 'Anthracene Sample Locations (Z = 0)' showing sample locations with a color scale from 0.80 to 5.50. The x-axis is labeled 'Easting' and the y-axis is labeled 'Northing'.

A green arrow points from the 'Data Type Box' to a pop-up window titled 'Soil' which contains the text: 'Soil', 'Results Gallery', and 'Background'.

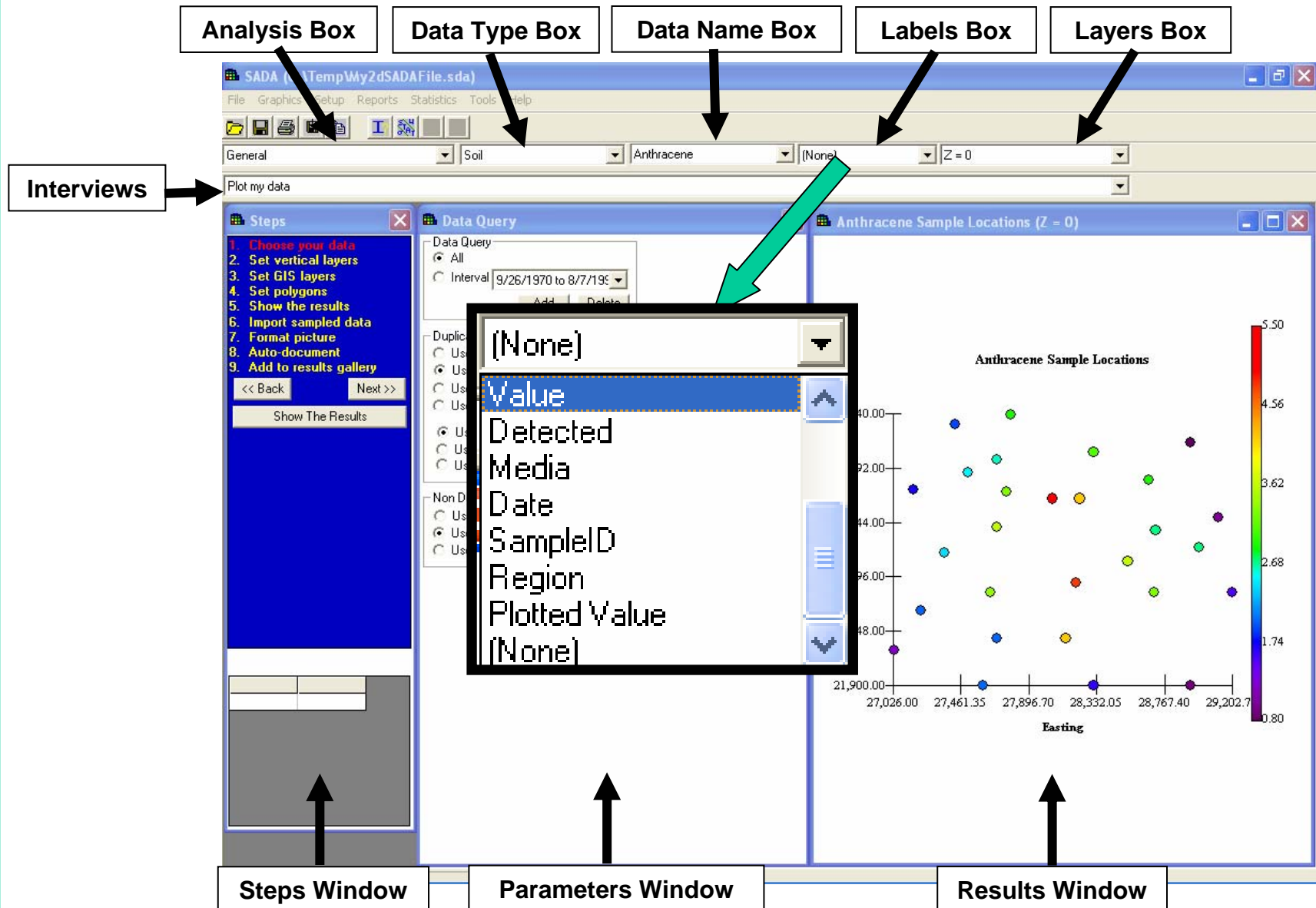
The New SADA Look

The screenshot displays the SADA software interface with the following components labeled:

- Analysis Box**: Points to the top toolbar area.
- Data Type Box**: Points to the 'Soil' dropdown menu.
- Data Name Box**: Points to the 'Anthracene' dropdown menu.
- Labels Box**: Points to the '(None)' dropdown menu.
- Layers Box**: Points to the 'Z = 0' dropdown menu.
- Interviews**: Points to the 'Plot my data' button.
- Steps Window**: Points to the 'Steps' panel on the left, which lists a workflow from 'Choose your data' to 'Add to results gallery'.
- Parameters Window**: Points to the 'Data Query' panel, which includes options for 'Data Query' (All, Interval), 'Duplicate Data' (Use all values, Use only detected values, etc.), and 'Non Detects' (Use zero, Use half the detection limit, etc.).
- Results Window**: Points to the 'Anthracene Sample Locations (Z = 0)' plot, which shows a scatter plot of sample locations with a color-coded legend for 'Ac-225', 'Anthracene', 'Arsenic', 'Barium', and 'Pooled Data'.

The main window title is 'SADA (TempMy2dSADAFfile.sda)'. The menu bar includes 'File', 'Graphics', 'Setup', 'Reports', 'Statistics', 'Tools', and 'Help'. The toolbar contains icons for file operations, plotting, and data management.

The New SADA Look

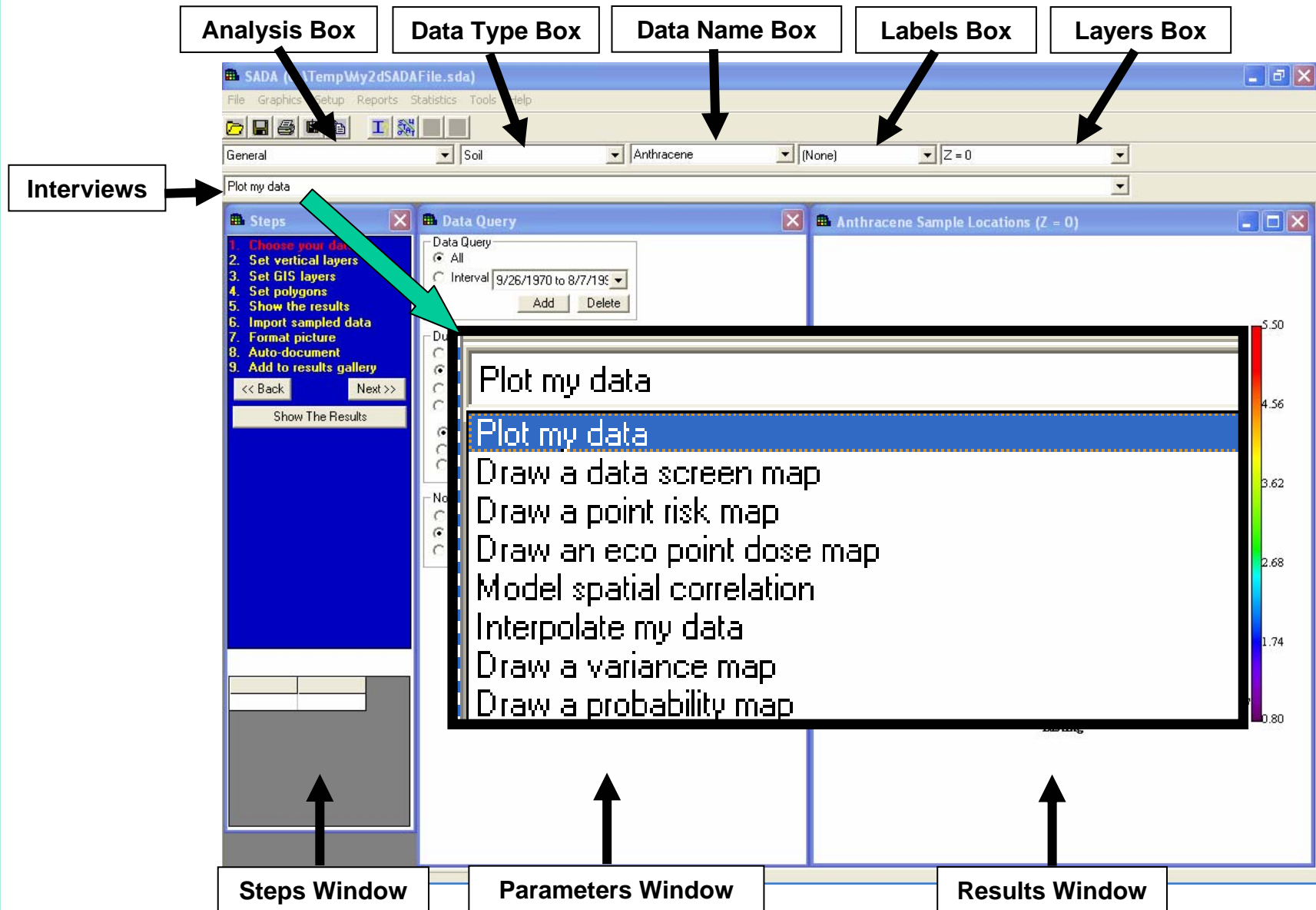


The New SADA Look

The screenshot displays the SADA software interface with the following components labeled:

- Analysis Box**: Points to the top menu bar (File, Graphics, Setup, Reports, Statistics, Tools, Help).
- Data Type Box**: Points to the 'Soil' dropdown menu.
- Data Name Box**: Points to the 'Anthracene' dropdown menu.
- Labels Box**: Points to the '(None)' dropdown menu.
- Layers Box**: Points to the 'Z = 0' dropdown menu.
- Interviews**: Points to the 'Plot my data' button.
- Steps Window**: Points to the 'Steps' panel on the left, which lists a sequence of tasks from 'Choose your data' to 'Add to results'.
- Parameters Window**: Points to the 'Data Query' panel, which includes a 'Data Query' section with 'All' selected and a date interval '9/26/1970 to 8/7/199'. Below this is a dropdown menu showing depth ranges: $0 \leq Z < 10$, $10 \leq Z < 20$, $20 \leq Z < 30$, $30 \leq Z < 40$, and $40 \leq Z < 50$.
- Results Window**: Points to the 'Anthracene Sample Locations (Z = 0)' plot, which is a scatter plot of sample locations. The x-axis is labeled 'Easting' and the y-axis is labeled 'Northing'. A color scale on the right indicates values ranging from 0.80 to 5.50.

The New SADA Look

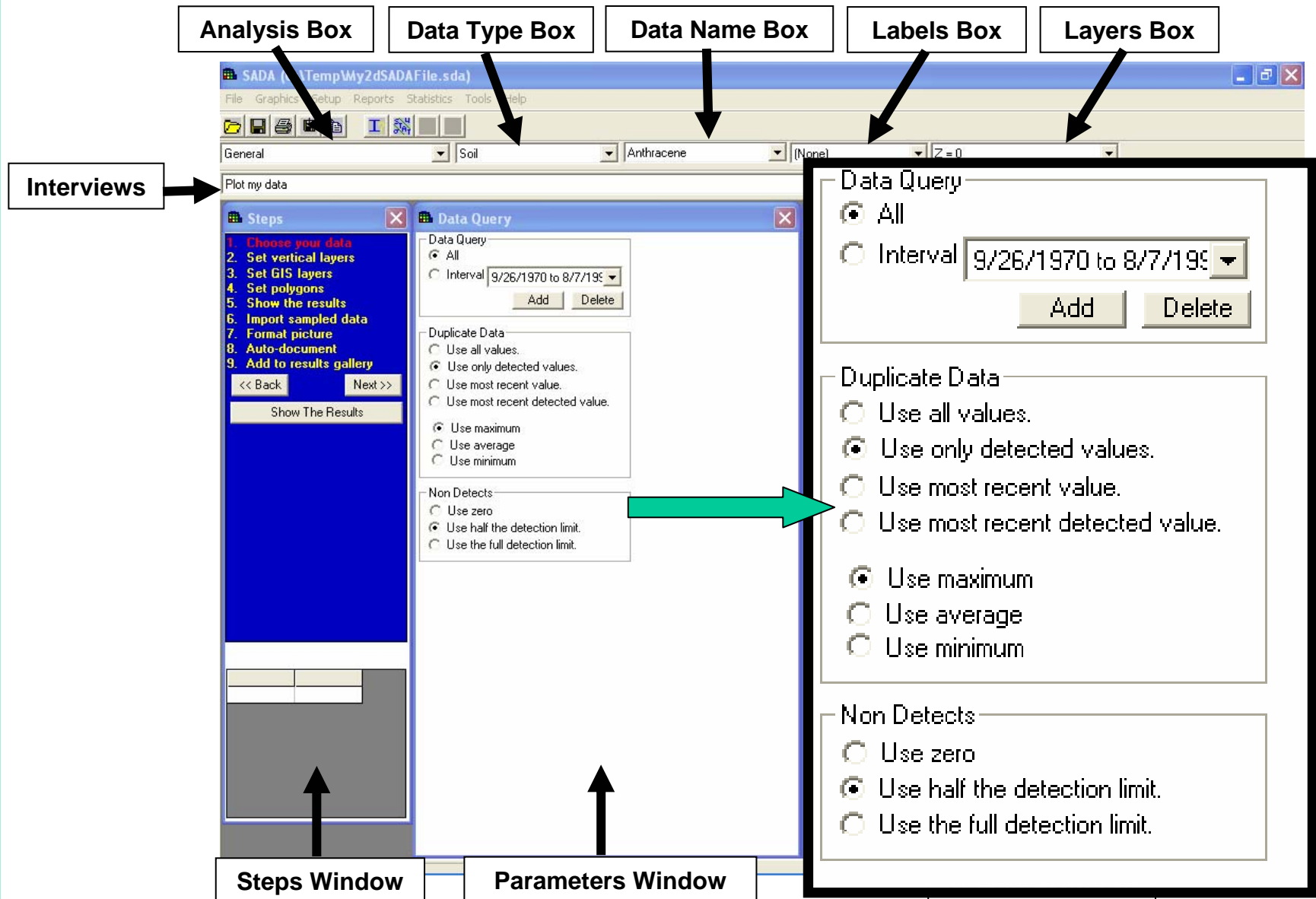


The New SADA Look

The image shows the SADA software interface with several components highlighted by callouts:

- Analysis Box**: Points to the 'General' dropdown menu.
- Data Type Box**: Points to the 'Soil' dropdown menu.
- Data Name Box**: Points to the 'Anthracene' dropdown menu.
- Labels Box**: Points to the '(None)' dropdown menu.
- Layers Box**: Points to the 'Z = 0' dropdown menu.
- Interviews**: Points to the 'Plot my data' button.
- Steps Window**: A window on the left containing a list of steps: 1. Choose your data, 2. Set vertical layers, 3. Set GIS layers, 4. Set polygons, 5. Show the results, 6. Import sampled data, 7. Format picture, 8. Auto-document, 9. Add to results gallery. It includes '<< Back' and 'Next >>' buttons and a 'Show The Results' button.
- Parameters Window**: A central window titled 'Steps' containing the same list of steps as the Steps Window, but with a 'Show The Results' button at the bottom.
- Results Window**: A window on the right showing a scatter plot titled 'Sample Locations'. The plot has a color scale on the right ranging from 0.80 to 5.50. The x-axis is labeled 'Easting' with values 28,332.05, 28,767.40, and 29,202.7. The y-axis is labeled 'Northing' with values 16,700.00, 16,700.00, and 16,700.00.

The New SADA Look



The New SADA Look

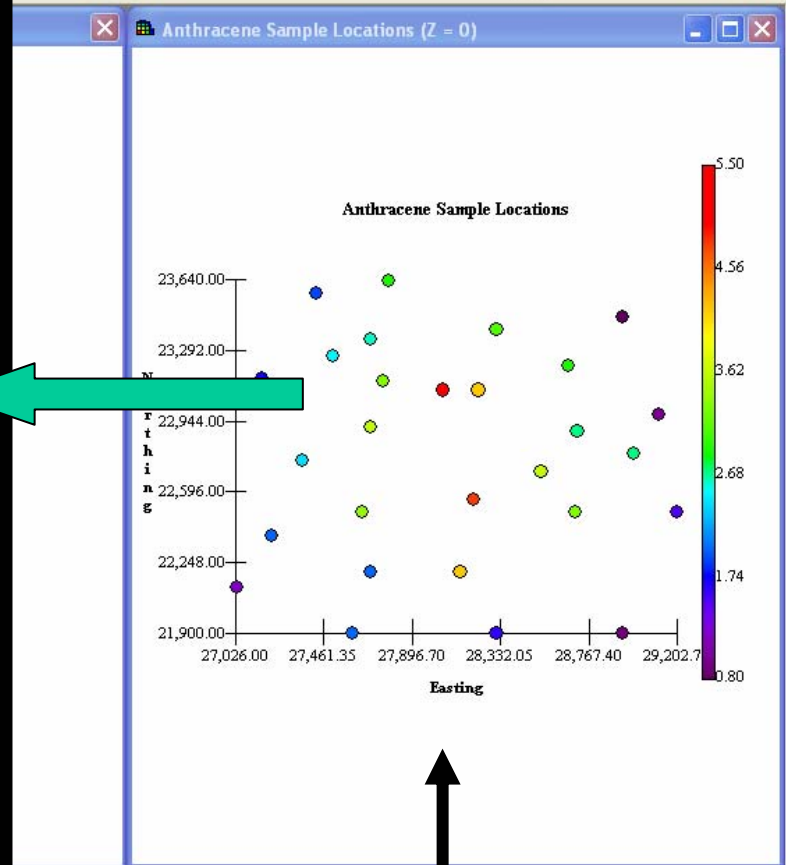
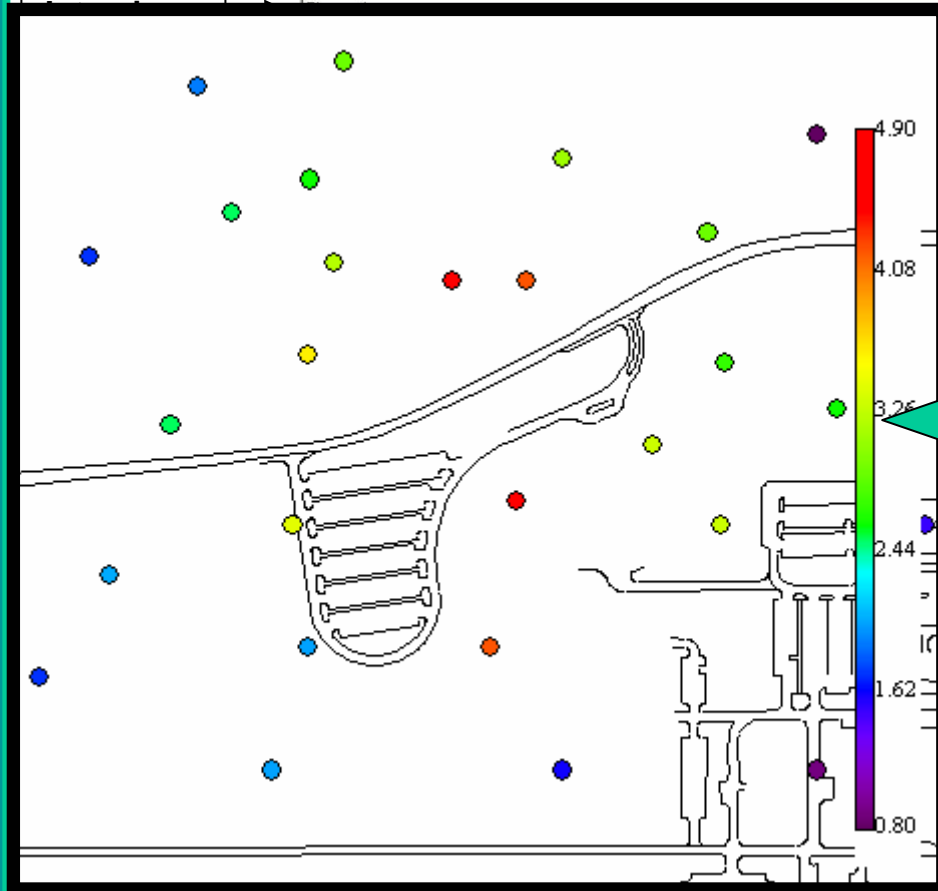
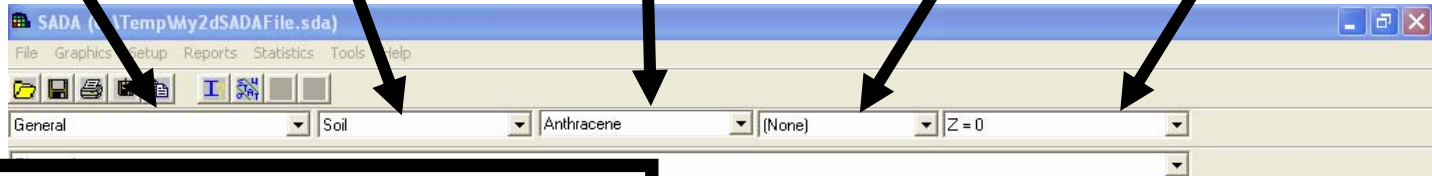
Analysis Box

Data Type Box

Data Name Box

Labels Box

Layers Box



Steps Window

Parameters Window

Results Window

What exactly can you do in SADA?

Create initial sample designs

Import data

Plot data

Import GIS layers

Aggregate sections of the site

Calculate statistics (univariate, bivariate)

Model spatial correlation

Create contour maps

Create a kriging variance map

Perform traditional HH and Eco risk assessments
(tabular risk, screens, prgs, benchmarks)

Create a HH or Eco contoured risk map

Create a HH or Eco point risk map

Create a data screen map for HH, Eco, Custom

Create an eco point dose map

Create an contoured eco dose map

Create probability maps

Define areas of concern

Calculate cost vs cleanup

Draw a LISA Map

Develop secondary sample designs

Perform a MARSSIM data analysis

Detect and Define MARSSIM elevated area

Visualize results in 3d

Autodocument results

Create a geobayesian site conceptual model

Draw area of concern maps based on conceptual model

Calculate cost vs cleanup based on conceptual model

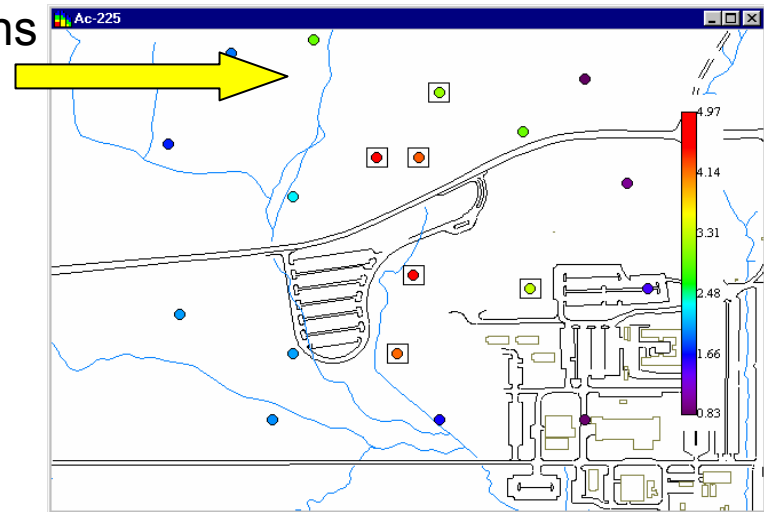
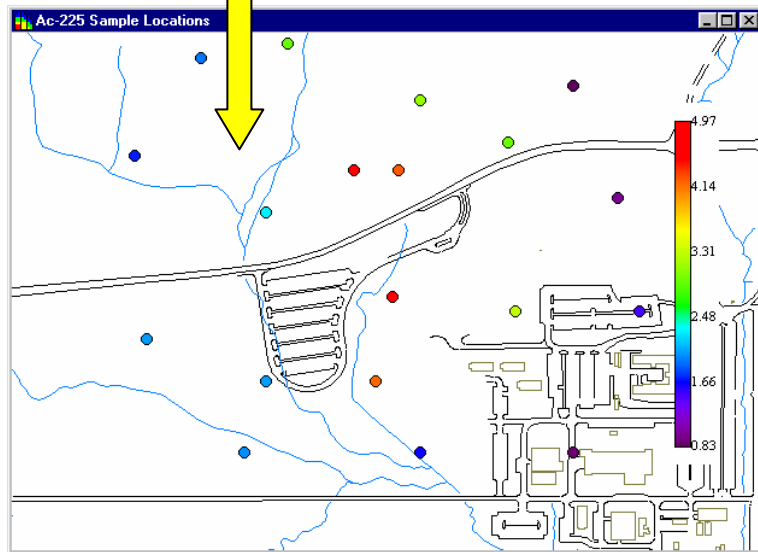
Update the site conceptual model

Export to ESRI or Earthvision or common window applications

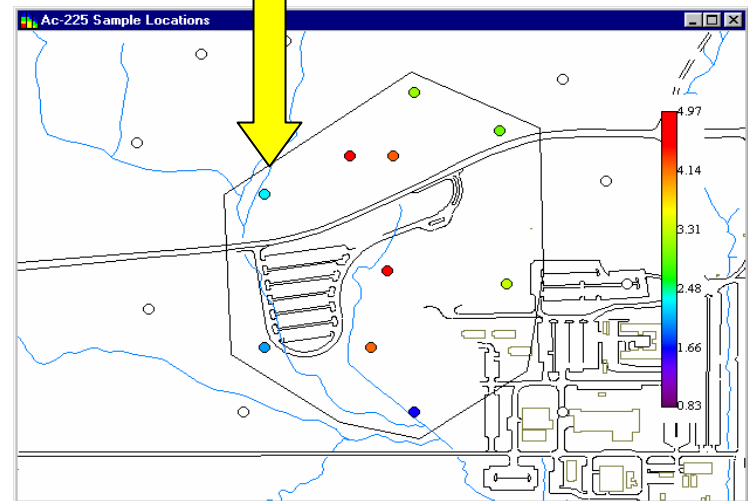
Data Exploration

Spatial Data Screens

Data Plot/GIS Overlays



Polygon Selection/Cutaways



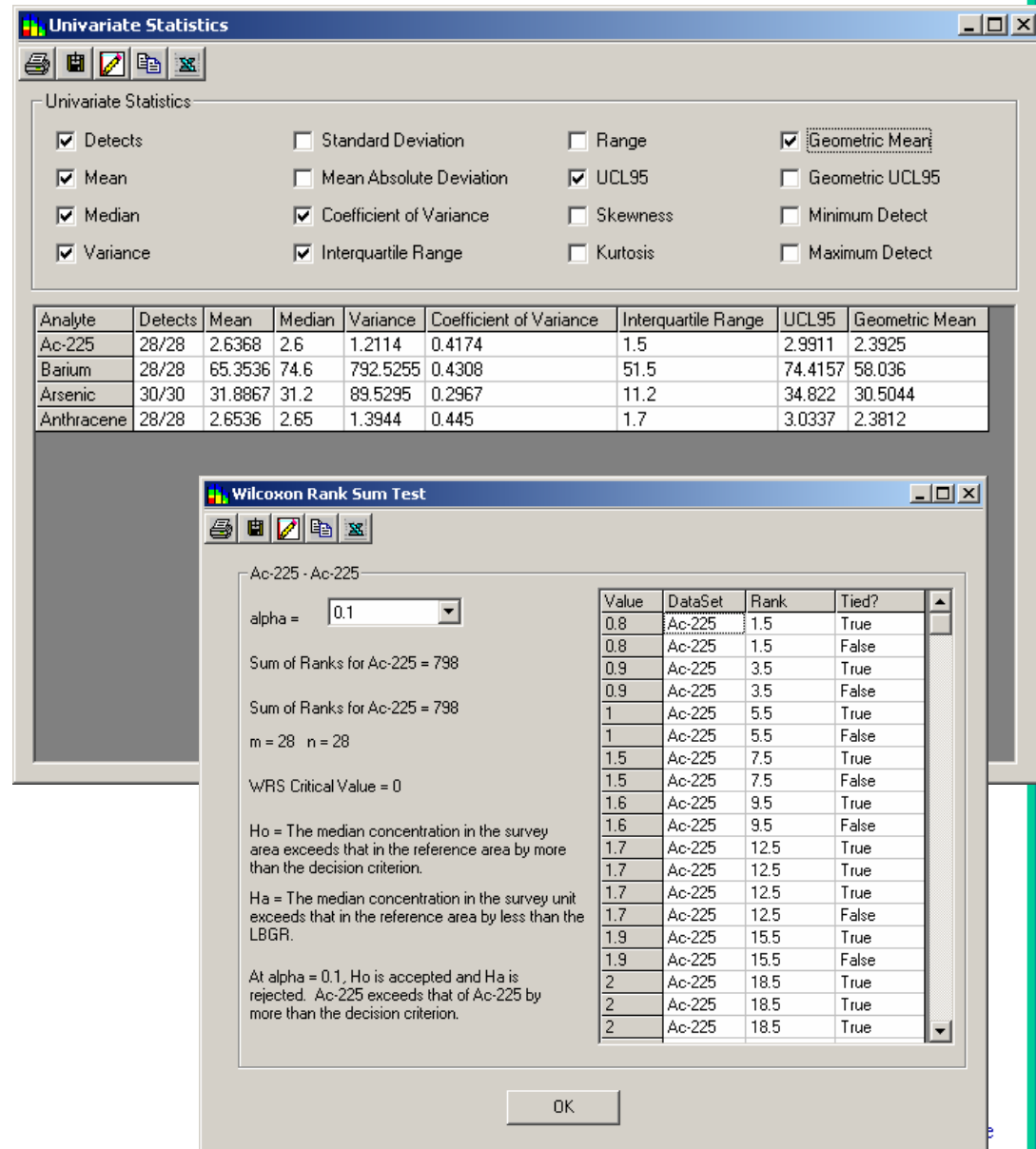
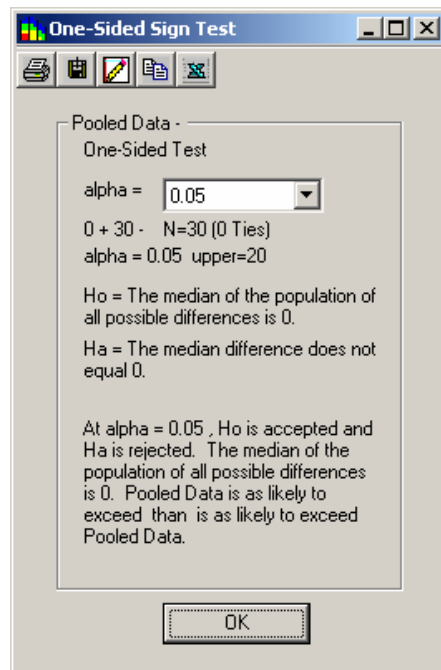
Statistics

A screenshot of a 'Statistics' window. It has a menu bar with 'Options' and 'Format'. Below the menu bar are three icons: a printer, a clipboard, and a pencil. A yellow arrow points to the table below.

Name	CAS Number	Mean	Variance	Number of Data
Ac-225	14265851	3.4	1.3	10
Beryllium and compo	7440417	75.6	640.9	10
Arsenic, Inorganic	7440382	8.3	6.	10

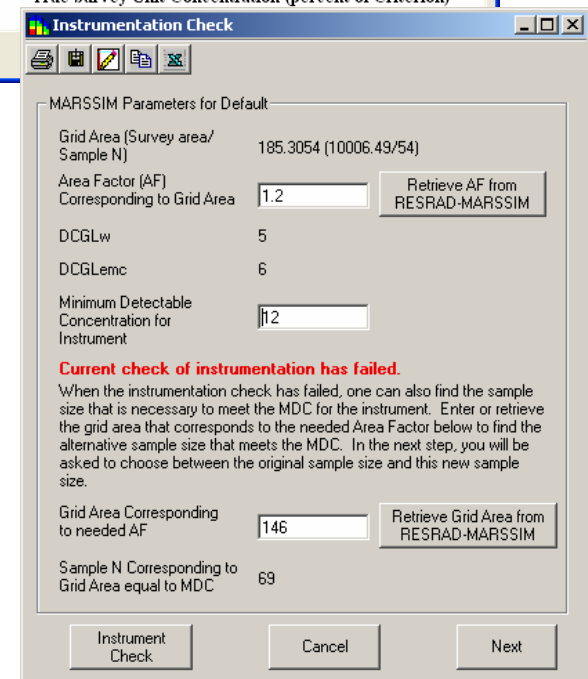
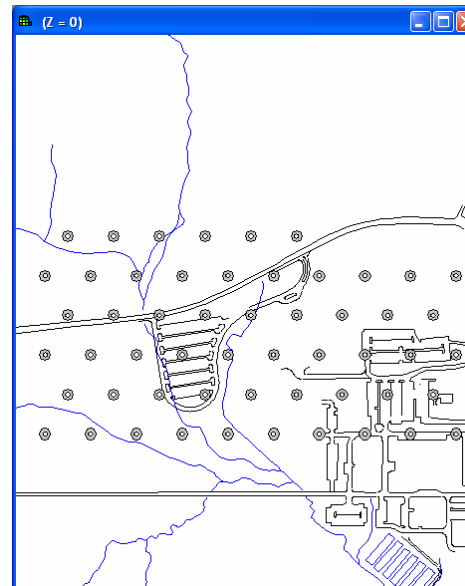
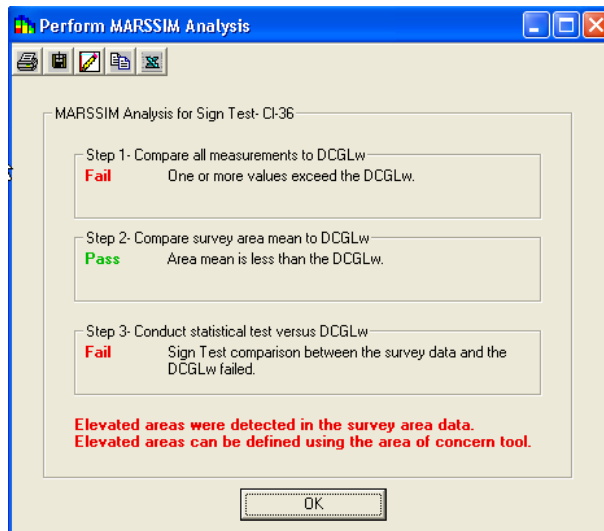
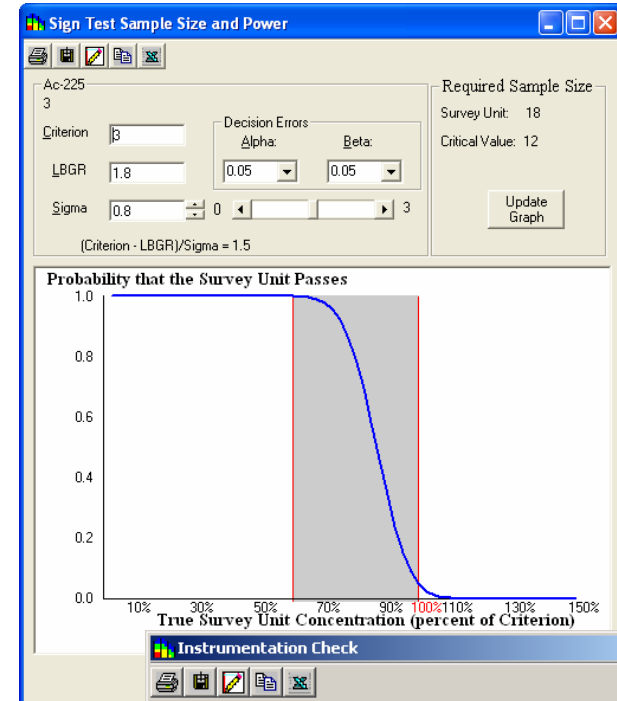
Statistics

- Additional univariate statistics
- Non-parametric hypothesis testing

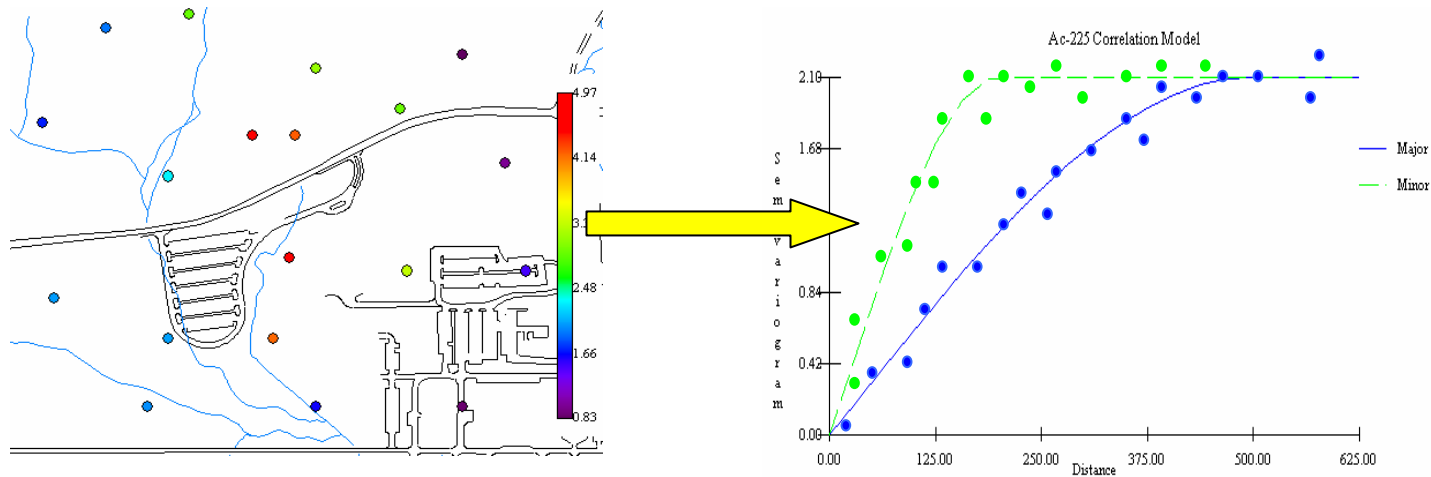


MARSSIM Functionality

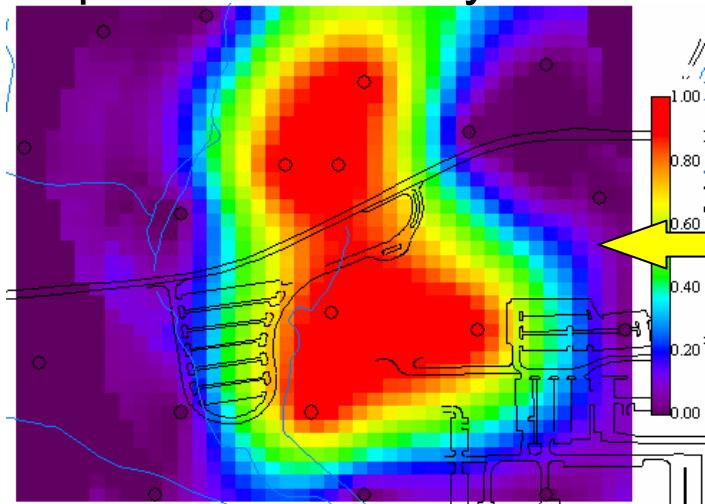
- Calculate sample size based on Sign Test and WRS Test
- Develop initial sample design incorporating DCGLS, Area Factors, Instrument sensitivity
- Post sampling analysis (A site passes or fails)
- Detecting and Defining Elevated Areas



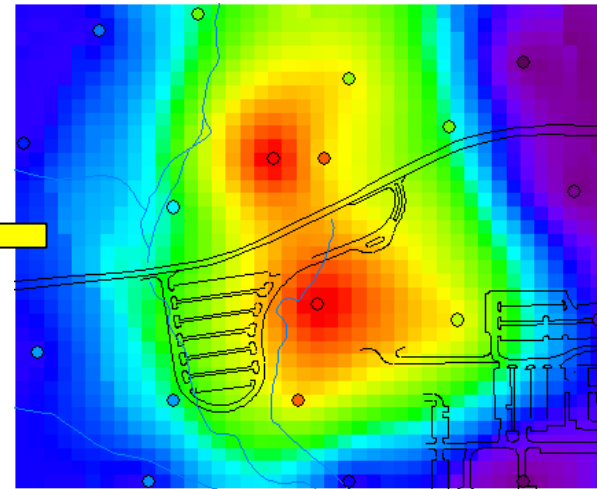
Spatial Analysis



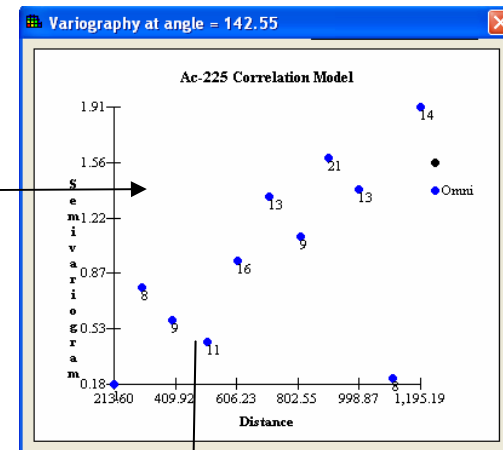
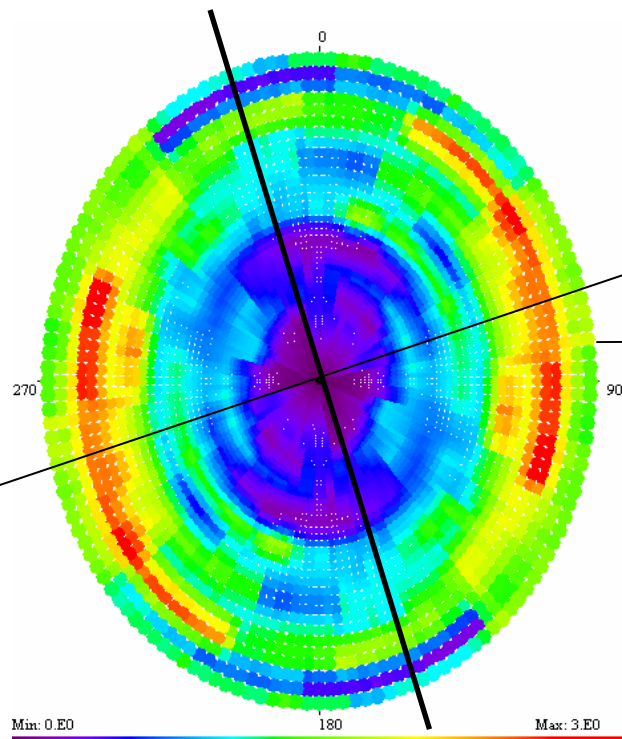
Spatial Uncertainty



Spatial Estimation



Version 4.0 Correlation Modeling Tools



Correlation Autofit

Choose the number of correlation structures to use.

☒ Single Structure ☐ Both Structures

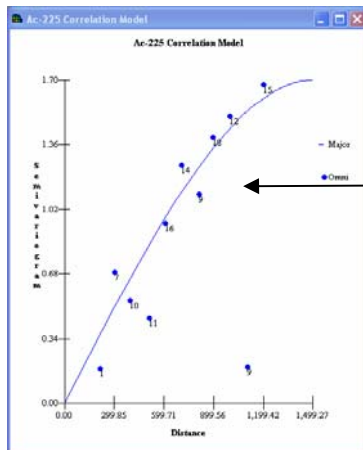
Select the models to be considered during Autofit

☒ Spherical

☒ Exponential

☐ Gaussian

OK Cancel



SADATM

Spatial Analysis and Decision Assistance

Geostatistics

SADA provides two kriging (geostatistical) models: Ordinary and Indicator kriging. Ordinary kriging assumes a normal or lognormal distribution for the data. Indicator kriging is a non parametric approach that does not assume any distribution.

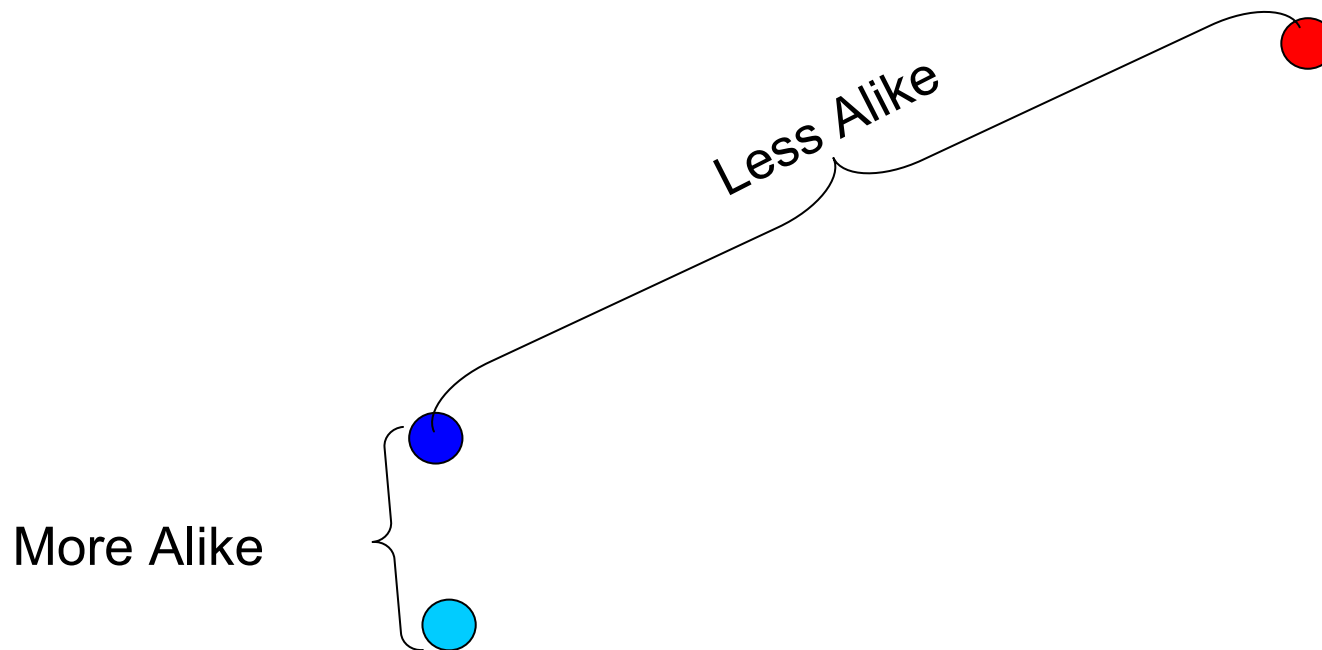
Like the methods discussed in Basic Spatial Analysis Tools, both methods are based on a weighted combination of nearby samples. However, the development and expression of these weights is quite complex and beyond the scope of this training guide.

It may be helpful 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 $C(d)$.

Spatial Correlation

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).

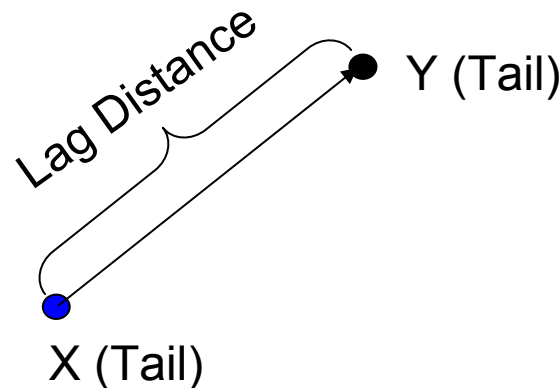


Spatial Correlation

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*.

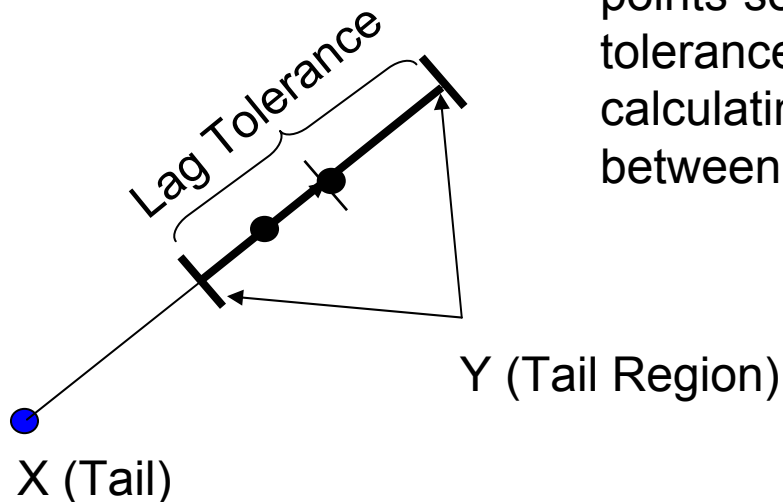
$$\gamma(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} (x_i - y_i)^2$$

where $N(\mathbf{h})$ 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).



Spatial Correlation

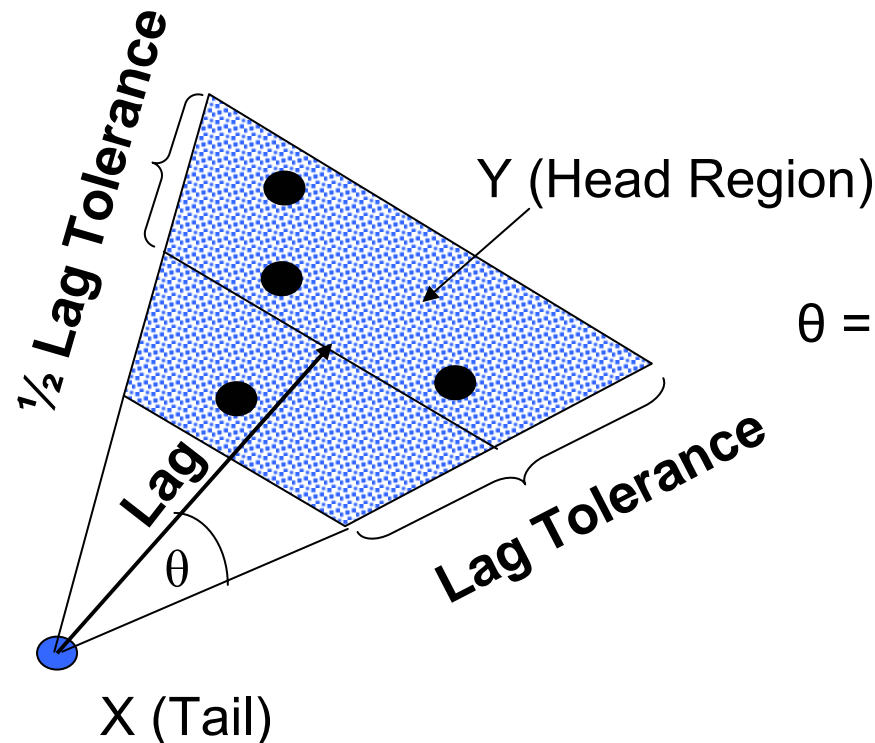
Rarely in practice, will you ever have any sample points separated by exactly a lag distance h . Therefore, a lag tolerance centered about the lag distance will permit a capture of more data points in the calculation of $\gamma(h)$. In the figure below, all data points within the blue shaded area will 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.

Spatial Correlation

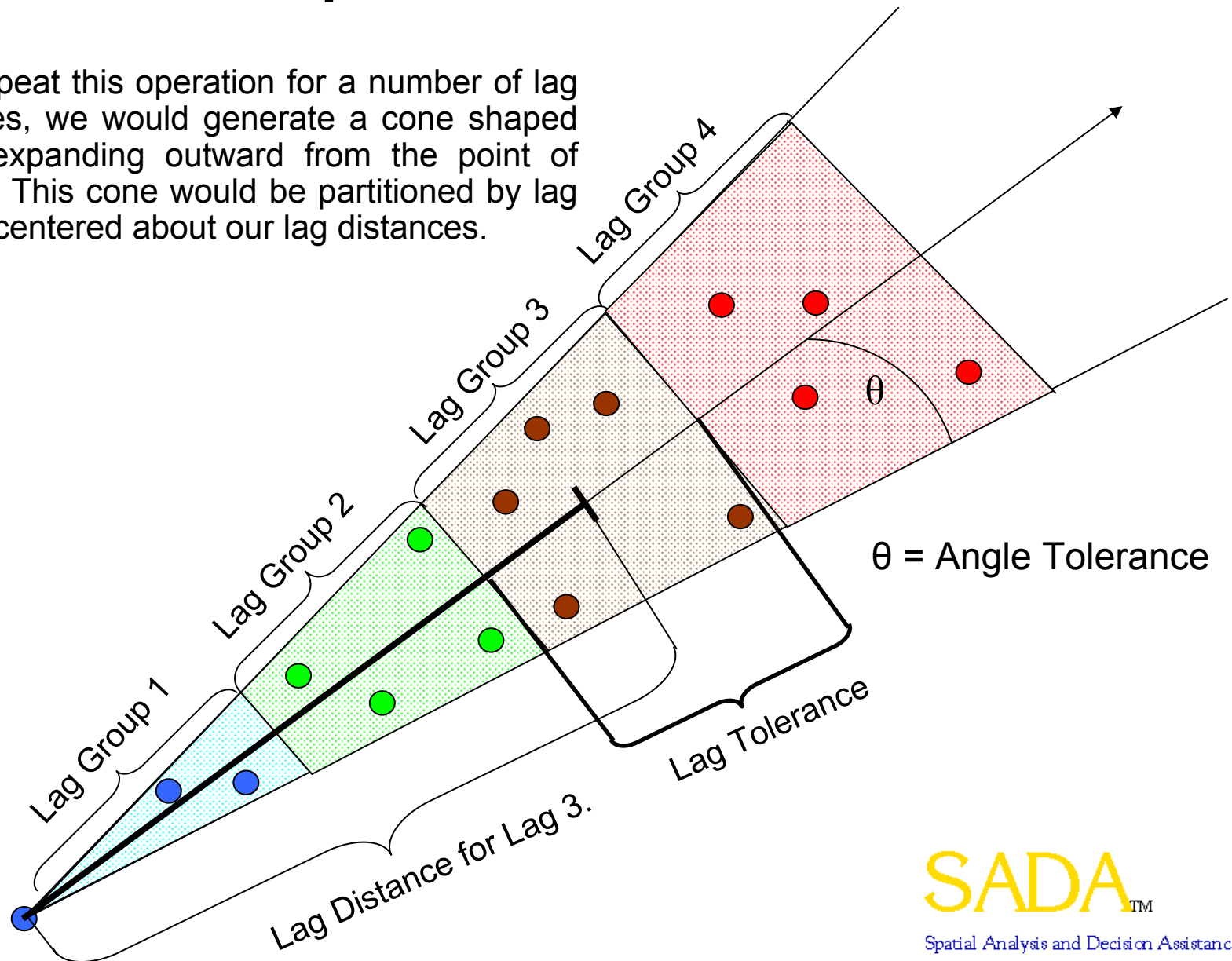
Although assigning a lag tolerance helps, most cases will never have enough samples separated by a lag - tol/2 to lag + tol/2 along a straight line to calculate the semivariogram value. Therefore, an angle tolerance, θ , is also introduced to expand the region and to include more points in the calculation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



θ = Angle Tolerance

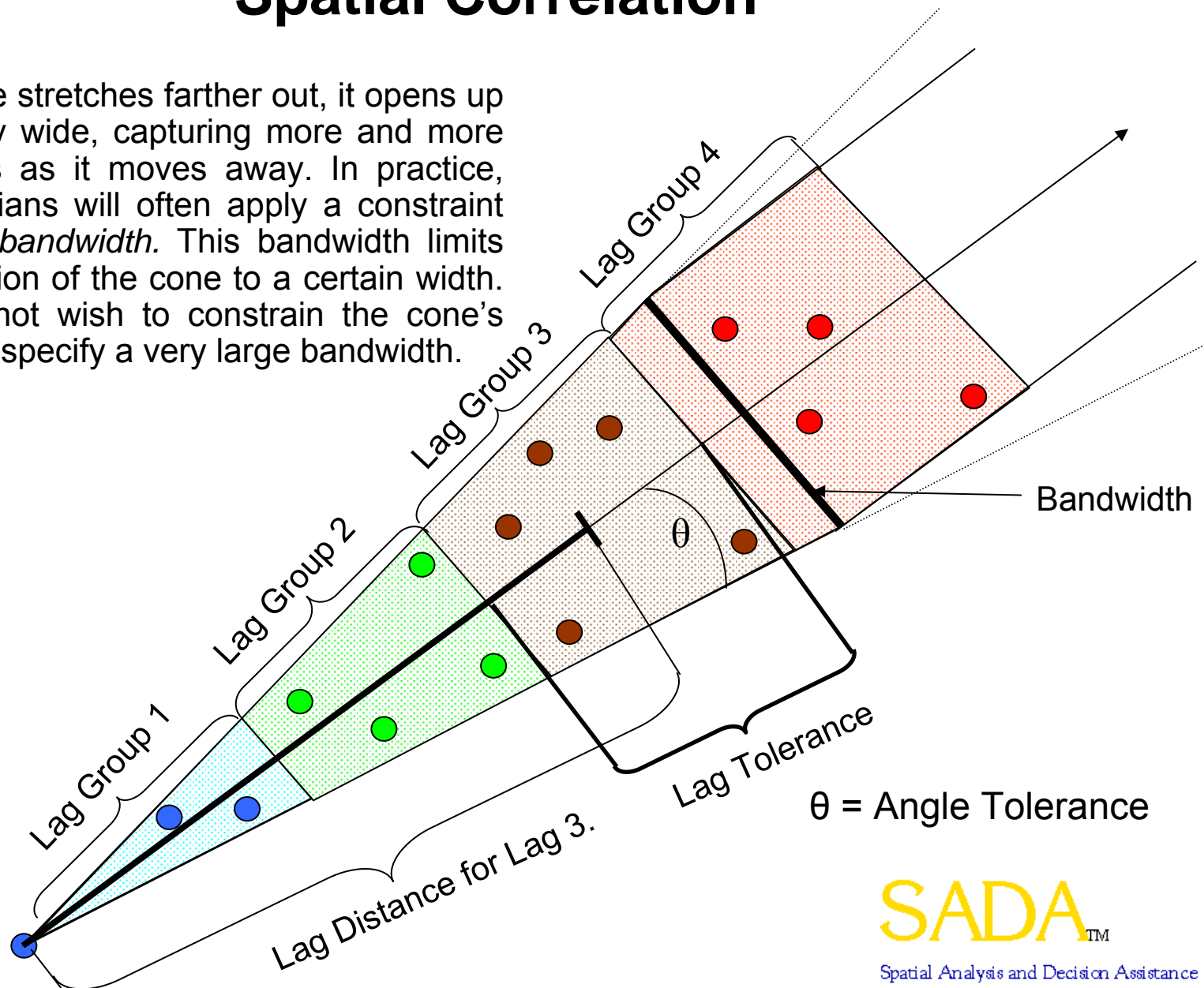
Spatial Correlation

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.



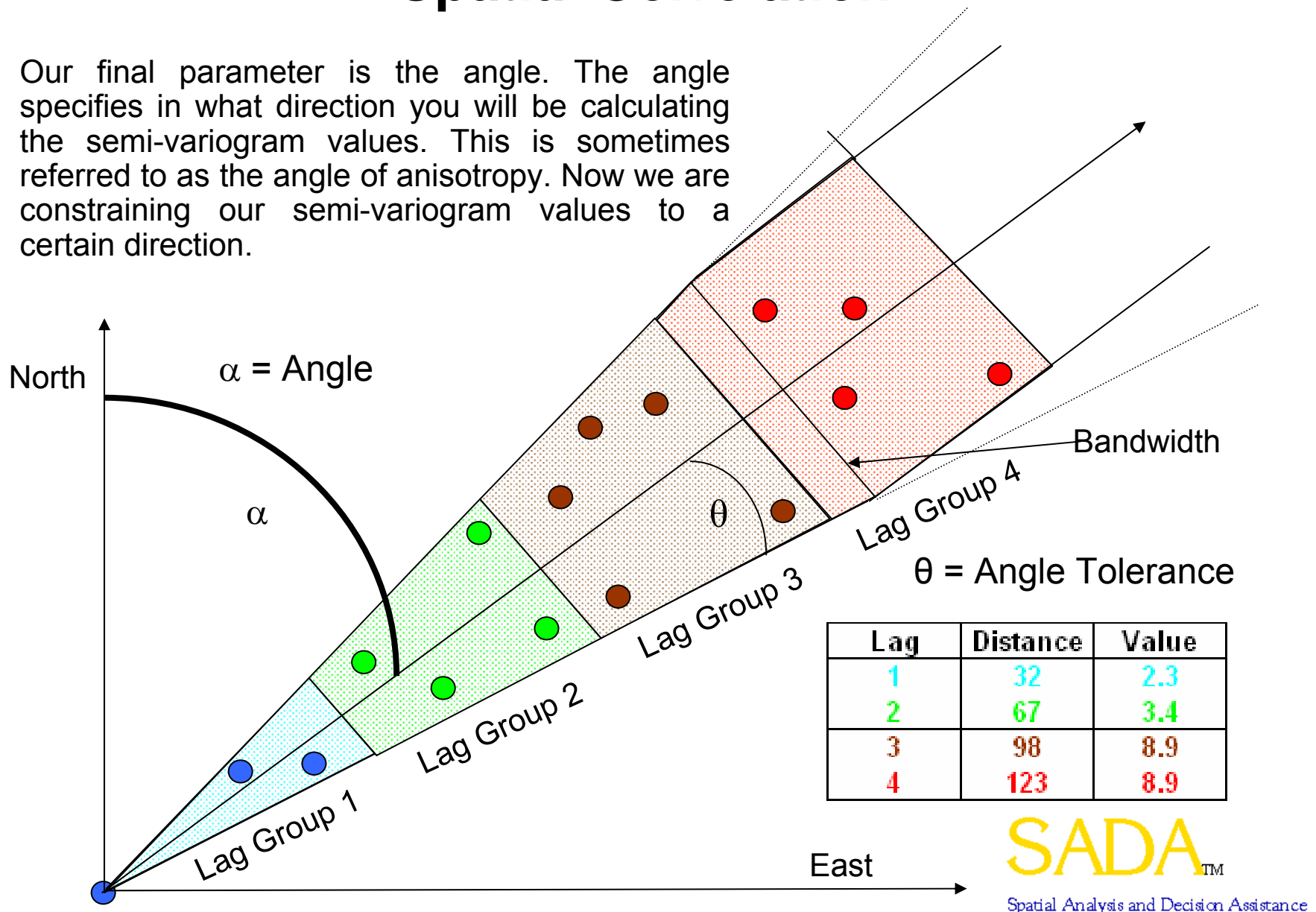
Spatial Correlation

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.

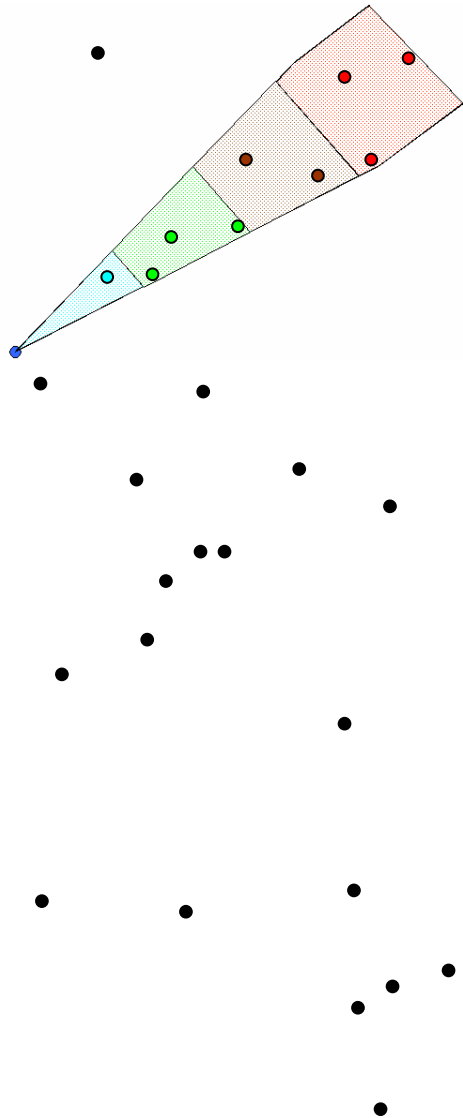


Spatial Correlation

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.



Spatial Correlation



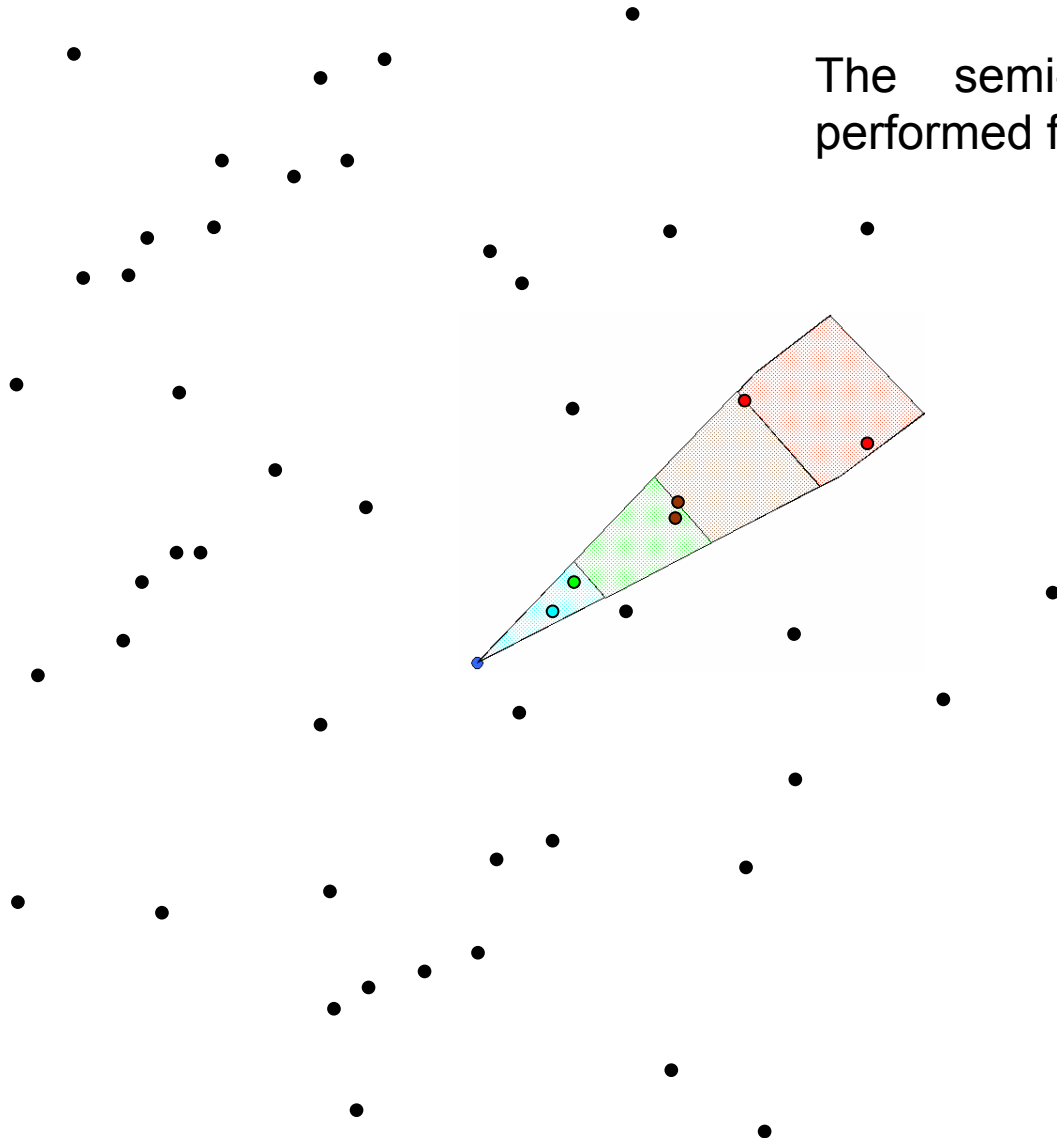
The semi-variogram calculation is performed for every sampled point.

SADATM

Spatial Analysis and Decision Assistance

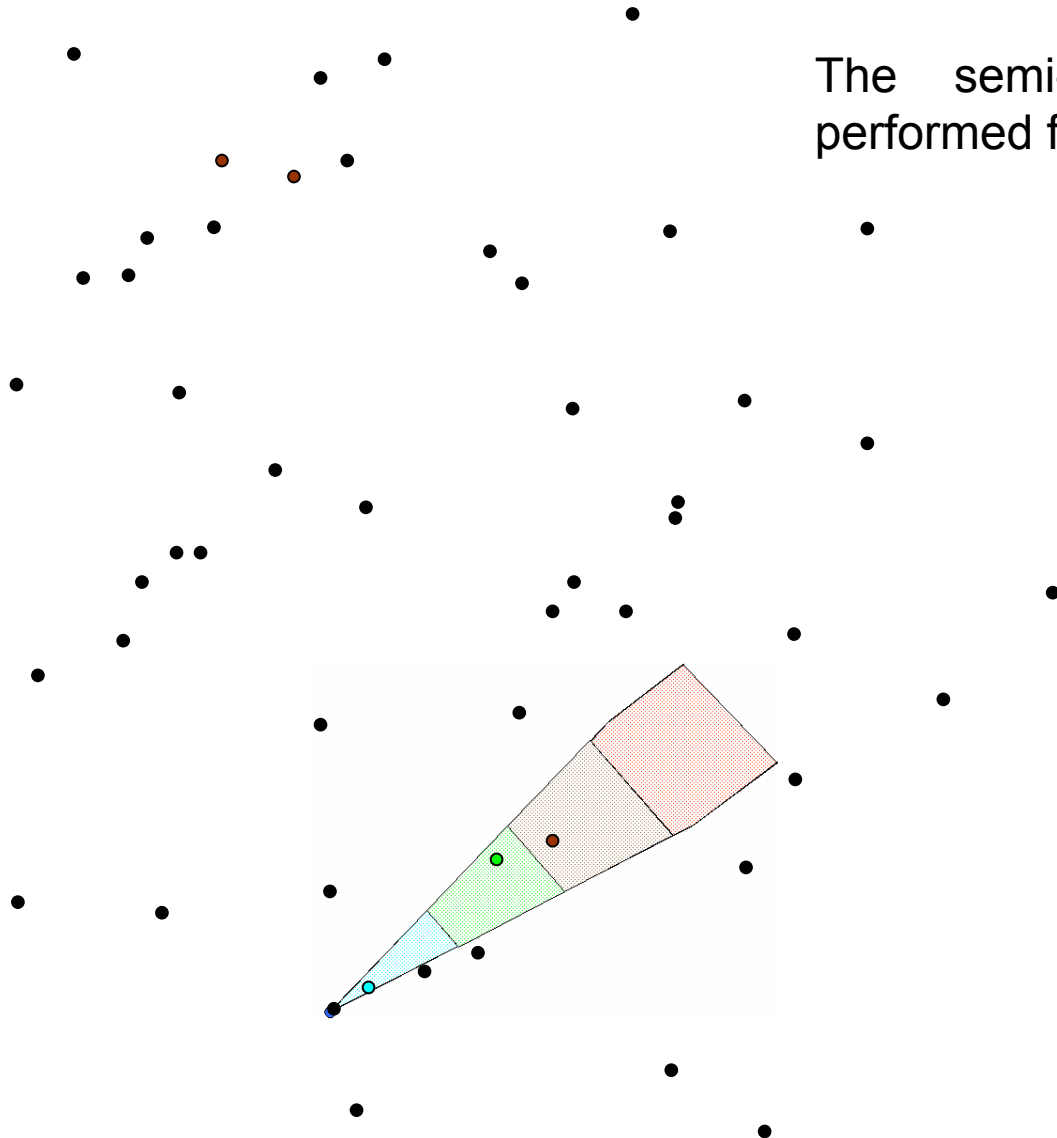
Spatial Correlation

The semi-variogram calculation is performed for every sampled point.



Spatial Correlation

The semi-variogram calculation is performed for every sampled point.

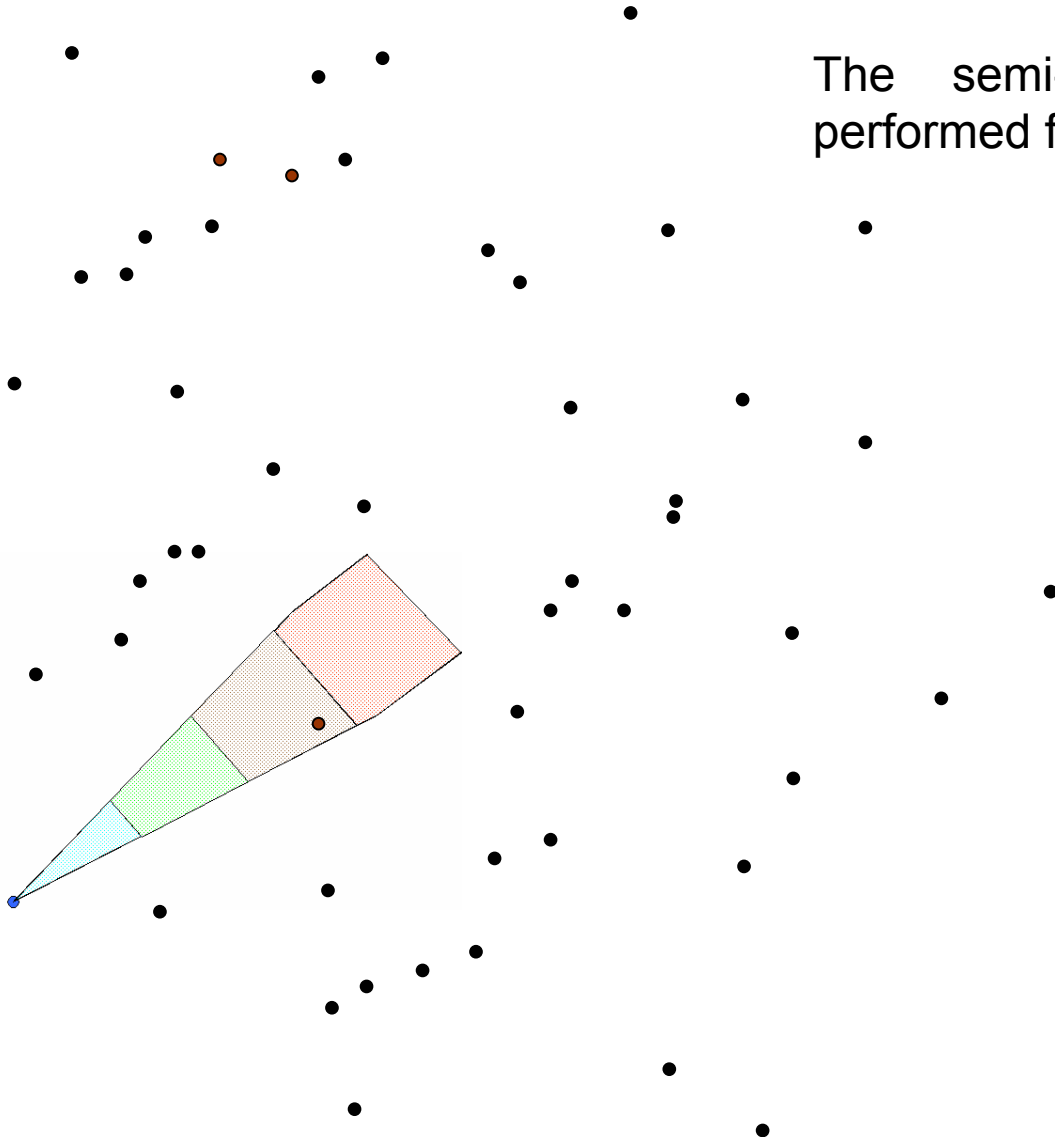


SADATM

Spatial Analysis and Decision Assistance

Spatial Correlation

The semi-variogram calculation is performed for every sampled point.



SADATM

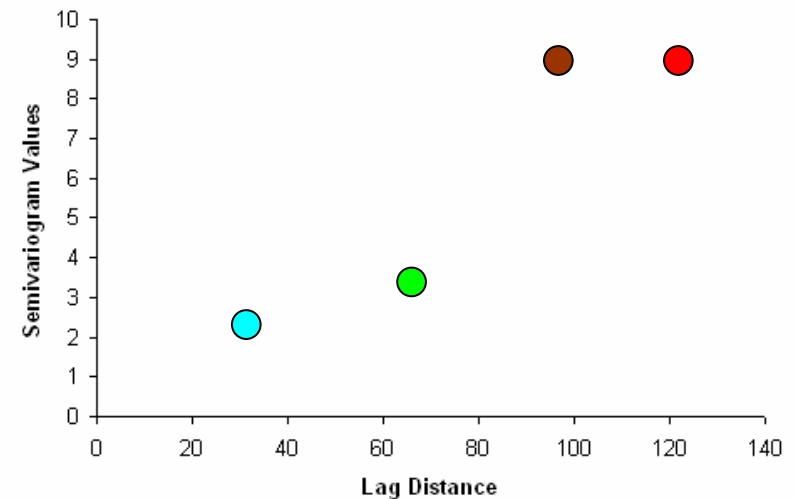
Spatial Analysis and Decision Assistance

Spatial Correlation

The semi-variogram values are then plotted.

Lag	Distance	Value
1	32	2.3
2	67	3.4
3	98	8.9
4	123	8.9

Semi-variogram Plot

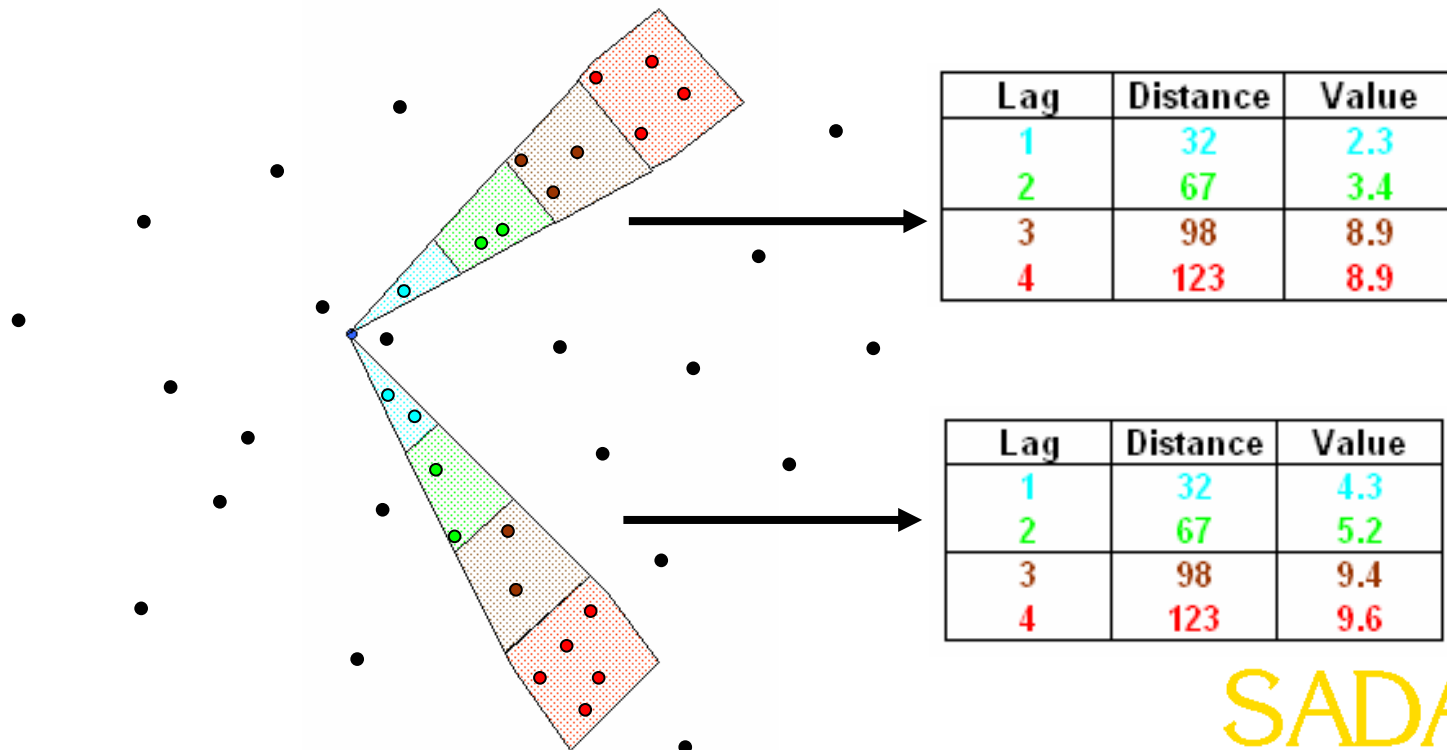


SADATM

Spatial Analysis and Decision Assistance

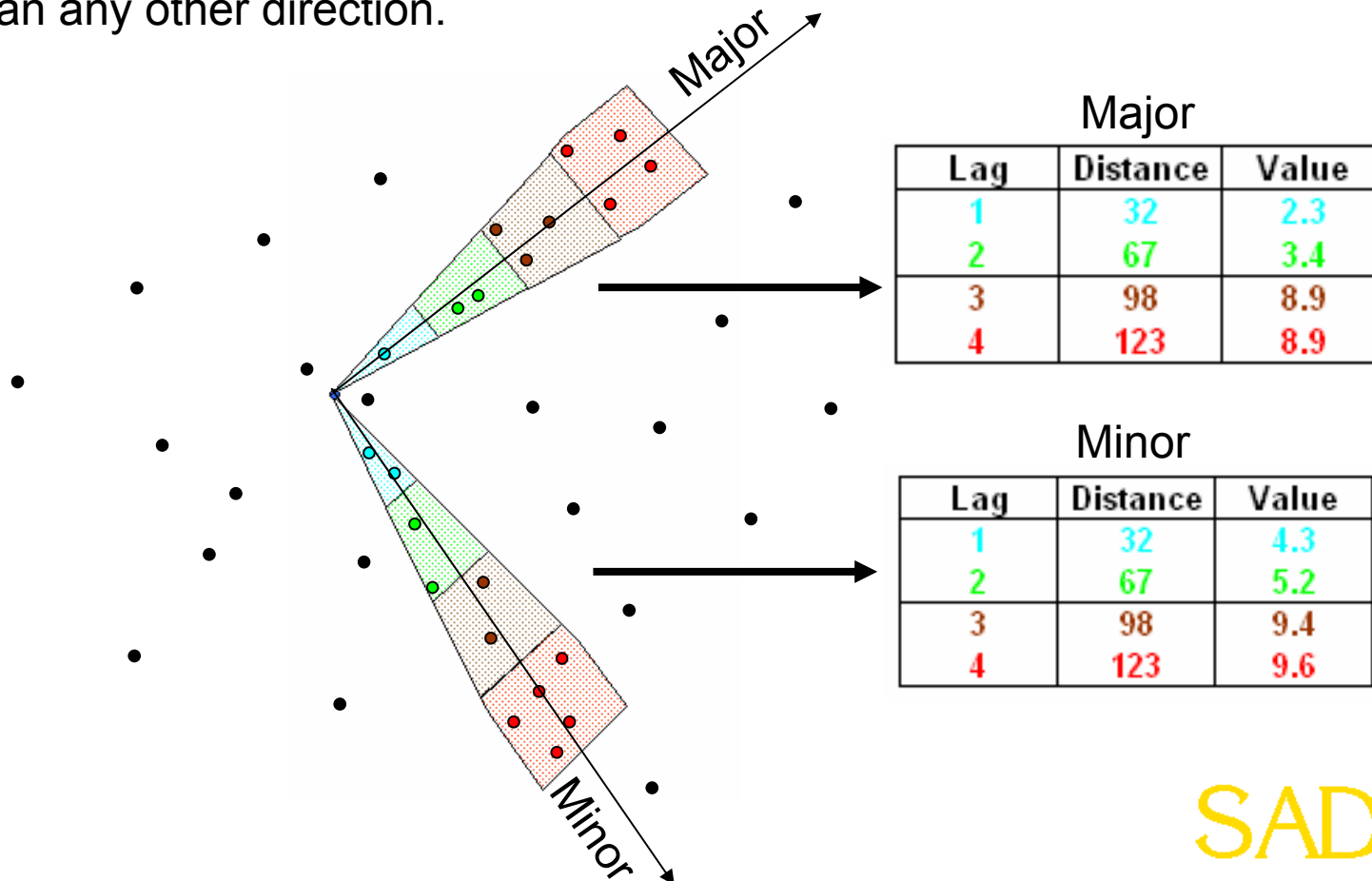
Spatial Correlation

Note though, by specifying an angle α , we are excluding all those data points located outside of the cone from $\alpha - \theta$ degrees to $\alpha + \theta$ 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.



Spatial Correlation

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.

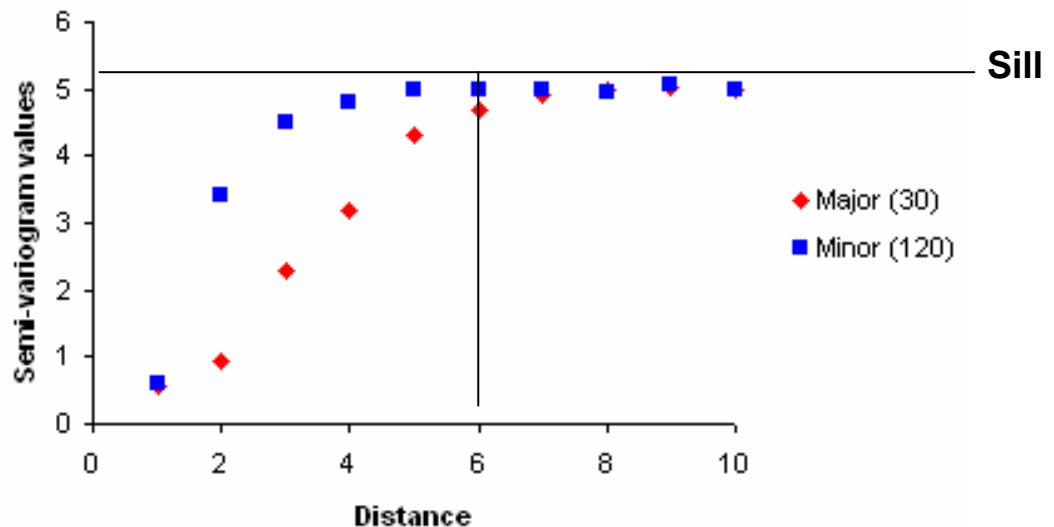


Spatial Correlation

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.

Semi-variogram plot

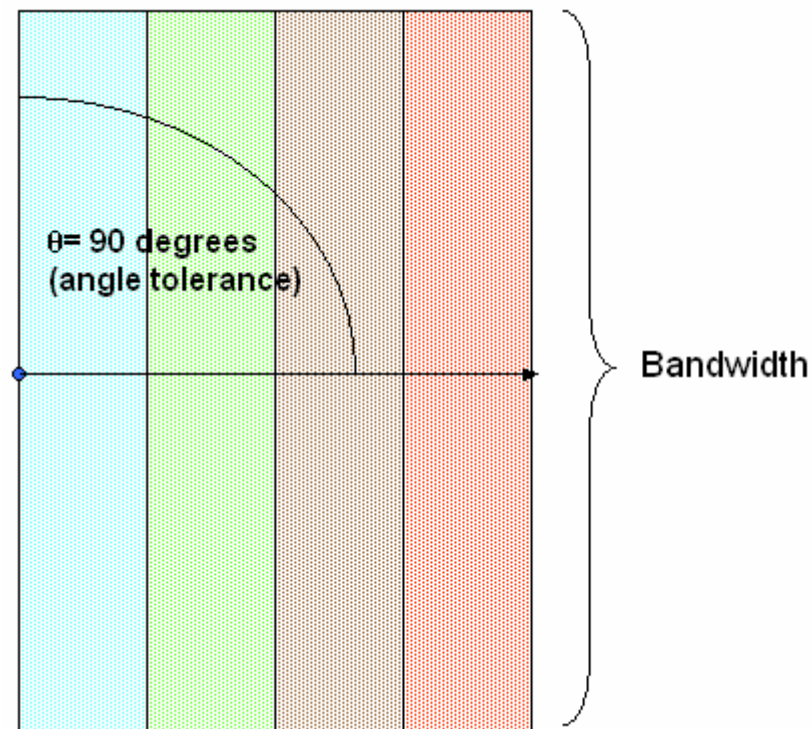
Lag	Major (30)	Minor (120)
1	0.56	0.6
2	0.95	3.4
3	2.3	4.5
4	3.2	4.8
5	4.3	5
6	4.7	5
7	4.9	5
8	5	4.95
9	5.01	5.05
10	4.99	5



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.



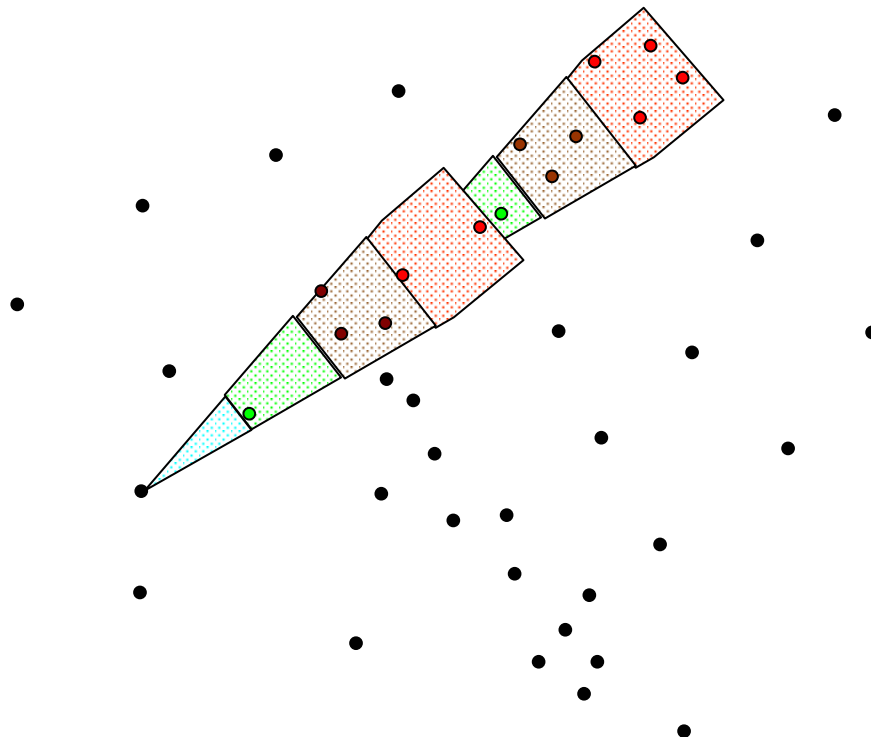
Spatial Correlation

What about in the opposite direction?

It is assumed that correlation is symmetrical. If data are varying a certain amount at 30 degrees, then they are varying the same amount at 120 degrees.

Why don't we include those sample points in the 120 degrees direction to improve our semi-variogram calculation?

We do. Our current point of interest will be captured by the cone of those points behind it.



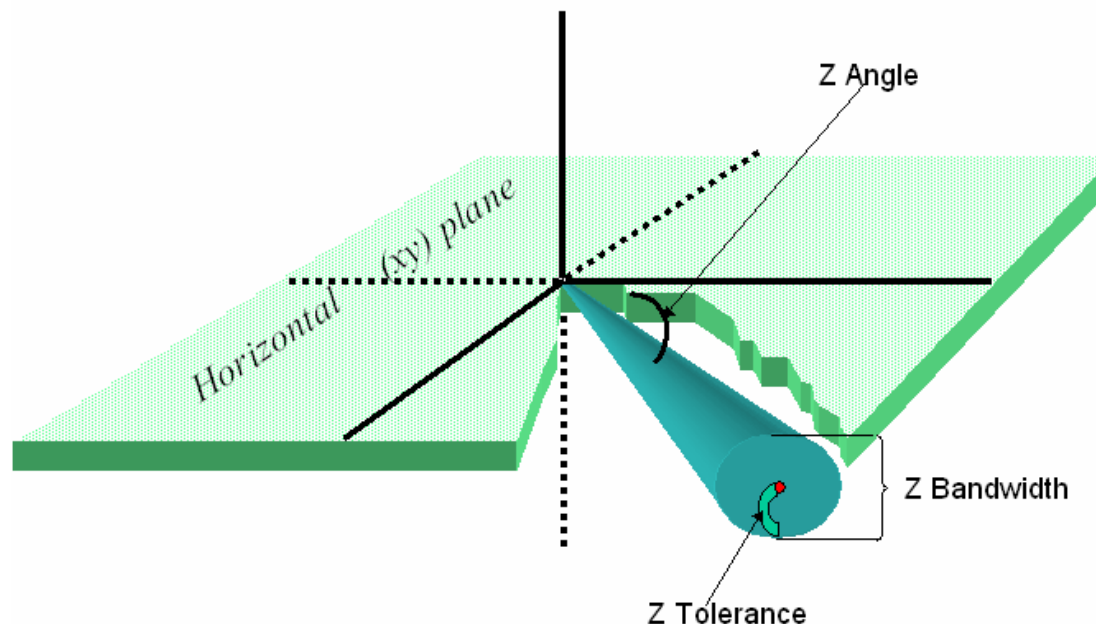
Three-Dimensional Variography

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.



SADATM

Spatial Analysis and Decision Assistance

Setting Variography Model

To calculate semivariogram values, select **Correlation Modeling** from the **Steps Window** and enter the appropriate information on the **Parameters Window**. The results of two separate cones are viewed at once to provide visual comparison and check for anisotropic correlation. Press **Show Me**.

Correlation Modeling

Correlation Modeling

Variogram Type

OK Edit

Variography

Variogram Both Rose

Name	Major	Minor
Caption	Major	Minor
Lag Number	6	6
Lag Distance	150	150
Lag Tol	100	85
Angle	90	45
Tol	45	70
Band	1000	1500
Dip	0	0
ZTol	90	90
ZBand	1000	1000

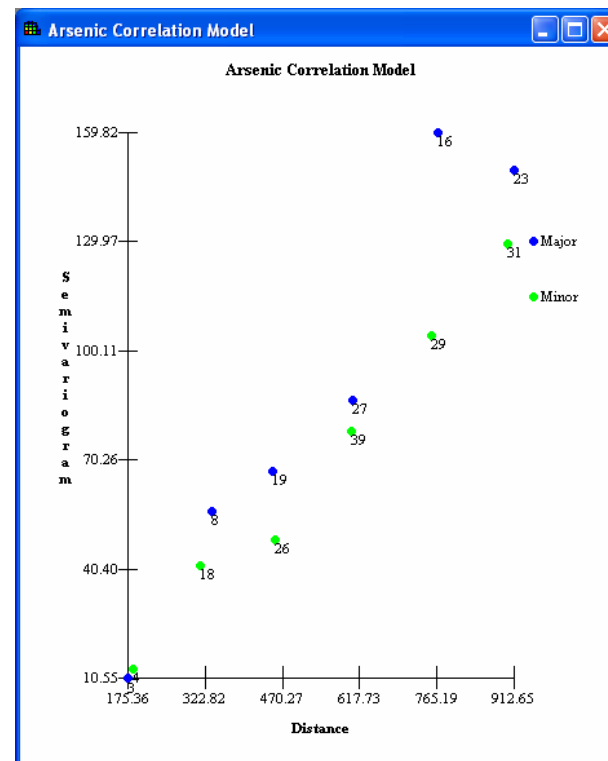
Modeling

Model Not Used Not Used

Major Range	
Minor Range	
Angle	
Contribution	
Z Angle	
Z Range	
Rotation	

Autofit Nugget

Show Me

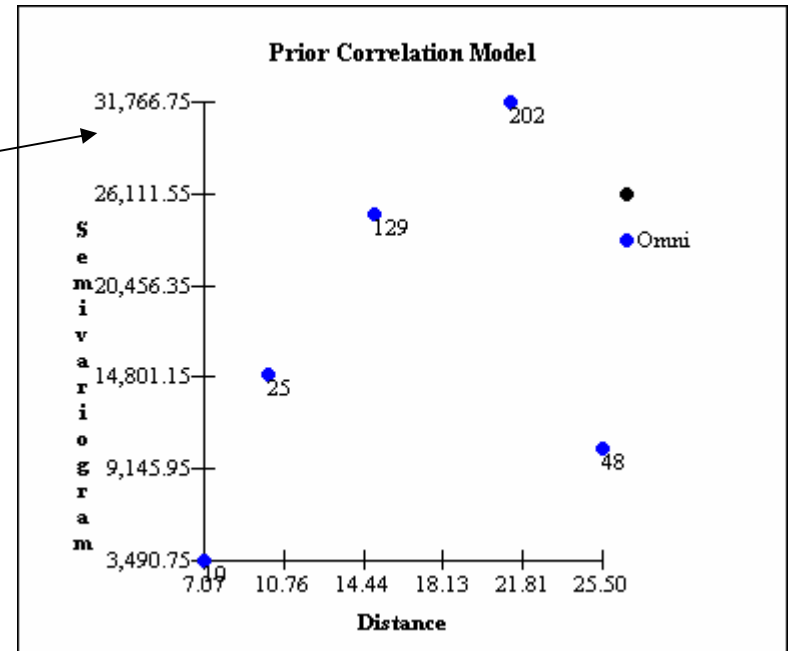
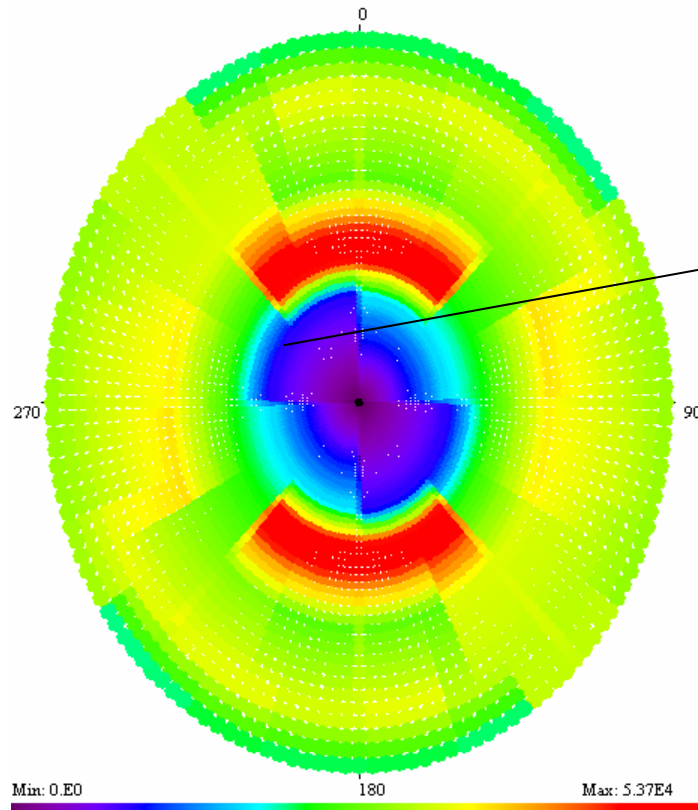


SADATM

Spatial Analysis and Decision Assistance

Rose Diagrams

Rather than viewing only on angle at a time, users can view semivariogram values in all directions at once. They can then choose an angle of interest by clicking on the rose diagram map. SADA will show the semivariogram values for that direction.



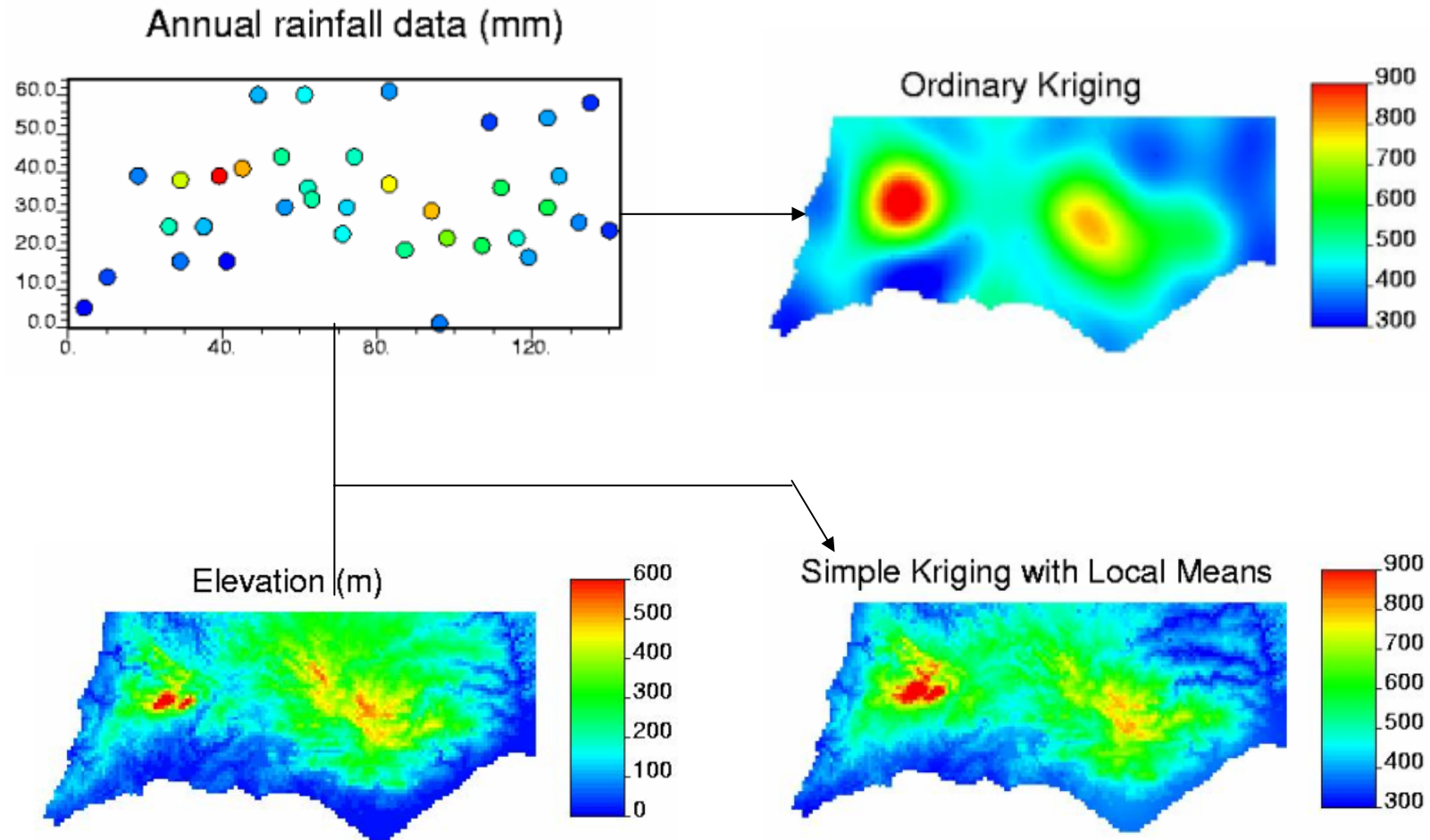
SADATM

Spatial Analysis and Decision Assistance

Secondary Information

- The term secondary information describes a collection of information that may be either quantitative or qualitative in nature. This collection of information is not the direct subject of interest. It is however related and may assist in characterization of the primary subject, particularly within a spatial context.
- Direct measurements of the subject may be costly or perhaps dangerous to obtain. This results in only a few explicit samples.
- If secondary information is available in great quantities, it may improve heterogeneity in the final results.

Impact of Secondary Information



*Taken from Pierre Goovaerts' Presentation "Performance comparison of geostatistical algorithms for incorporating elevation into the mapping of precipitation"

SADATM

Spatial Analysis and Decision Assistance

Geobayesian History

- The U.S. Nuclear Regulatory Commission is interested in explicitly using all relevant information about a contaminated site to create a better design strategy for subsurface (3d) sampling.
- The interest originates from final status decommissioning surveys conducted by NRC.
- Issues in two-dimensional applications have been worked out in the MARSSIM guidance developed by NRC, EPA, and DOE.
- MARSSIM guidance is best suited for 2d applications because of the role that walk over radiological scans play in the process.
- The goal is to identify an analogous approach to MARSSIM for 3d, particularly when faced with sparse data sets.

Geobayesian History

- In many cases, potentially useful information is known about the site that can drive sampling and characterization strategies. These may include site history, geology, and previous sampling.
- An approach is needed to explicitly use these varying sources of information in a formal geospatial framework to drive the location of final survey samples, to characterize the radiological risk, and to support closure decisions.
- A number of approaches are being evaluated. The first is a method originally formalized as the Adaptive Sampling and Analysis Program (ASAP) at Argonne National Laboratory. This method integrates a standard bayesian approach with indicator kriging.
- This approach was the basis for the SADA geobayesian module, created by NRC and the University of Tennessee in the freeware program SADA.
- Other models under consideration include Co-kriging, and Markov Bayes.

SADATM

Spatial Analysis and Decision Assistance

The Geobayesian Model

Historical Documents

EMPLOYEE/VENDOR/PAYEE IDENTIFIED BELOW HAS OVERPAID THE NUCLEAR REGULATORY COMMISSION FOR GOODS AND/OR SERVICES PROVIDED AND IS DUE A REFUND

EMPLOYEE/VENDOR/PAYEE CODE: _____

NAME: New England Community Hospital

ADDRESS: 444 Elm Street, CT 06102

CITY: St. Louis STATE: MO ZIP: 63121

TRANS CODE: 21

TRANS TYPE: _____ FUND: _____ JOB CODE: _____ AMOUNT: \$400.00

TRANS TYPE: 18 FUND: 81525 JOB CODE: INTR AMOUNT: _____

TRANS TYPE: 18 FUND: 81029 JOB CODE: ADSH AMOUNT: _____

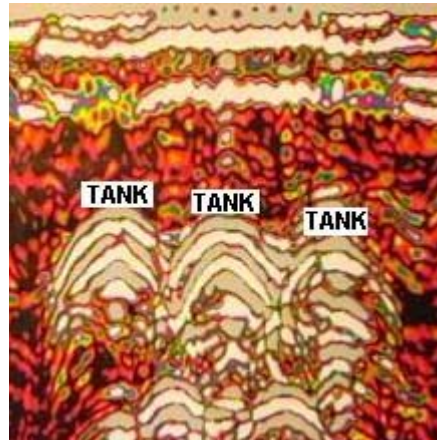
TRANS TYPE: 18 FUND: 81029 JOB CODE: FINE AMOUNT: _____

TOTAL REFUND AMOUNT: \$400.00

COMMENTS: App dtd 6/20/97/OK 2761/75 crq gmt

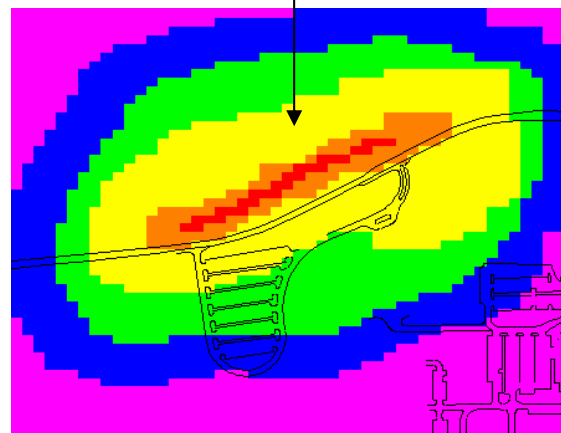
PREPARED BY: Shirley G. Hild DATE: July 11, 1997

Geotechnical data



Previous data
(secondary or direct)

X	Y	Z	CAS	Names	Values	Detect
27273.89	23261.16	2.6	1234123	Chlordane	0.00015	
27273.89	23261.16	5.2	1234123	Chlordane	0.00027	
27273.89	23261.16	7.8	1234123	Chlordane	0.00046	
27273.89	23261.16	10.4	1234123	Chlordane	0.00045	
27273.89	23261.16	13	1234123	Chlordane	0.00028	
27273.89	23261.16	15.6	1234123	Chlordane	0.00012	
27273.89	23261.16	18.2	1234123	Chlordane	0.00004	
27273.89	23261.16	20.8	1234123	Chlordane	0	
27273.89	23261.16	23.4	1234123	Chlordane	0	
27273.89	23261.16	26	1234123	Chlordane	0	
27512.91	23260.99	2.6	1234123	Chlordane	0.46147	
27512.91	23260.99	5.2	1234123	Chlordane	0.69144	
27512.91	23260.99	7.8	1234123	Chlordane	0.73024	
27512.91	23260.99	10.4	1234123	Chlordane	0.38323	



Probability of Exceedance

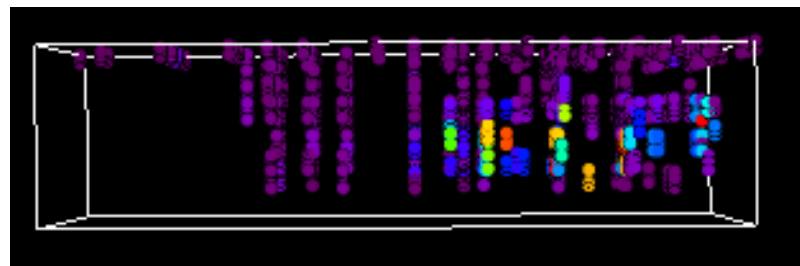
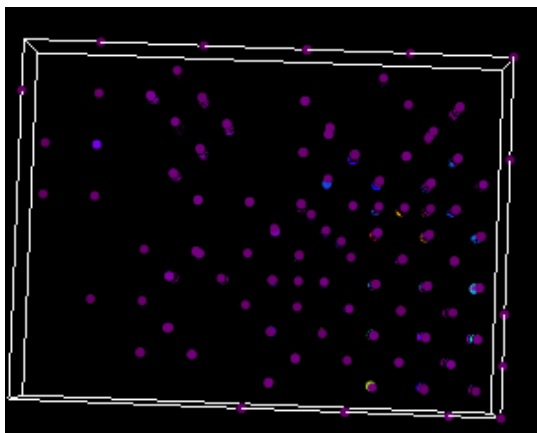
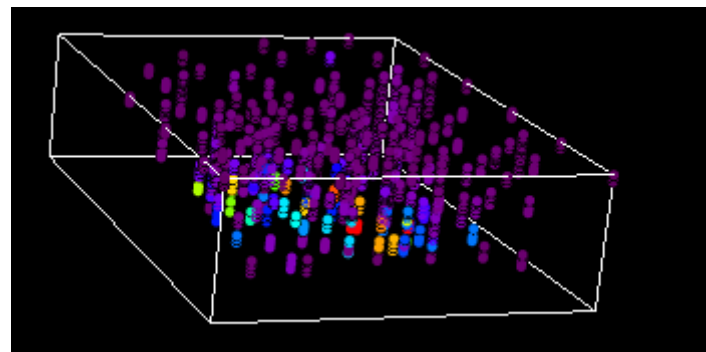
SADATM

Spatial Analysis and Decision Assistance

Case Study: Site Description

The KISKI Data Set

- Used as an example data set to test Geobayesian modeling.
- 1261 samples in shallow sediment.
- ~90 boreholes.
- Values range from near zero to 900 pCi/g.
- Contaminant name was changed.
- Large number of data, but typical spatial distribution.
- Good starting point for evaluating the new Geobayesian approach.

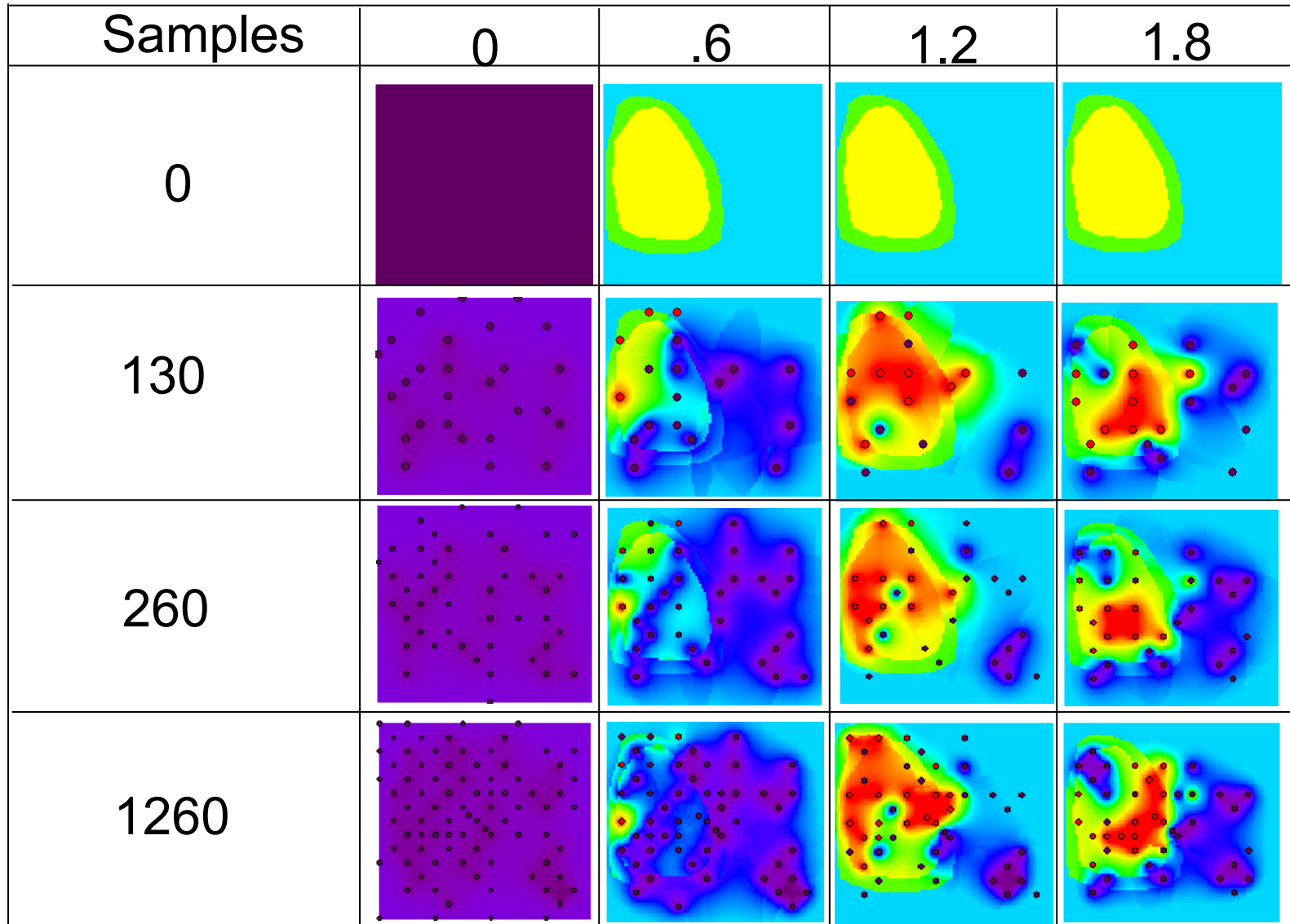


SADATM

Spatial Analysis and Decision Assistance

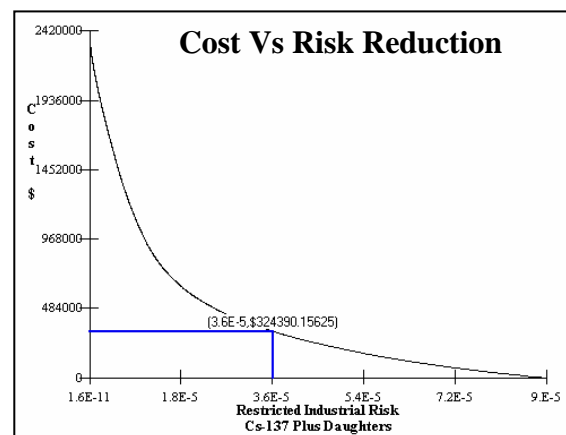
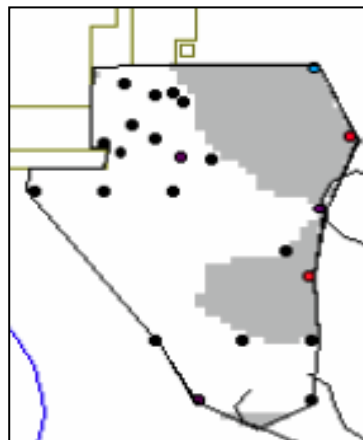
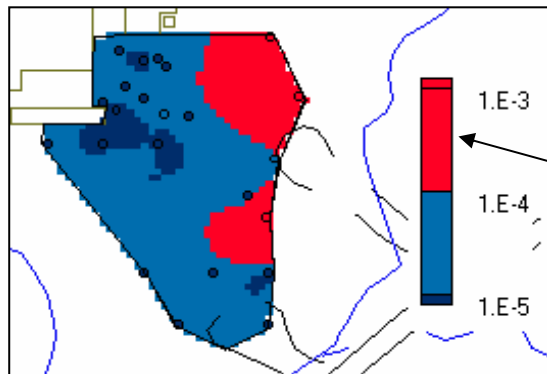
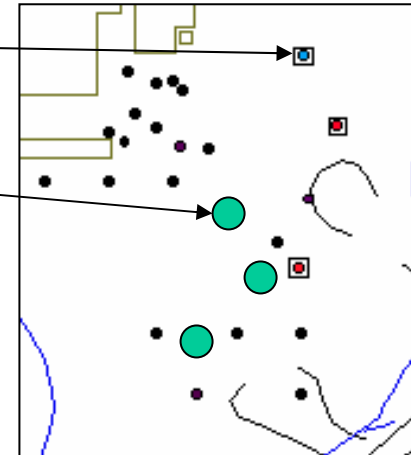
Case Study: Iterative Sampling

Depth



Decision Analysis

- Spatial Screens
- Sampling Strategies



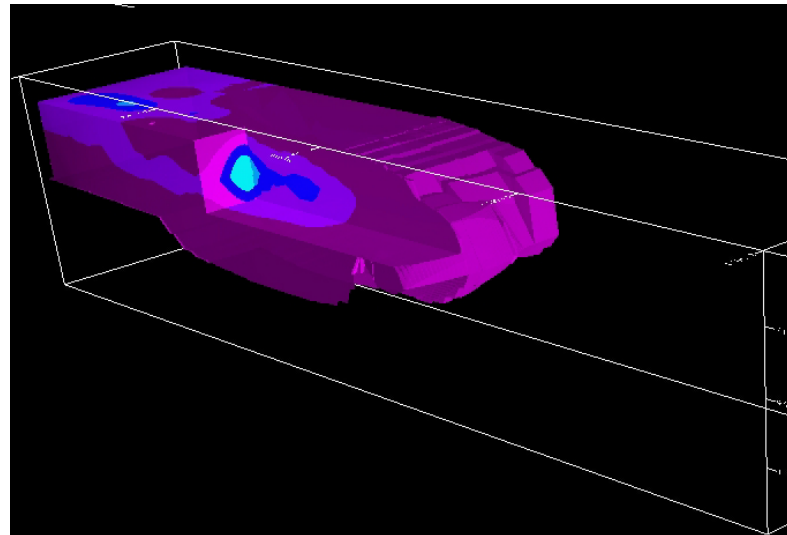
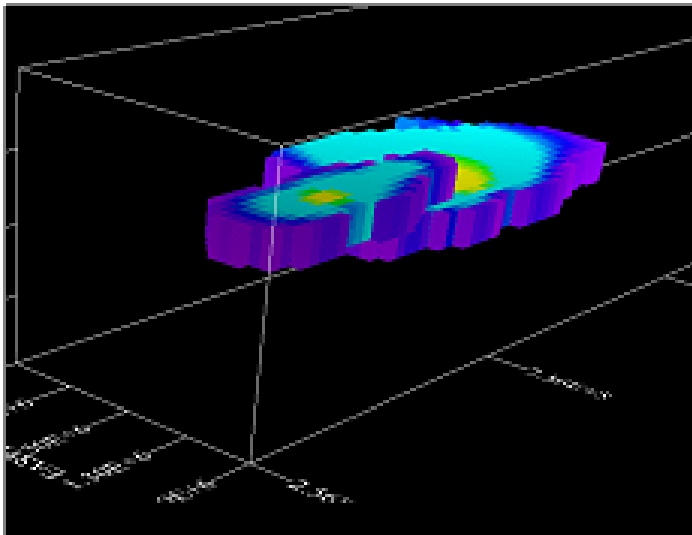
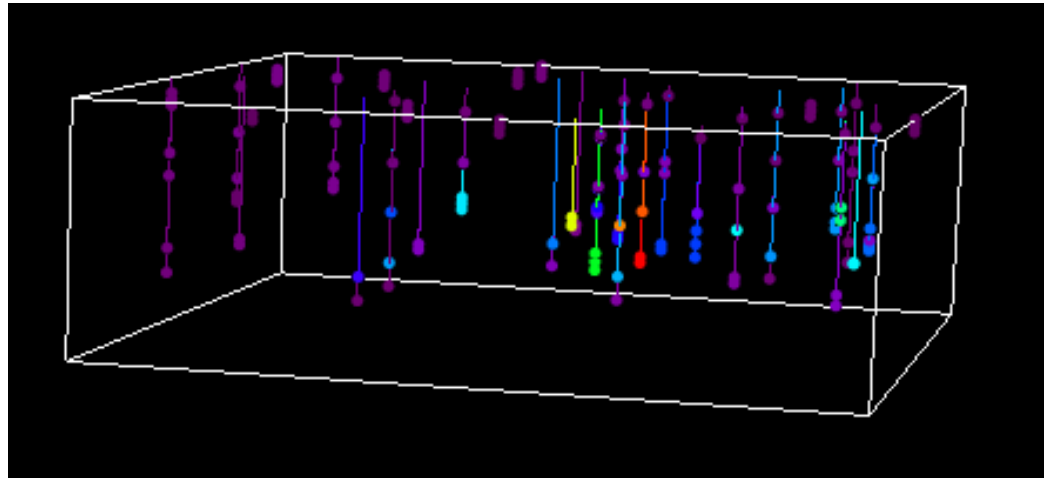
- Spatial Risk
- Area of Concern
- Cost Benefit

SADATM

Spatial Analysis and Decision Assistance

3D Visualization

True 3d Views: Points,
Blocks, and Isosurfaces



SADATM

Spatial Analysis and Decision Assistance

Sample Designs

SADA has a number of sample design strategies in Version 4.0. These strategies include initial and secondary designs. Some are based on data alone while others are based on modeling results. With the exception of a couple of exclusively 2d designs all are available in 3d dimensions.

Initial Sample Designs

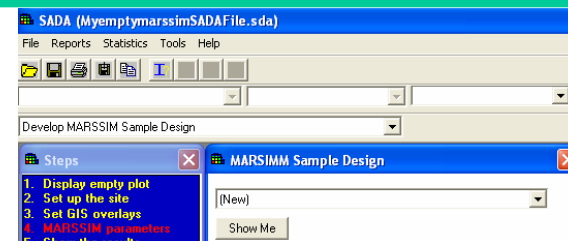
- Judgmental
- Simple Random
- Simple Grid
- Simple Unaligned Grid
- Standard Grid
- Standard Unaligned Grid
- ***MARSSIM Design***
- 2d and 3d Hot Spot search designs

Secondary Sample Designs

- Threshold Radial
- Adaptive Fill
- High Value
 - (soft, simulated & unsimulated)
- High Variance
 - (soft, simulated & unsimulated)
- Extreme Value
 - (soft, simulated & unsimulated)
- Area of Concern Boundary Design
 - (soft, simulated & unsimulated)
- Minimize/Maximize Area of Concern
- LISA Designs
 - (Ripley's K, Moran's I, Geary's C)

I want to create a MARSSIM sample design

- (1) Identify the survey area
- (2) Set Class I, II, or III based on extent of contamination suspected/known
- (3) Set WRS or Sign (background or not)
- (4) View/edit DCGL and associated values (DCGLw, LBGR, alpha, beta, sigma)
- (5) Show power curve, return N, alpha, beta
- (6) Get grid area (survey area/N)
- (7) Get grid area-area factor curve
- (8) Update AF for new grid area, calculate DCGL_{emc}, get MDC
- (9) Instrument sensitivity check
 - (1) If pass
 - (1) Show 2D Elipgrid results for circular hot spot of size grid area
 - (2) If fail
 - (1) Query for area factor based on updated grid area of (needed scan factor/DCGL)
 - (2) Recalculate N based on updated grid area and survey area
 - (3) Show elipgrid probabilities for both Ns and update grid area
 - (4) Accept original N and higher risk of missing circular hotspot or new N and lower risk of missing same hotspot size
- (10) Show MARSSIM grid or simple random sample design based on Class type



SADATM

Spatial Analysis and Decision Assistance

Calculate Grid Area and Enter Area Factor

Grid area is calculated based on the number of samples and the area of the site

Area Factor can be entered or retrieved from an excel file generated in RESRAD-MARSSIM

Click on Retrieve AF from RESRAD-MARSSIM

MARSSIM Parameters for Default

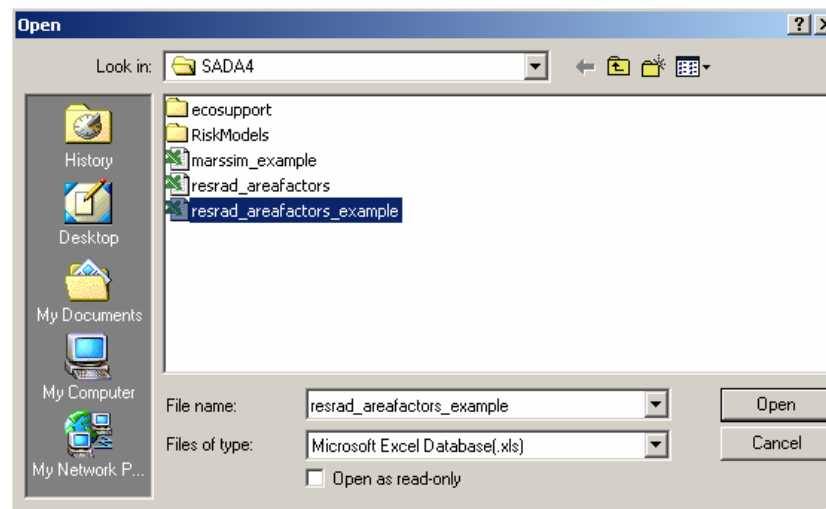
Step 1- Determine DCGLw and Sample Size

Sign Test		Sigma	0.2
DCGLw	1	alpha	0.05
LBGR	0.8	beta	0.05
		Sample Size	29

Step 2- Enter Area Factor

Grid Area (Survey area/ Sample Size) 0.0344828 (1/29)

Area Factor (AF) for Grid Area ≥ 1

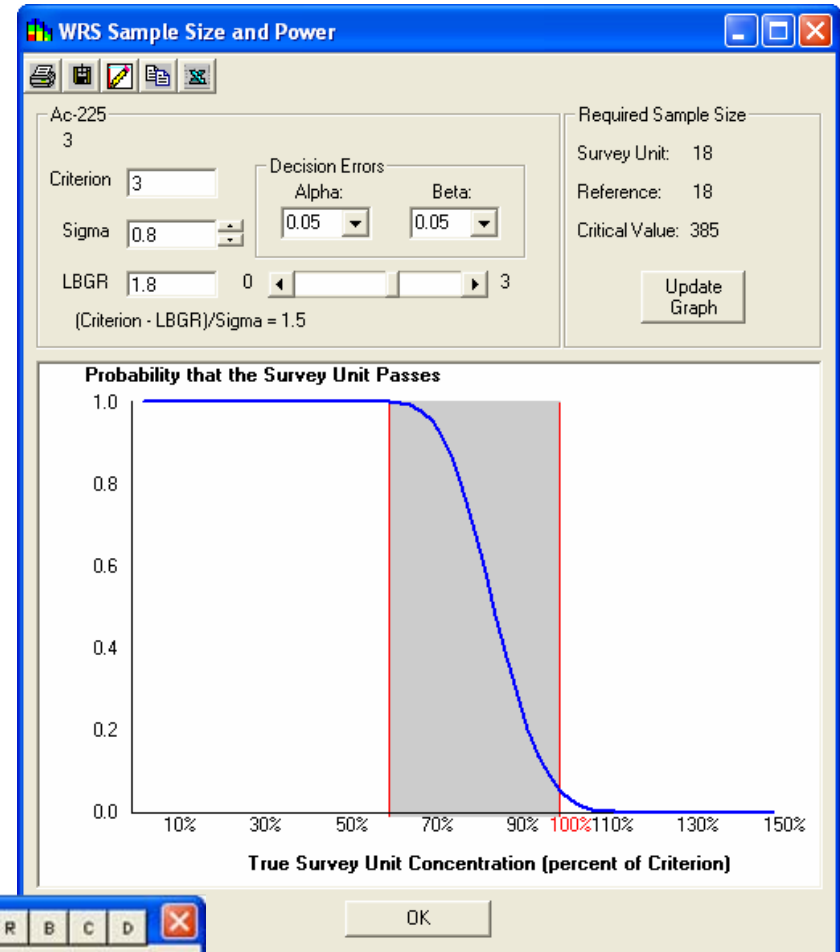


SADATM

Spatial Analysis and Decision Assistance

Determining Number of Samples – Wilcoxon Rank Sum

- User inputs DCGL, LBGR, and acceptable Type I and II error rates
- Appropriate for grid designs and simple random sampling
- Used when no background is available



SADA4 S W R B C D

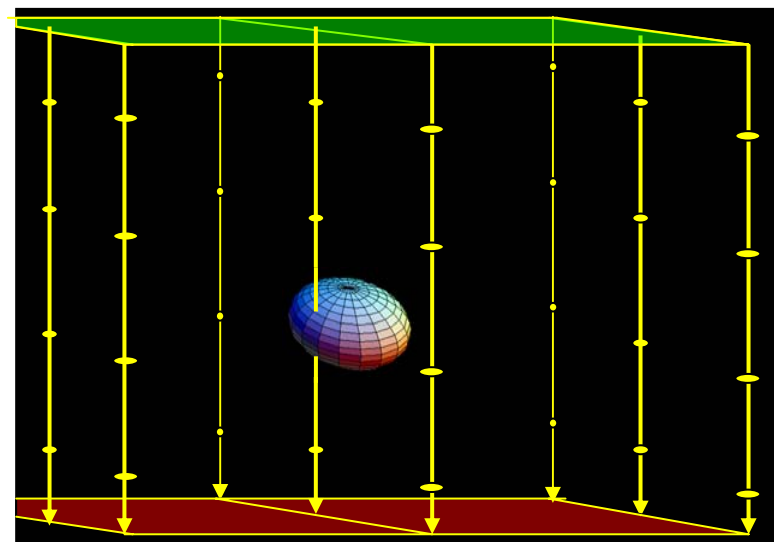
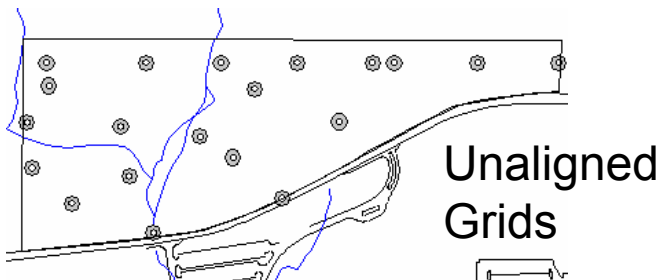
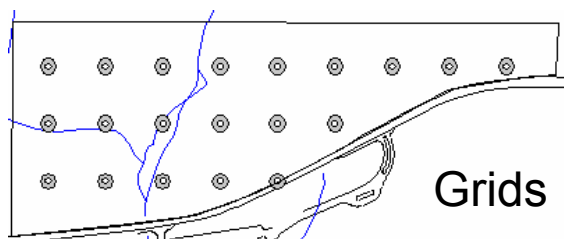
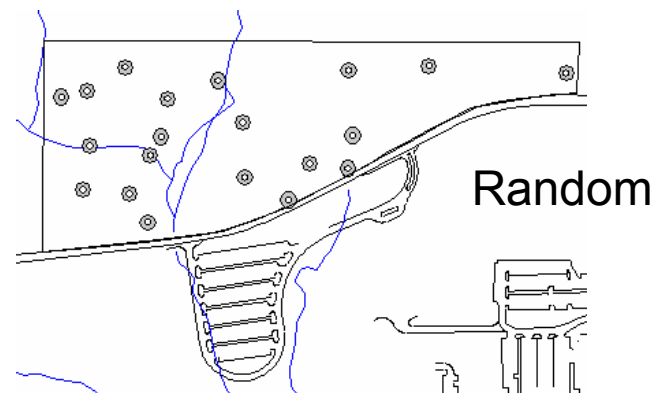
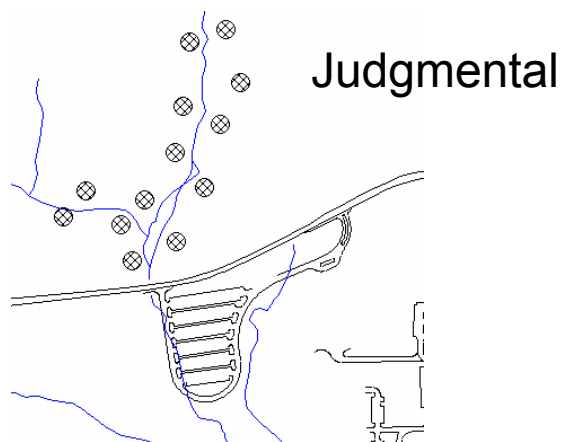
You need a total of 36 samples. You need 18 in your area of interest and 18 in your reference area. SADA will now plot the 18 samples in your area of interest.

OK

SADA™

Spatial Analysis and Decision Assistance

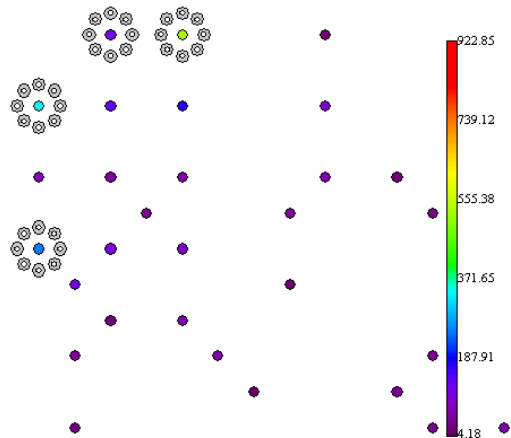
Some Example Initial Designs



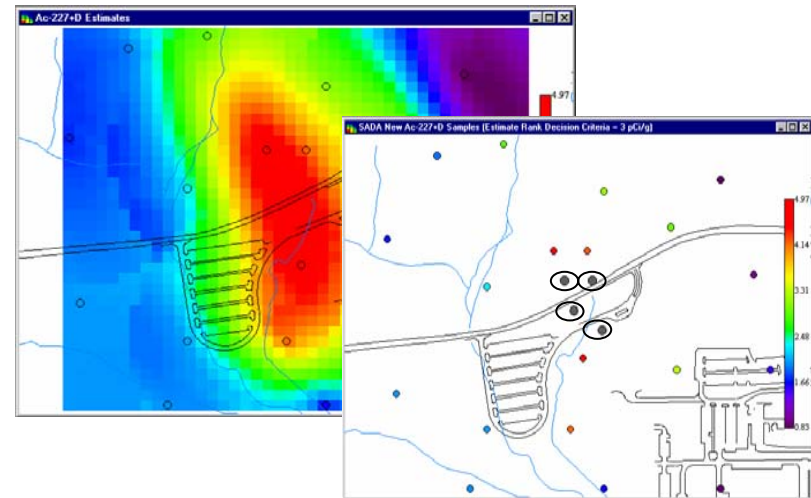
3d hotspot
search

Some Example Secondary Designs

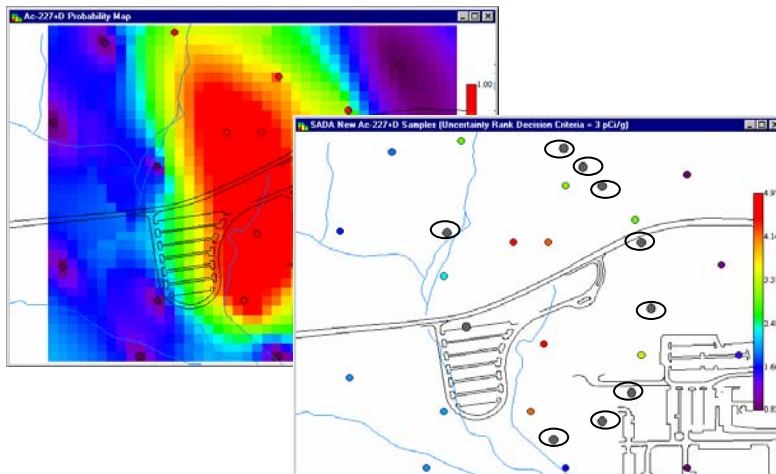
Threshold Radial



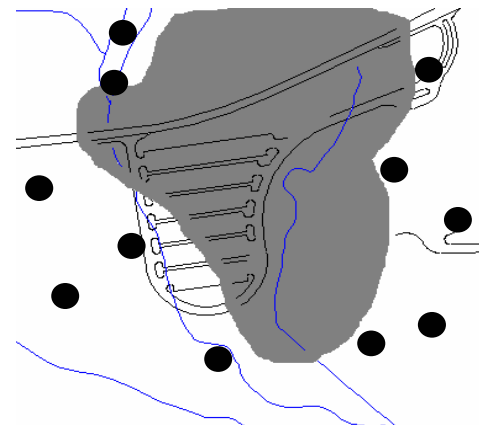
High Value Design



AOC Boundary Design



Min/Max AOC



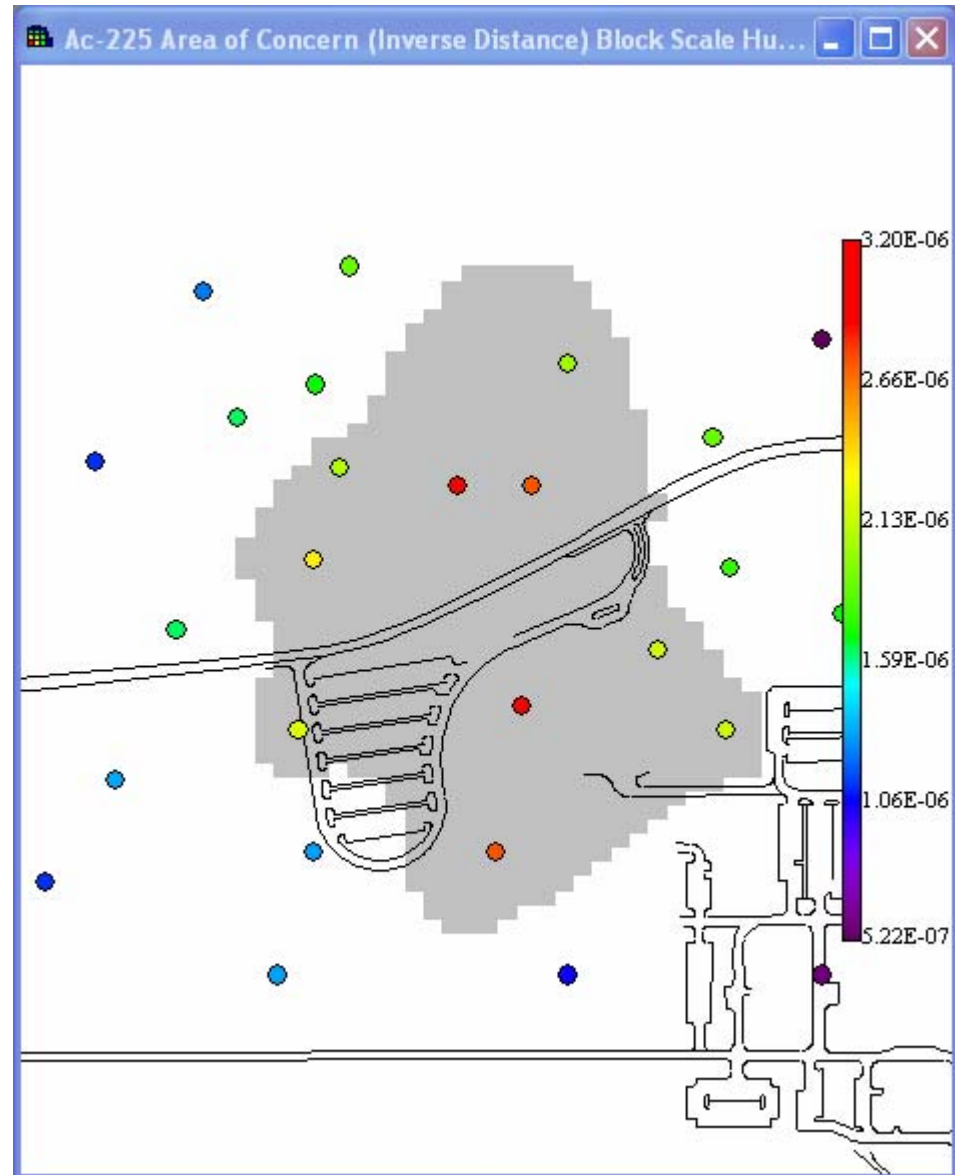
SADA Overview: Autodocumentation

- Provides transparency in the modeling process and facilitates reproducibility of results.
- SADA automatically analyzes any current result and determines what the “ingredients” of that result are. These ingredients are presented to the user, who can choose the level of documentation to create.
- Self-documentation of all parameters, models, and other relevant information.
 - Exposure concentrations
 - Risk models
 - Exposure variables
 - Geospatial parameters
 - Toxicity data
 - Images as bitmaps
- HTML format, can be exported to popular word processors

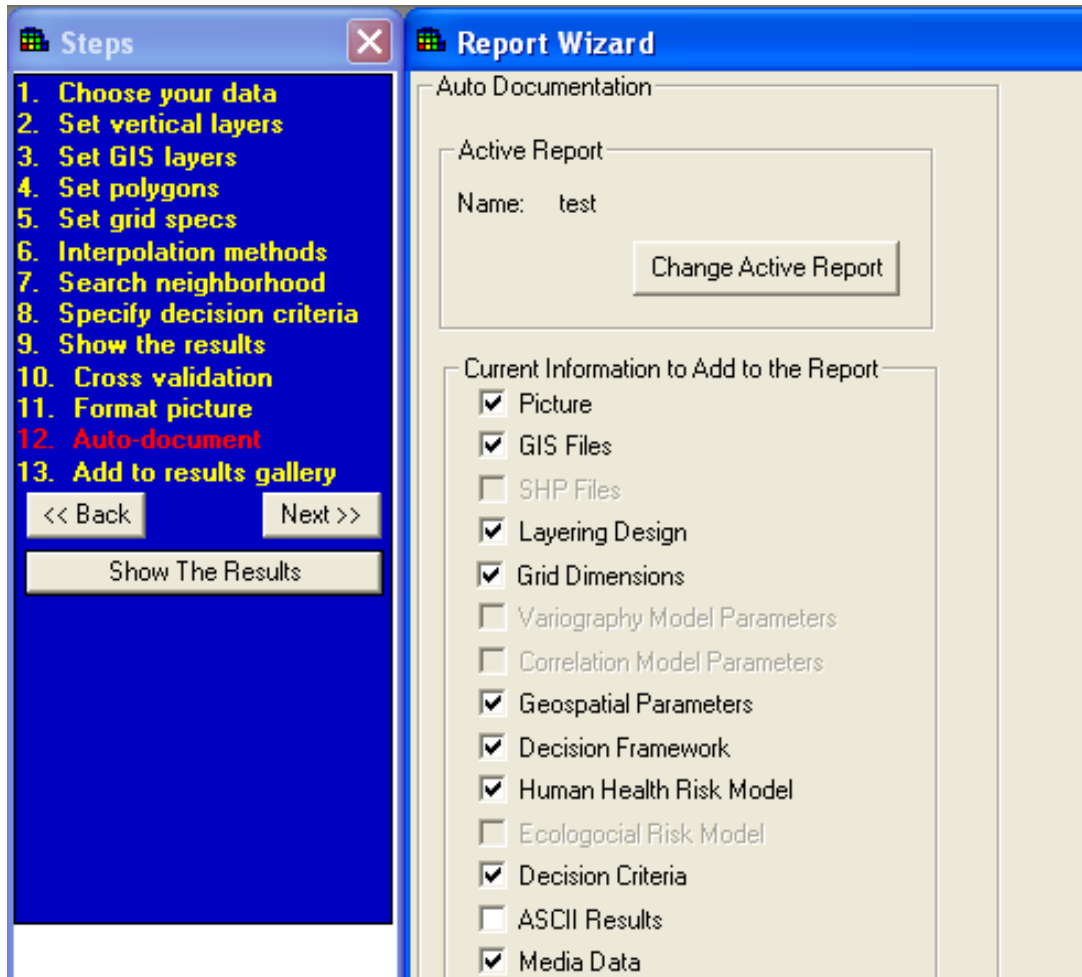


SADA Overview: Autodocumentation

- Area of concern map
- Based on HH Risk
- Utilized inverse distance as geospatial model
- Block based area of concern framework.



SADA Overview: Autodocumentation



Steps

1. Choose your data
2. Set vertical layers
3. Set GIS layers
4. Set polygons
5. Set grid specs
6. Interpolation methods
7. Search neighborhood
8. Specify decision criteria
9. Show the results
10. Cross validation
11. Format picture
12. Auto-document
13. Add to results gallery

<< Back Next >>

Show The Results

Report Wizard

Auto Documentation

Active Report

Name: test

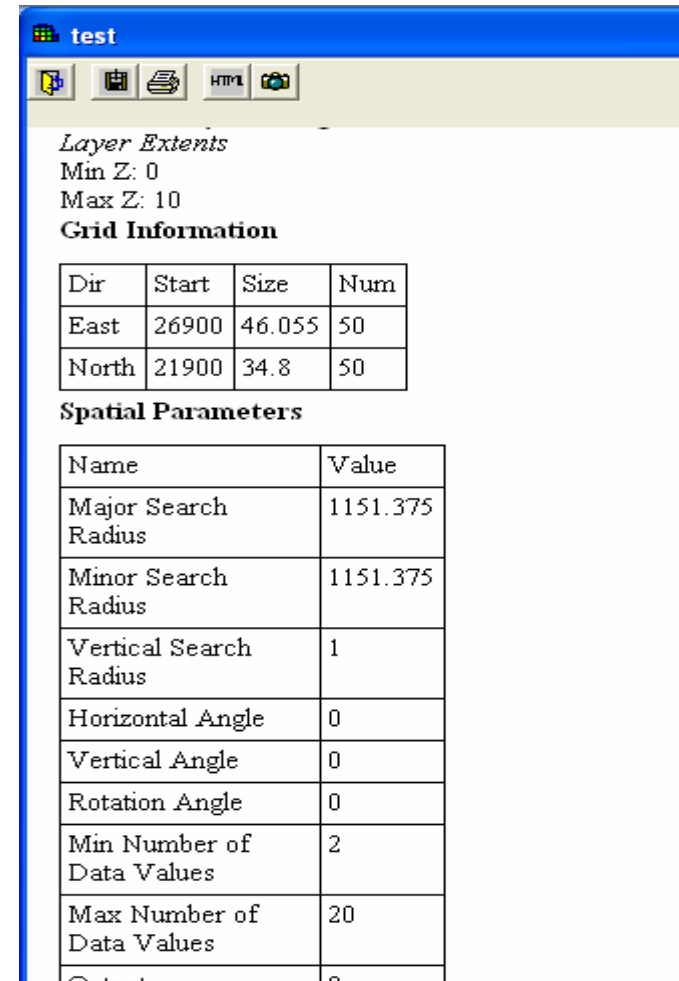
Change Active Report

Current Information to Add to the Report

- ☒ Picture
- ☒ GIS Files
- ☐ SHP Files
- ☒ Layering Design
- ☒ Grid Dimensions
- ☐ Variography Model Parameters
- ☐ Correlation Model Parameters
- ☒ Geospatial Parameters
- ☒ Decision Framework
- ☒ Human Health Risk Model
- ☐ Ecological Risk Model
- ☒ Decision Criteria
- ☐ ASCII Results
- ☒ Media Data

Step

Analysis of Model
Elements



test

Layer Extents
Min Z: 0
Max Z: 10

Grid Information

Dir	Start	Size	Num
East	26900	46.055	50
North	21900	34.8	50

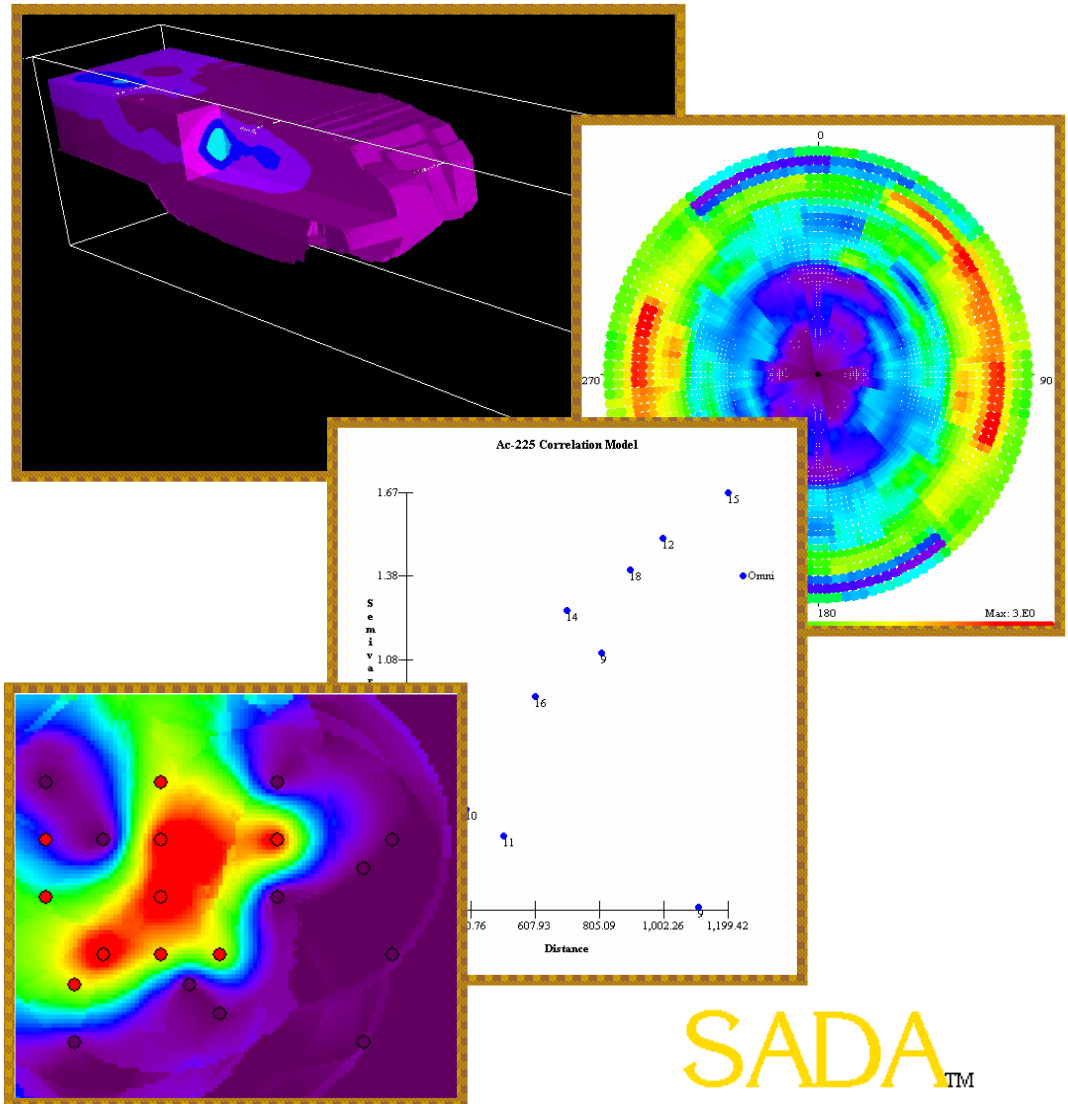
Spatial Parameters

Name	Value
Major Search Radius	1151.375
Minor Search Radius	1151.375
Vertical Search Radius	1
Horizontal Angle	0
Vertical Angle	0
Rotation Angle	0
Min Number of Data Values	2
Max Number of Data Values	20
Output	0

Documenation Output

Results Gallery

- Users can now save “static” results to the results gallery
- Users can view them, format them, and change various viewing properties
- Prevents users from having to regenerate a picture each time they want to see it
- Version 5.0 will allow dynamic results to be saved for further modeling



SADATM

Spatial Analysis and Decision Assistance

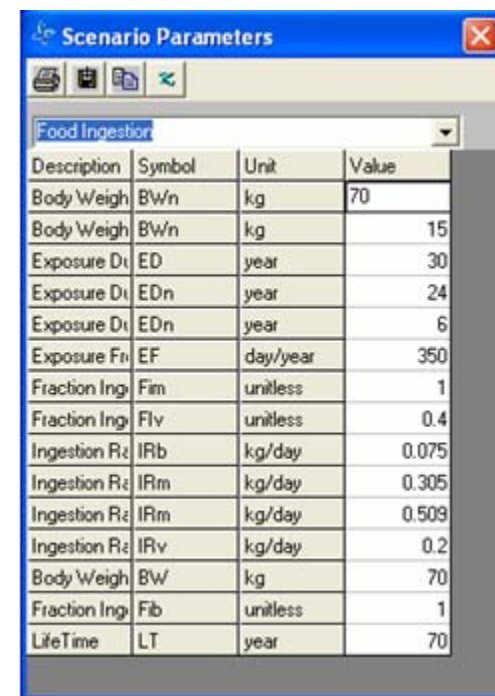
Human Health and Ecological Risk

- SADA implements EPA methods for conducting ecological and human health risk assessments
- Calculation of site-specific preliminary remediation goals
- Benchmark database for contaminant effects on ecological receptors
- Exposure modeling for humans and over 20 other terrestrial species
- Contains IRIS/HEAST toxicity databases for calculating risk from exposure
- Contains EPA default exposure parameters for the risk models
- Tabular screening and risk results
- Point screens
- Risk and dose mapping

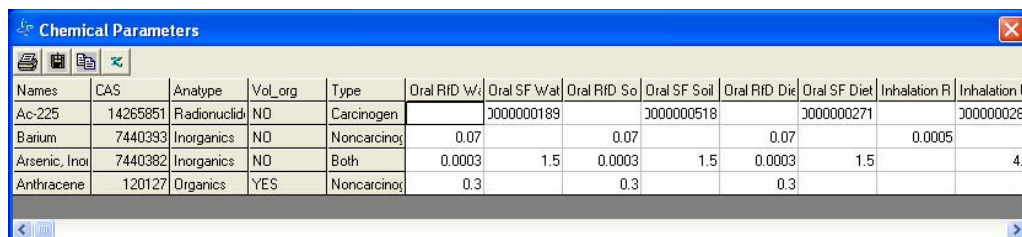


Human Health Risk Calculations

- For each media
 - Soil, Sediment, Surface Water, Groundwater
- Exposure Scenarios
 - Residential, Industrial, Recreational, Agricultural, Excavation
- Exposure Pathways
 - Ingestion, Inhalation, Dermal Contact, Food Chain (Beef, Milk, and Vegetable Ingestion)
- IRIS and HEAST Toxicity Databases for Carcinogenic and Noncarcinogenic Effects
- Physical Parameters for Modeling
 - Bioaccumulation Factors
 - Volatilization, Particulate Emission Factors
 - Permeability Constants, Absorption Factors
 - Saturation Coefficients, Radionuclide Half-Lives



Description	Symbol	Unit	Value
Body Weigh	BWn	kg	70
Body Weigh	BWn	kg	15
Exposure Di	ED	year	30
Exposure Di	EDn	year	24
Exposure Di	EDn	year	6
Exposure Fr	EF	day/year	350
Fraction Ing	Fim	unitless	1
Fraction Ing	Flv	unitless	0.4
Ingestion R	IRb	kg/day	0.075
Ingestion R	IRm	kg/day	0.305
Ingestion R	IRm	kg/day	0.509
Ingestion R	IRv	kg/day	0.2
Body Weigh	BW	kg	70
Fraction Ing	Fib	unitless	1
LifeTime	LT	year	70



Names	CAS	Anatype	Vol_org	Type	Oral RID W	Oral SF Wat	Oral RID So	Oral SF Soil	Oral RID Die	Oral SF Diet	Inhalation R	Inhalation L
Ac-225	14265851	Radionuclid	NO	Carcinogen		3000000189		3000000518		3000000271		3000000288
Barium	7440393	Inorganics	NO	Noncarcinog	0.07		0.07		0.07		0.0005	
Arsenic, Inot	7440382	Inorganics	NO	Both	0.0003	1.5	0.0003	1.5	0.0003	1.5		4.3
Anthracene	120127	Organics	YES	Noncarcinog	0.3		0.3		0.3			

SADATM

Spatial Analysis and Decision Assistance

Human Health Risk

- PRG Calculation
- PRG Screens
- Human Health Risk

Risk Based Screening Goals: Target risk = 0.0001...

Pathways:

☒ Ingestion ☐ Dermal ☐ Fish ☐ Beef ☒ All

☒ Inhalation ☐ External ☐ Vegetables ☐ Dairy

Rads and Nonrads/Soil/Residential/Carcinogenic

Name	CAS	Ingestion	Inhalation	All
Ac-225	14265851	5.6E+2	1.5E+5	5.6E+2
Arsenic, Inoi	7440382	4.3E+1	7.4E+4	4.3E+1

Human Health Risk Results

Pathways:

☒ Ingestion ☐ Dermal ☐ Fish ☐ Beef ☒ All

☒ Inhalation ☒ External ☐ Vegetables ☐ Dairy

Rads and Nonrads/Soil/Residential/Carcinogenic

Name	CAS	Ingestion	Inhalation	External	All
Ac-225	14265851	5.5E-7	2.E-9	5.5E-7	1.1E-6
Arsenic, Inoi	7440382	1.8E-5	1.E-8		1.8E-5
Total		1.8E-5	1.2E-8	5.5E-7	1.9E-5

Screening Results: Target risk = 0.0001/Target ...

Pathways:

☒ Ingestion ☐ Dermal ☐ Fish ☐ Beef ☒ All

☐ Inhalation ☐ External ☒ Vegetables ☐ Dairy

Rads and Nonrads/Soil/Residential/Carcinogenic

Name	CAS	Ingestion	Vegetables	All
Ac-225	14265851		Yes	Yes
Arsenic, Inoi	7440382		Yes	Yes

SADATM

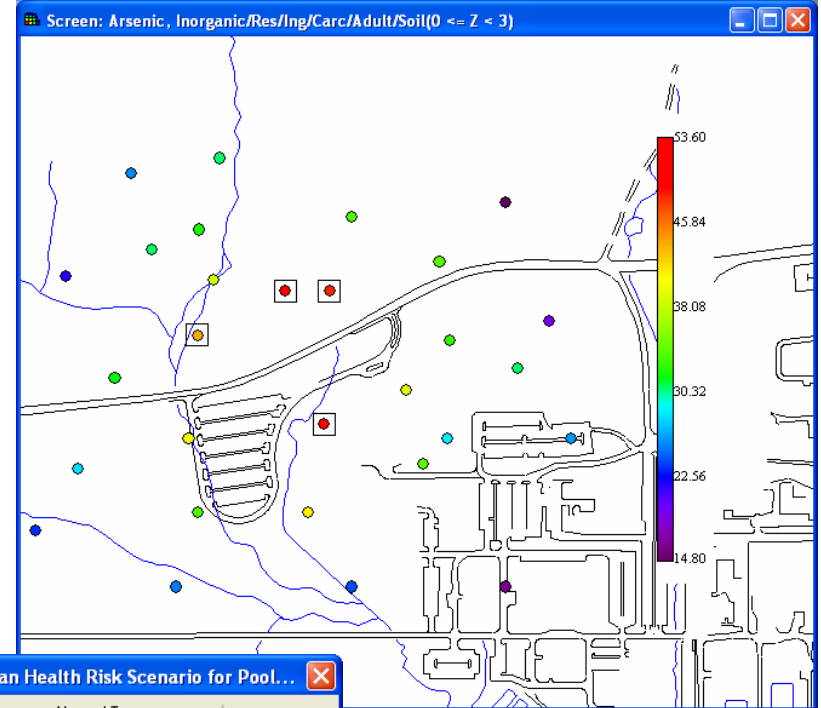
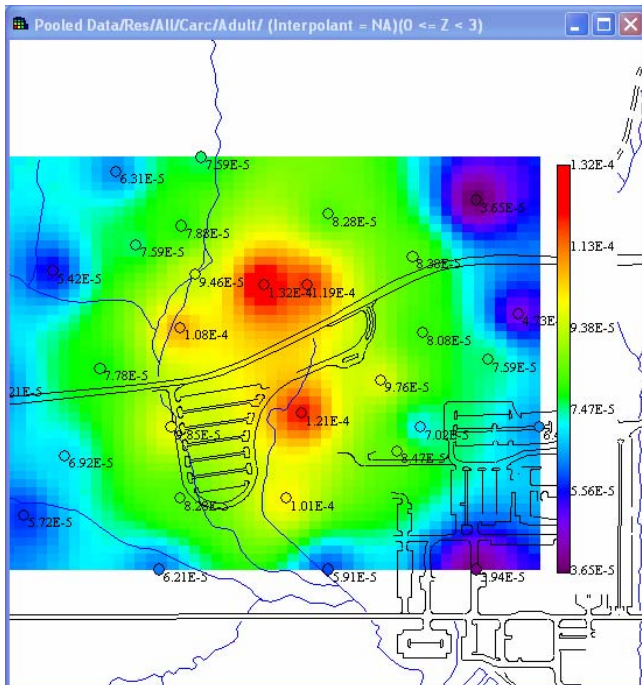
Spatial Analysis and Decision Assistance

Spatial Implementation of Risk Assessment

- Conventional Risk Assessment Limitations
 - Typically regulatory exposure assessment guidance recommends a summary statistic for the exposure concentration
 - Spatial information is lost when a summary statistic is used in the RA-exposure is assumed to be continuous in space and time
 - Often this lost info not recovered in the rest of the remediation process
- Reasons for incorporating spatial statistics into risk assessment
 - Maximize the use of limited resources
 - Efficiently collect data
 - Retain collected spatial info in the risk assessment
 - Use all types of available data, including expert judgment
 - To more adequately characterize the exposure distribution
 - Extrapolate from known data to cover data gaps
 - Account for spatial processes related to exposure
 - Better understand uncertainties in the exposure assessment

Human Health Spatial Risk Maps

- SADA calculates risk for each sampling point based on contaminant and exposure scenario
- Legend scale changes to risk



Human Health Risk Scenario for Pool...

Analyte	Nonrad Type	Age
<input type="radio"/> Rad	<input checked="" type="radio"/> Carcinogen	<input type="radio"/> Child
<input type="radio"/> Nonrad	<input type="radio"/> Noncarcinogen	<input checked="" type="radio"/> Adult
<input type="radio"/> Both		
Landuse	<input type="radio"/> Industrial	<input checked="" type="radio"/> Residential
<input type="radio"/> Agricultural	<input type="radio"/> Recreational	
<input type="radio"/> Excavation		
Pathway	<input type="radio"/> External	<input type="radio"/> Beef
<input checked="" type="radio"/> Ingestion	<input type="radio"/> Fish	<input type="radio"/> Milk
<input type="radio"/> Inhalation	<input type="radio"/> Vegetables	<input type="radio"/> Total
<input type="radio"/> Dermal		
Total Pathway Components	<input type="checkbox"/> External	<input type="checkbox"/> Beef
<input type="checkbox"/> Ingestion	<input type="checkbox"/> Fish	<input type="checkbox"/> Milk
<input type="checkbox"/> Inhalation	<input type="checkbox"/> Vegetables	
<input type="checkbox"/> Dermal		

OK

SADA™

Spatial Analysis and Decision Assistance

Ecological Risk

Ecological Risk Benchmarks

- Suitable for screening ERAs
- Compilation of ecological benchmarks for surface water, soil, and sediment
- Benchmarks a function of environmental variables where appropriate

Surface Water Ecological Benchmark Retrieval (mg/L)

Freshwater Benchmarks

☐ EC20 Daphnids ☐ LCV Daphnids

☒ EC20 Fish ☐ LCV Fish

☒ EC25 Bass Population ☐ LCV Non-Daphnid Inverts

☐ EC20 Sensitive Species ☒ NAWQC- Chronic

☐ EPA Region 4- Acute ☐ NAWQC- Acute

☒ EPA Region 4- Chronic ☒ Tier II SAV

☒ LCV Aquatic Plants ☐ Tier II SCV

Water Analysis Type

☒ Total

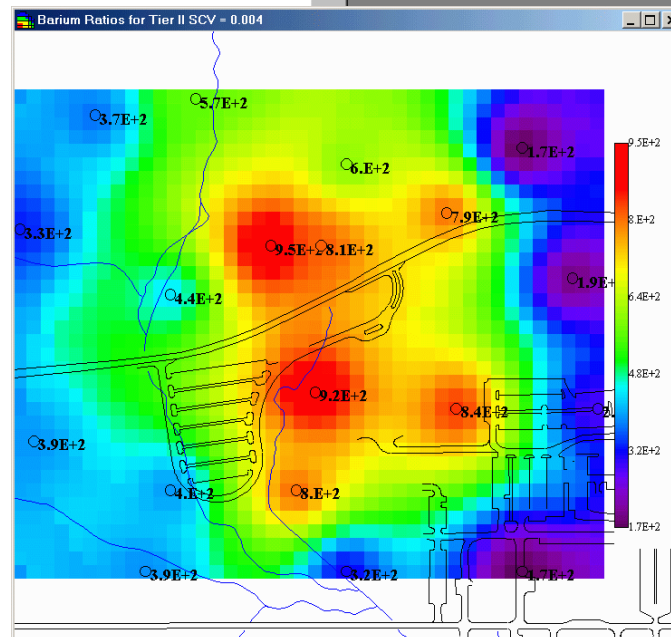
☐ Dissolved

Surface Water Constants

Hardness

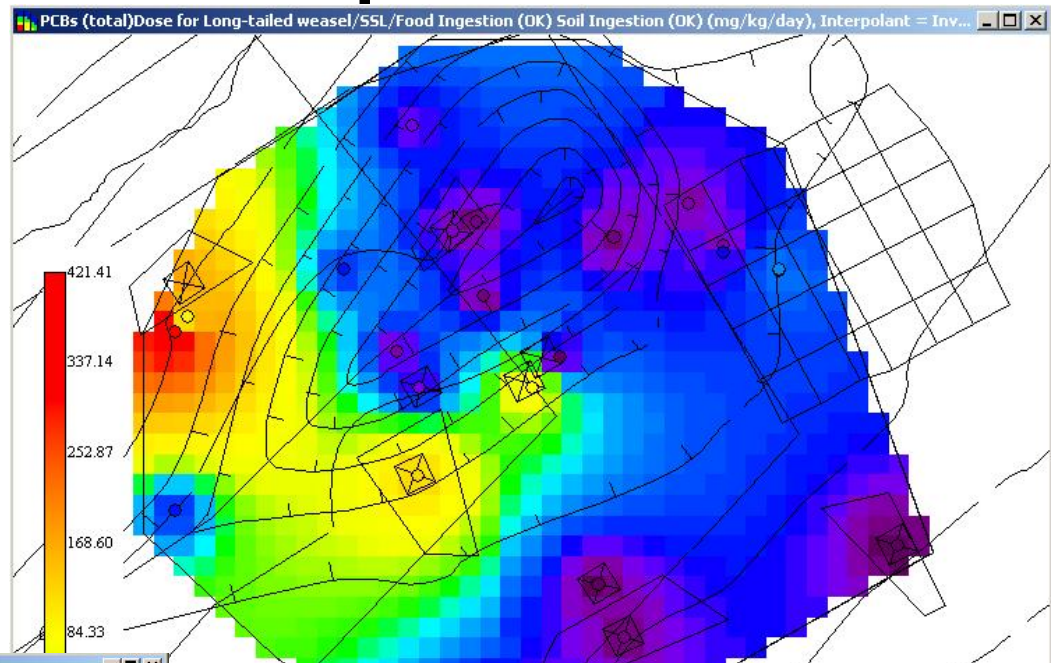
pH

Analyte	EPA Region 4- Chronic	NAWQC- Chronic	Tier II SAV	LCV Aquatic Plants	EC20 Fish	EC25 Bass Population
Dieldrin	0.0000019	0.000062				
Tetrachloromethane			4.4			
1,1,1-Trichloroethane	0.53		0.2	669	2.46	0.2
Carbon disulfide			0.02		5.72	1
Trichloroethylene (TCE)			0.44		5.76	0.2
Arsenic	0.19	0.19		2.32	2.13	2
Zinc	0.11	0.11		0.03	0.05	0.0



Ecological Dose Exposures


- SADA calculates dose (mg/kg BW d) from food ingestion, soil ingestion, dermal contact, and inhalation for terrestrial exposures
- SSL, Female, Male, or Juvenile
- Over 20 different species



Set Species-Specific Terrestrial Exposure Parameters

Select a species to view (and change if necessary) the default exposure parameters used in determining the daily contaminant dose received from exposure to soil.

Long-tailed weasel
Mustela frenata



Food Ingestion Parameters

Food ingestion rate: 0.1 kg dw / kg bw day

Fraction foliage: 0 0-1

Fraction seed: 0 0-1

Fraction invert: 0 0-1

Fraction mammal: 1 0-1

Mammalian Prey Diet

Fraction foliage: 0.485 0-1

Fraction seed: 0 0-1

Fraction invert: 0.485 0-1

Fraction soil: 0.029 0-1

Soil Ingestion Parameters

Soil ingestion: 0.039 fraction of food IR

Soil Inhalation Parameters

Inhalation Rate: 0.456 m³/day

Dermal Contact Parameters

Adherence Factor: 0.000001 kg/cm²

Surface area: 388 cm²

Physical Parameters

Body weight: 0.202 kg

Area usage factor: 1 fraction

Range:
Extending from just north of the United States-Canadian border through Central America to northern South America.

Save Changes Exit

Set Terrestrial Modeling Contaminant Parameters

PCBs (total)

Chemical Constants

Log Octanol/Water Partitioning Coefficient (Log Kow): 7.31 (mg/L)/(mg/L)

Inhalation

☒ Volatile ☐ Non-Volatile

Volatilization Factor (VF): m³/kg

Particulate Emission Factor (PEF): 1316239339 kg/m³

Soil -> Plant Concentration: Foliage

☐ Custom BAF (mg/kg)/(mg/kg)

☒ Kow-based BAF: 0.0313 (mg/kg)/(mg/kg)

☐ Tissue Regression

Log-linear slope

Log-linear intercept

Soil -> Plant Concentration: Seed

☐ Custom BAF (mg/kg)/(mg/kg)

☒ Kow-based BAF: 0.0313 (mg/kg)/(mg/kg)

☐ Tissue Regression

Log-linear slope

Log-linear intercept

Dermal Contact

Absorption Fraction: 0.06 mg/mg

Soil -> Invertebrate Concentration

☐ Custom BAF (mg/kg)/(mg/kg)

☒ Kow-based BAF: 33.4187 (mg/kg)/(mg/kg)

☐ Tissue Regression

Log-linear slope

Log-linear intercept

Soil -> Small Mammal Concentration

☐ Custom BAF (mg/kg)/(mg/kg)

☐ Tissue Regression

Log-linear slope

Log-linear intercept

Diet -> Small Mammal Concentration

☒ Custom BAF: 2.63 (mg/kg)/(mg/kg)

☐ Tissue Regression

Log-linear slope

Log-linear intercept

Save Changes Exit