
Kriging and conditional indicator simulation are valuable tools for evaluating uncertainty in the subsurface, but are limited by the assumption of stationarity (i.e. the mean and spatial variance are constant across the site). If the regions are distinct and unrelated, then zonal kriging can be accomplished manually by modeling each region and merging the results into a single model. However, merging results is expensive in human resources and computer processing time, and merged results cannot represent gradational transitions. A technique called zonal kriging was developed and is presented in this chapter so that different spatial equations can be applied to separate regions of a site. This zonal kriging algorithm is applied to a synthetic data set, to data from an extensively sampled outcrop in Yorkshire, England, and to a subsurface site at the Colorado School of Mines (CSM) survey field, in Golden, Colorado. Estimation of the synthetic data set demonstrates the advantages and shortcomings of the technique, and conditional multiple indicator simulations of both the Yorkshire outcrop and the CSM survey field data sets illustrate the improvement attained through use of zonal kriging.

5.1: Introduction and Previous Work

Significant variation of spatial statistics across a site can violate the basic assumptions of stationarity and this can lead to strongly biased estimates. Depending on the magnitude of the deviation from stationarity and the importance of the results, two approaches are often taken. One assumes the problem can be controlled with the local stationarity of the neighboring data samples, and a spatial model which reflects the mean behavior of the entire site. The other divides the area into an appropriate number of zones, describes the spatial statistics for each zone (Figure 5.1), estimates each zone, and merges the results. One problem with the second method is that the boundary between zones is often abrupt. The second method may be appropriate where the contact is a fault or an unconformity, but the results are unsatisfactory for sites with gradational transitions.

One alternative approach, that can be applied in cases with gradational boundaries is to transform all the points in the data set to match the spatial statistics of the cell currently being estimated in

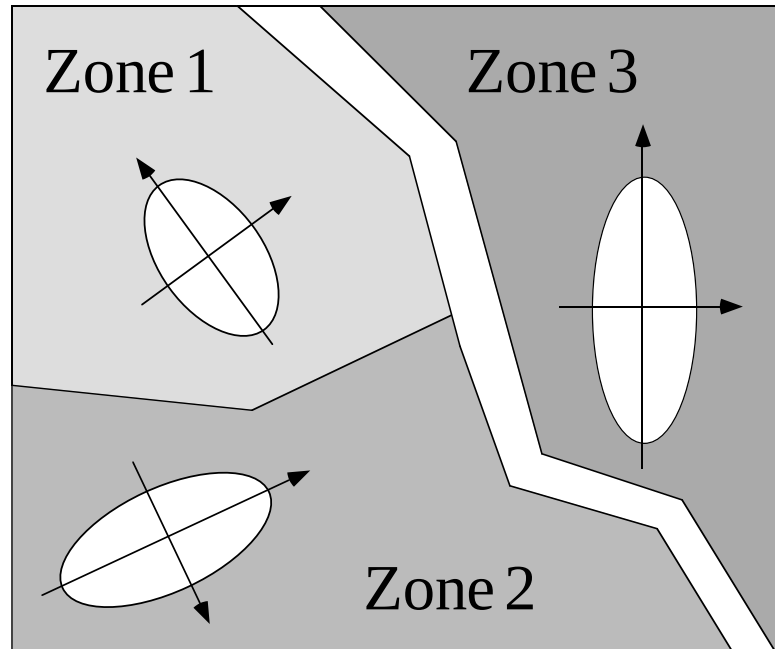


FIGURE 5-1. Spatial statistics may vary across a site, such that a single semivariogram model may not be appropriate for the entire site.

which case all the site data are considered, whether they are from the zone of interest or not. This method eliminates the problem of sharp zone boundaries, and addresses the possible gradation of properties between zones, and eliminates the need to manually merge individual zones into a single model. However, it does not accommodate sharp boundaries, nor does it recognize that some points from neighboring zones may have no bearing on the estimated value, even though they are in the transformed search neighborhood.

The approach presented here has the advantage of both of the zoning techniques described above and adds utilities to define inter-zone relationships. Such relationships describe how data, located in one zone are treated when sampled for a cell calculation located in another zone. This technique is applied using Simple (SK) and Ordinary Kriging (OK), and Multiple Conditional Indicator Simulation (MCIS). MCIS is used in two ways in this chapter. It is first used to define the zonation boundaries. MCIS can generate multiple, unique, realizations of zone boundaries, which honor the statistics of the data. This is a useful technique when the data are limited. This makes the simulation a two step process; first the zone boundaries are defined using discrete MCIS, then the interior of each zone is estimated using SK or OK. The second use of MCIS, is to populate the zones (predefined with some other method) with indicator based parameter estimates. MCIS can be used to generate zone boundaries; and MCIS, SK, or OK can be used to estimate parameter variations within the zones.

5.2: Methodology

Existing kriging algorithms (ktb3dm ; and SISIM3D) were modified to implement zonal kriging. Both codes have the required mathematical tools, but the calculation sequence was reordered, additional input describing zones and their transitional character was defined, and the search algorithm was modified. The standard kriging algorithm and modifications for zonal kriging (*italicized steps*) are shown in a flowchart in Figure 5.2. The key new aspects that have been added

Standard Kriging Algorithm with Zonal Kriging Modifications

(*Additional Zonal Kriging tasks in italics*).

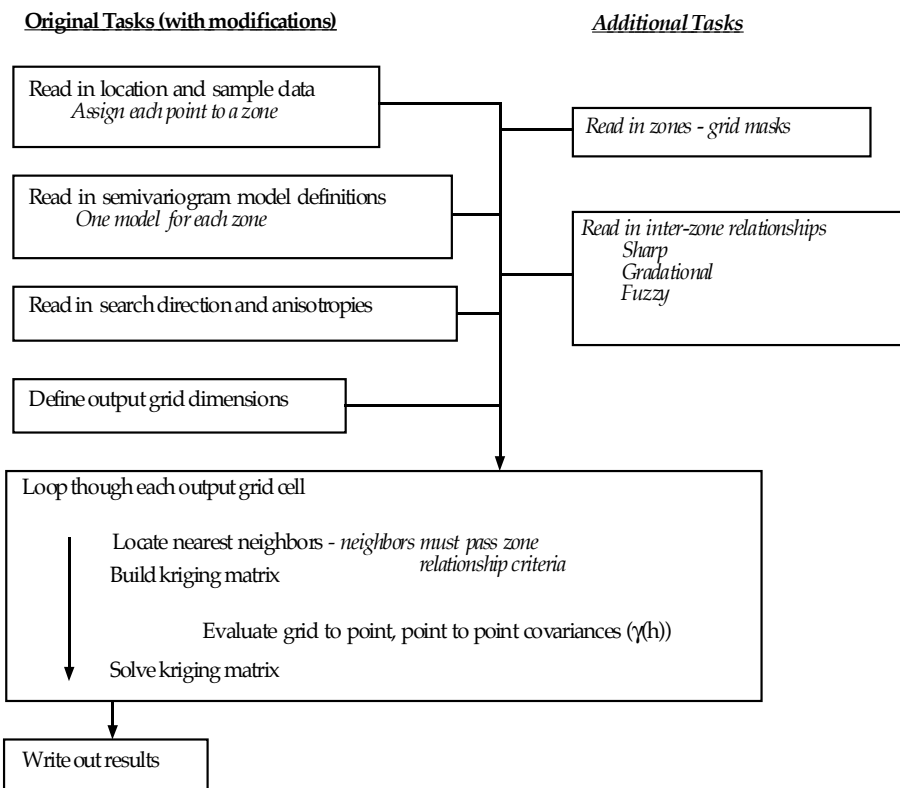


FIGURE 5-2. Basic steps involved in the standard kriging algorithm with additional steps needed to implement Zonal Kriging indicated in italics.

to the previous algorithms are:

- 1) Defining zones: Defining zones is the most arbitrary portion of the process. Choosing the location of boundaries between zones is subjective, particularly when the data are

sparse. In the synthetic example described below, a realization from a MCIS is used to define the zonation. Repeating the process using zones from a series of realizations, addresses much of the uncertainty associated with the location of the boundaries (although this is not done in this chapter). In the Yorkshire, England example MCIS is not used to define zones because there are sufficient data to determine zones by stratigraphic interpretation. At the CSM survey field, the zonal boundary is defined along the contact between two geologic formations.

- 2) Defining inter-zone relationships. Inter-zone relationships may be *Sharp* (the zones are completely unrelated), *Gradational* (one zone merges infinitely into the other), or *Fuzzy* (the zones are gradational over a limited distance and then are distinct). If the inter-zone relationship is *Fuzzy* (this term does not refer to fuzzy logic), the width of the boundary must also be defined. The rationale behind each type of transition is as follows:

Sharp: In many cases, two units are in contact with one another (Figure 5.3a), but otherwise are unrelated. Examples are faults and geologic unconformities. In this situation, it is not appropriate to use data from one zone to estimate the spatial distribution in the other.

Gradational: In some environments, units grade into one another (Figure 5.3b). This is typical of coastal deposits where beach sands grade into marine clays and shales. In each environment, the depositional systems are very different, but the change is gradational. Selecting this option requires the assumption that the region being estimated lies fully within the gradational region.

Fuzzy: Fuzzy inter-zone relationships are similar to the Gradational relationship, however the transitional distance is limited in extent (Figure 5.3c). Beyond a defined distance, data from the other zone is no longer correlated to the location being considered. The width of this boundary is subjective, as it is not necessarily related to the range of the semivariogram of either zone. Defining the width of the zone is left to the modeler, and is based on their experience and knowledge of site conditions. The gradational method, is a special case of the Fuzzy method with an infinite boundary width.

Once the user has defined the zones and the inter-zone relationships, the algorithm:

- 3) assigns each data point to a zone.
- 4) creates a mask for each zone, describing how each cell within the grid will be treated.

Once these prerequisite details are defined, each grid cell is evaluated. When estimating a grid cell value, the modified programs use the properties associated with the zone in which that cell lies. This includes search criteria and semivariogram model information. The algorithm then:

- 5) finds the nearest neighboring data points: Neighboring points may be selected in several ways depending on the zone inter-relationship. If the points are in the same zone they are treated normally. If the boundary is *sharp*, no points from across the

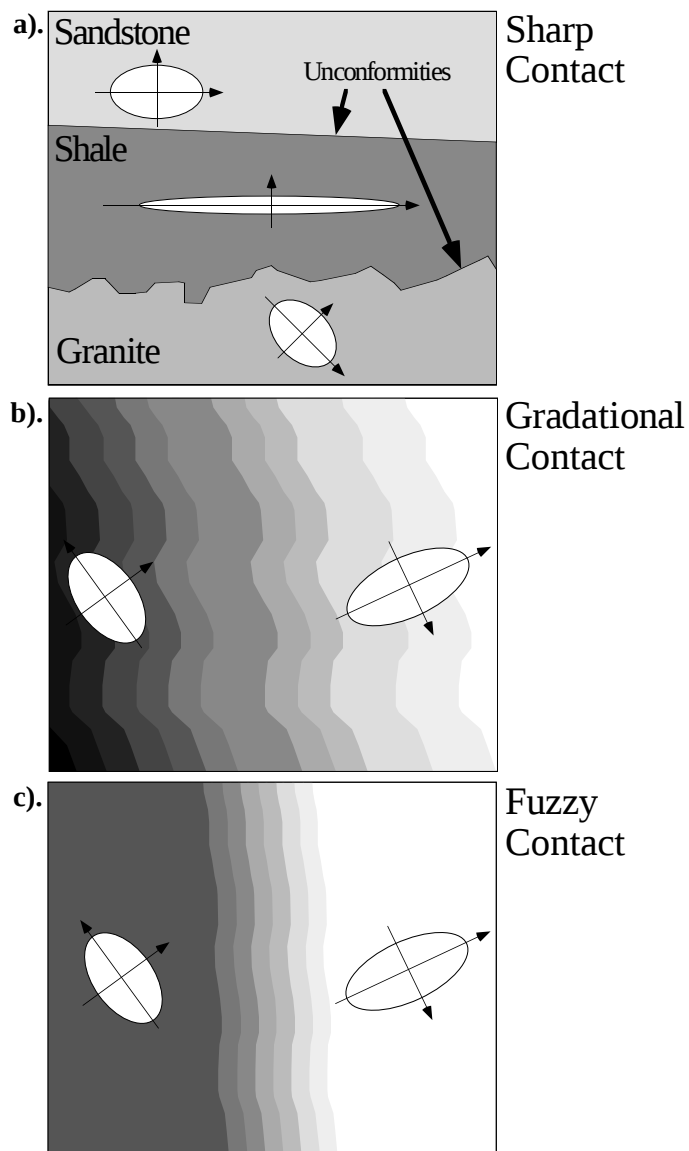


FIGURE 5-3. Different methods for describing zone contacts: a) sharp, b) gradational, and c) fuzzy.

boundary will be used (Figure 5.4a). If the boundary is *gradational*, the nearest points will be used regardless of the zone they belong to (Figure 5.4b). If the transition is fuzzy, points can be selected from the neighboring zone, but only to a limited distance (Figure 5.4c), and all other points are ignored.

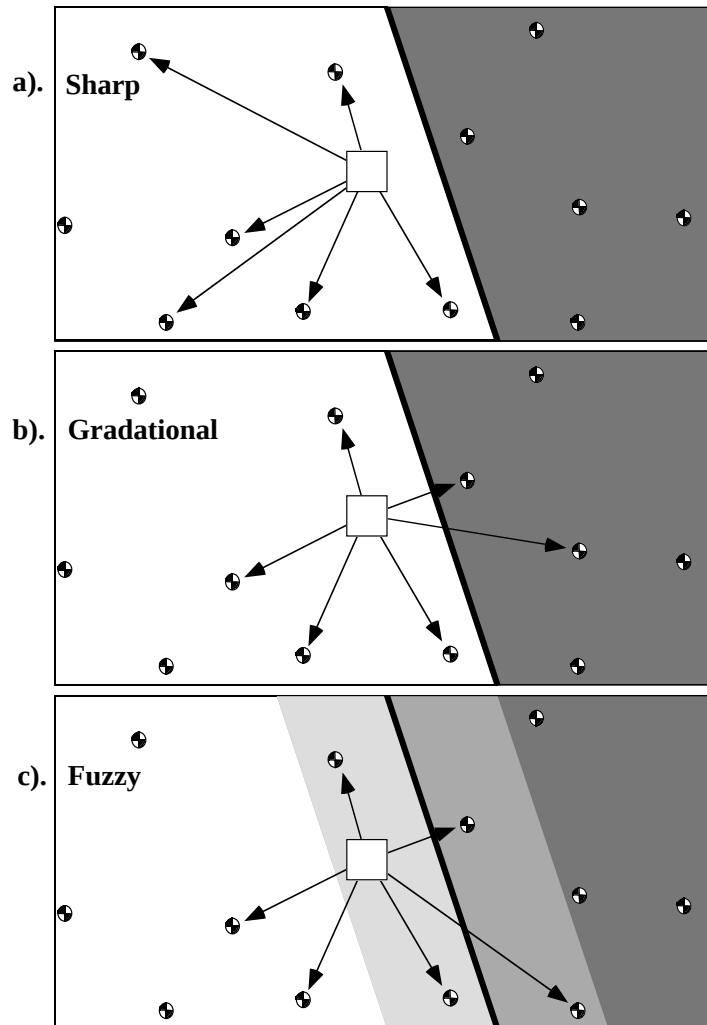


FIGURE 5-4. The search for nearest neighbors varies with zone boundary type: a) sharp, b) gradational, and c) fuzzy.

- 6) generates and solves the kriging matrix: Once the nearest neighbor points have been selected, the kriging algorithm proceeds, evaluating all points using the semivariogram model from the zone of the cell being estimated. This is regardless of which zone each point is from, or how deep the point is into the neighboring zone. The information describing the semivariogram from the neighboring zone is ignored because merging the models may lead to a matrix that is not positive definite, a basic requirement for kriging.

5.3: Examples

Three examples are used to demonstrate the utility of zonal kriging. The first is based on a small sample of eleven synthetic data points. This example demonstrates 1) the differences between SK with and without zoning; 2) how different inter-zone relationships can effect results; and 3) some of the shortcomings of zonal kriging. The second example applies zonal kriging and indicator simulation to an extensively sampled fluvio-deltaic outcrop in Yorkshire, England. This outcrop exhibits two zones with different spatial characteristics, and was "sampled" on seven vertical lines representing bore-holes, thus allowing the comparison of simulations based on a small sample of points, with the full, known section. This example demonstrates that dividing the cross-section into two zones yields more realistic realizations as compared with modeling the site using a single zone. The final example uses a data set from the Colorado School of Mines survey field, Golden, Colorado. Zonal kriging is applied to field data in combination with techniques described in earlier chapters (directional semivariograms, class indicators).

5.3.1: Synthetic Data Set Example

A simple, synthetic, two-dimensional data set of porosity's (%) with eleven data points (Figure 5.5a) was evaluated using SK (Figure 5.5b). If the spatial statistics of the site are relatively consistent, this may be a good interpretation. However, if the data reflect material properties from three distinctly different areas, where the spatial statistics are substantially different, another interpretation is needed (three zones is excessive given the amount of data, but is useful for this demonstration). Given a map of the material zones (Figure 5.5c is one possible zonation realization created using MCIS), the site can be modeled using one of several assumptions: 1) the zones are completely unrelated; 2) there is an infinite gradation between zones; or 3) there is a short distance over which the zones are gradational. For these examples, the following isotropic semivariogram models were used:

	Single Zone	Zone 1	Zone 2	Zone 3
Model Type	Spherical	Spherical	Spherical	Spherical
Range	150 m	150 m	175 m	200 m
C₁	0.24	0.24	0.22	0.22
Nugget	0.01	0.01	0.03	0.03

Although these models are not dramatically different, some interesting results are obtained.

Simple kriging with sharp, non-gradational contacts yields the map shown in Figure 5.5d. In this model, Zone 1 (Figure 5.5c - blue) and Zone 3 (red) are gradational, but Zone 2 (green) is completely unrelated. This yields sharp transitions between Zone 2 and the other two zones. Within each zone though, the surfaces are smooth as would be expected with SK. The sharp transitions match the zone definition shown in Figure 5.5c.

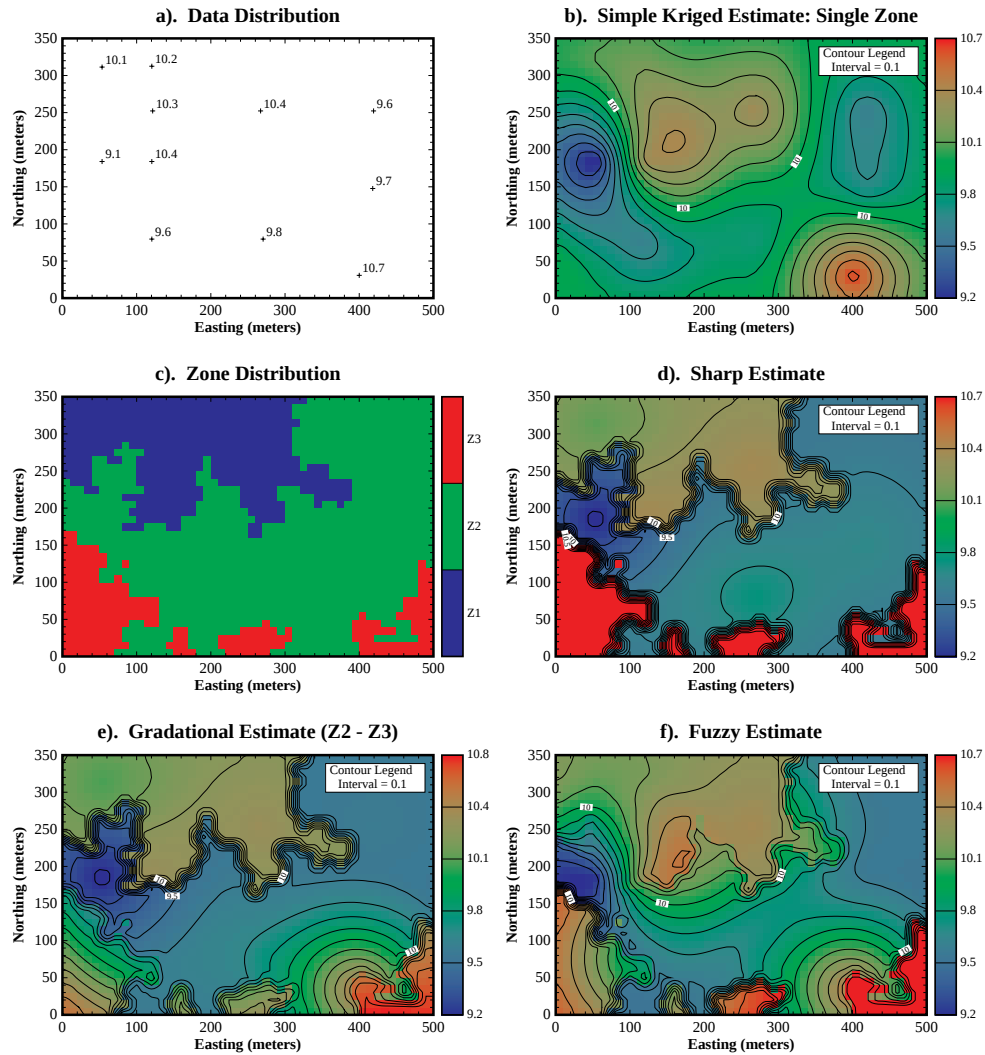


FIGURE 5-5. Different forms of ordinary and zonal kriging. (a) Sample data, (b) a traditional Simple Kriged map, (c) one possible zone map from a conditional indicator simulation, (d) sharp transition, (e) gradational transition, (f) fuzzy transition.

A gradational contact between adjacent units produces a different map. In this example (Figure 5.5e), Zones 1 and 3 are gradational, and the contact between Zones 2 and 3 is also defined as gradational (the contact between Zone 1 and 2 is sharp). The contour lines in Zone 1 are unchanged from Figure 5.5d, but there is substantial smoothing between Zones 2 and 3, although it is not complete, and the boundaries are somewhat abrupt. This incomplete smoothing occurs because the

data control for neighboring cells in different zones is coupled with different semivariogram models, and model ranges near the sample spacing of the data. These rough transitions will disappear with finer sampling.

The final option is explored by defining transitional (fuzzy) contacts of finite thickness. In this example, Zone 1 was defined to have a fuzzy boundary of 20 meters with Zone 2, and Zone 3 was defined to have a fuzzy boundary of 40 meters with Zone 2. Although there is substantial smoothing, each zone maintains much of its own character (Figure 5.5f), and the map is still substantially different than the map generated using single zone SK (Figure 5.5b). In some of the fuzzy boundary zones, particularly near the southern map border (Zone 2 vs. Zone 3), the contacts are still quite abrupt. This is, as noted for the gradational method, due to lack of data within the model range.

Which of the above models best represents the synthetic site is not an issue for this discussion. What is important, is that the modeler can control zonal differences and inter-relationships as appropriate for the site under evaluation, thus providing additional flexibility. Neither SK, nor OK could generate anything other than the first map (Figure 5.5b) or the second map if results were merged manually (Figure 5.5d).

As shown in Figures 5.5e and 5.5f, gradational and fuzzy boundaries can be abrupt. This is a numerical phenomenon, not a physical feature. The problem occurs when the set of nearest neighboring points are substantially and consistently different. In this model, this occurs because sample spacings are close to the range of the zonal semivariogram models. If the site was modeled with 1) more data points, or 2) with longer semivariogram model ranges, the contacts would be less abrupt.

5.3.2: Yorkshire, England Example

An outcrop cross-section from a fluvio-deltaic deposit near Yorkshire, England was sampled on a 17 x 20 cm grid spacing. For practical reasons (computation time) the data were upscaled to 2m x 1m grid blocks. The full 600m x 30m cross-section is shown in Figure 5.6a, and is composed of three materials: shale (SH), shaley-sandstone (SH-SS), and sandstone (SS).

To demonstrate the advantages of zonal kriging, the cross-section was "sampled" with seven vertical lines representing wells (Figure 5.6b) (1m vertical samples). Two distinct zones exist: a lower zone with high continuity in the SS and SH-SS units, and an upper zone dominated by SH, with small lenses of SS and SH-SS. To confirm that the spatial statistics were indeed different, the original section was divided into two zones (Figure 5.6d) based on a fence diagram of the bore-hole data. The facies frequencies and semivariograms for the entire cross-section were compared with those from the two zones. Whether examining the data from the exhaustive data set of the full

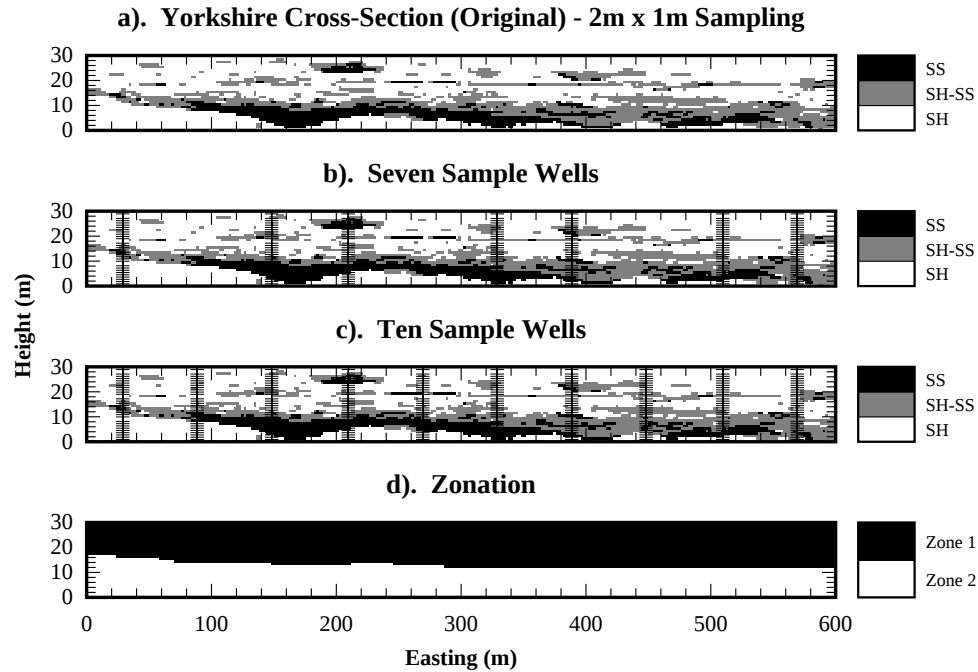


FIGURE 5-6. Model definition information for the Yorkshire cross-section: a) actual cross-section sampled with 2m x 1m cells; the locations of b) 7 and c) 10 “well samples;” d) zone definition.

cross-section, or the well samples, the results are similar. About 61% of the samples are SH, 24% SH-SS, and 15% SS. When the individual zones are separated, the distributions are very different:

	SH (%)	SH-SS (%)	SS (%)
Cross-Section (All)	63.4	23.6	13.0
Wells (All)	59.0	24.8	16.2
Cross-Section (Top)	81.2	16.4	2.4
Cross-Section (Bottom)	40.2	32.9	26.9
Wells (Top)	80.7	16.8	3.5
Wells (Bottom)	37.2	32.6	30.2

The top zone contains more than twice as much SH as the bottom zone and almost no SS, whereas the materials are more evenly distributed in the bottom zone. When semivariograms are calculated for the entire section, and for the top and bottom zones, using both the full cross-section and the well data, the zonation is again apparent. The horizontal and vertical indicator semivariograms are shown in Figures 5.7 through 5.10. The spatial statistics of the top zone are clearly different than

Exhaustive Yorkshire Cross-Section Indicator Semivariograms: SH/SH-SS Threshold

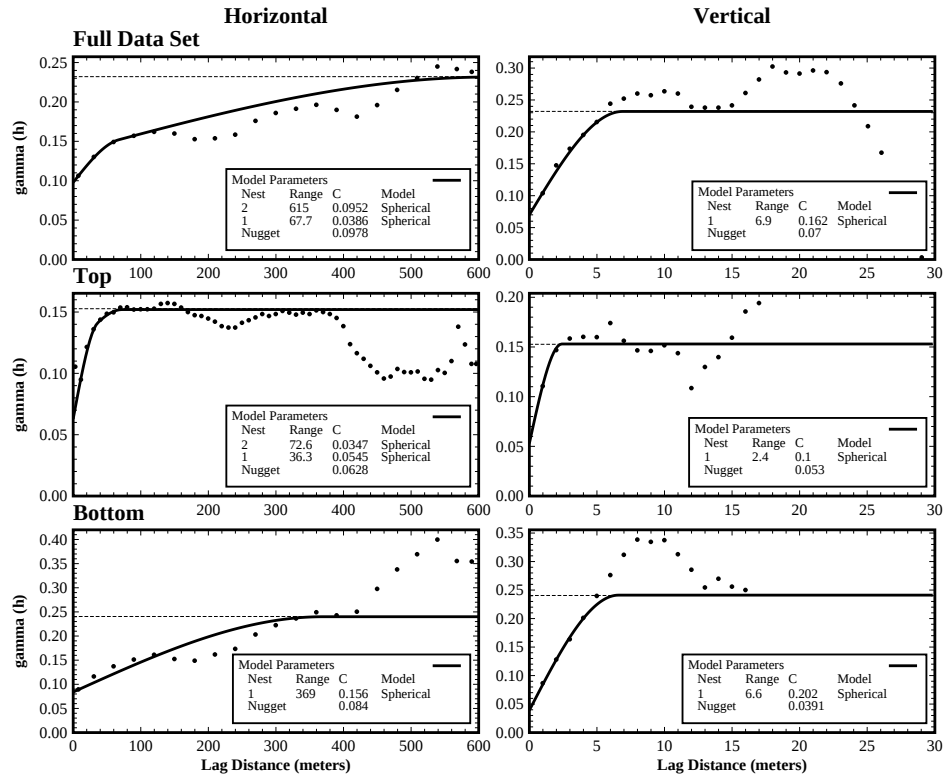


FIGURE 5-7. Exhaustive experimental and model indicator semivariograms for SH/SH-SS threshold (full cross-section) of the Yorkshire data set.

those of the bottom zone, and those of the entire cross-section. The horizontal range for the SH/SH-SS threshold in the top zone is only about 15% to 25% that of the bottom zone (Figures 5.7 and 5.9). For the SH-SS/SS threshold, the horizontal range in zone 1 is about 35% of that in zone 2 (Figures 5.8 and 5.10).

The assumption of stationarity is not applicable to this cross-section, thus modeling this site using a single set of semivariograms is statistically inappropriate. Two ensembles of 100 realizations each, demonstrate that zonal kriging yields more accurate results than a single semivariogram model. The first ensemble is based on the assumption that stationarity is valid, and only one semivariogram model set is required (the full well data set semivariogram models, Figures 5.9 and 5.10). The second set is based on the observation that stationarity is violated between zones, but is valid within each zone, thus a sharp transition is assumed, and the "Top" and "Bottom" semivariogram models (Figures 5.9 and 5.10) are used. The ranges of both horizontal and vertical semivariograms, in the top zone are much shorter than the ranges for the full section and the bottom zone. The model grid

Exhaustive Yorkshire Cross-Section Indicator Semivariograms: SH-SS/SS Threshold

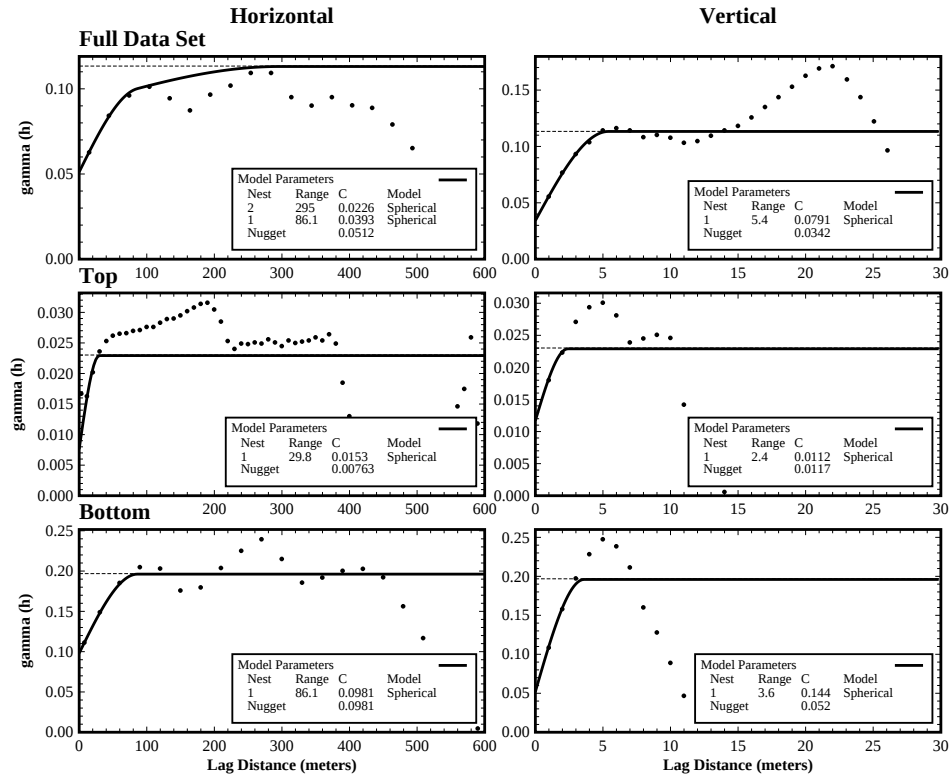


FIGURE 5-8. Exhaustive experimental and Model indicator semivariograms for SH-SS/SS threshold (full cross-section) of the Yorkshire data set.

matches the original cross-section grid exactly; with cells 2m x 1m in 300 columns and 30 layers. Results of several realizations using one and two zones respectively, are shown in Figures 5.11 and 5.12. Differences between the two sets are subtle, but discernible in the top zone. There are differences in the bottom zone, but they are minor. To evaluate the results, realization pairs were compared in order (Realization #1 [1 Zone] vs. #1 [2 Zones], #2 [1 Zone] vs. #2 [2 Zones], etc.) because matched realizations use the same "random" search path to evaluate the grid, and start using the same random seed. Therefore differences are only attributed to the use of zoning, and the realization with the smallest number of misclassifications, is defined to be the more accurate model. The number of misclassifications was calculated by subtracting the actual grid value from the estimated grid value at each grid location; any cell with a non-zero error was misclassified.

In the four realizations shown (using either method), there is a largely continuous SH-SS/SS unit in the bottom zone, which is present in the actual section, though there are random fluctuations in the realizations. The random fluctuations in the realizations are due to the large nugget terms (up to

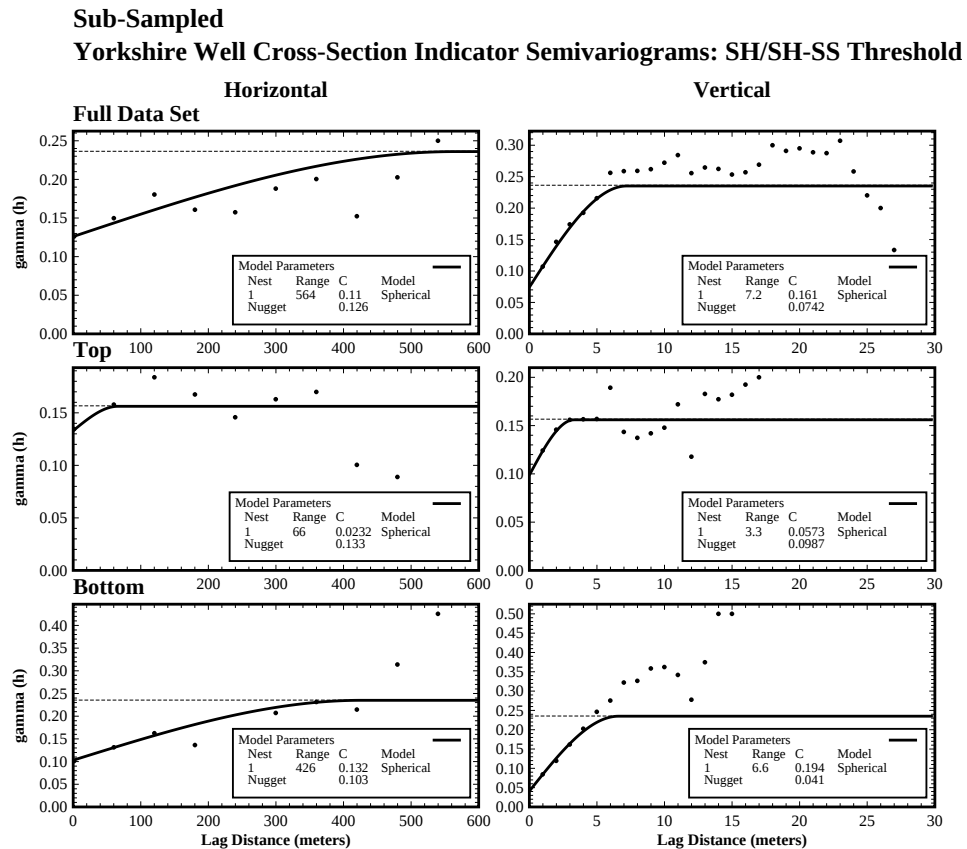


FIGURE 5-9. Sub-sampled experimental and Model indicator semivariograms for SH/SH-SS threshold (well data) of the Yorkshire data set.

64% of variance) in the model semivariograms. The difference between the lower portions of the two series of simulations is minor, because the global semivariogram models are fairly similar to the semivariogram models for the lower zone.

There are greater differences between the one and two zone simulations in the top zone. Both exhibit a great deal of randomness due to large nuggets, but the results of using two zones create more solid lenses without as many small isolated cells (compare the single and two-zone versions of realization #66, this was one of the most accurate models, and #22, one of the least accurate). The single zone model tends to create long, thin layers with considerable scatter, as compared to the more connected, but less laterally continuous units in the upper section of the two zone realization.

One-hundred pairs of realizations (one zone and two zone) were generated using the same random number sequence, and random path through the model grid. These pairs were compared to the actual cross-section. Realizations generated with the two-zone approach had fewer (10 to 328; both

**Sub-Sampled
Yorkshire Well Cross-Section Indicator Semivariograms: SH/SH-SS Threshold**

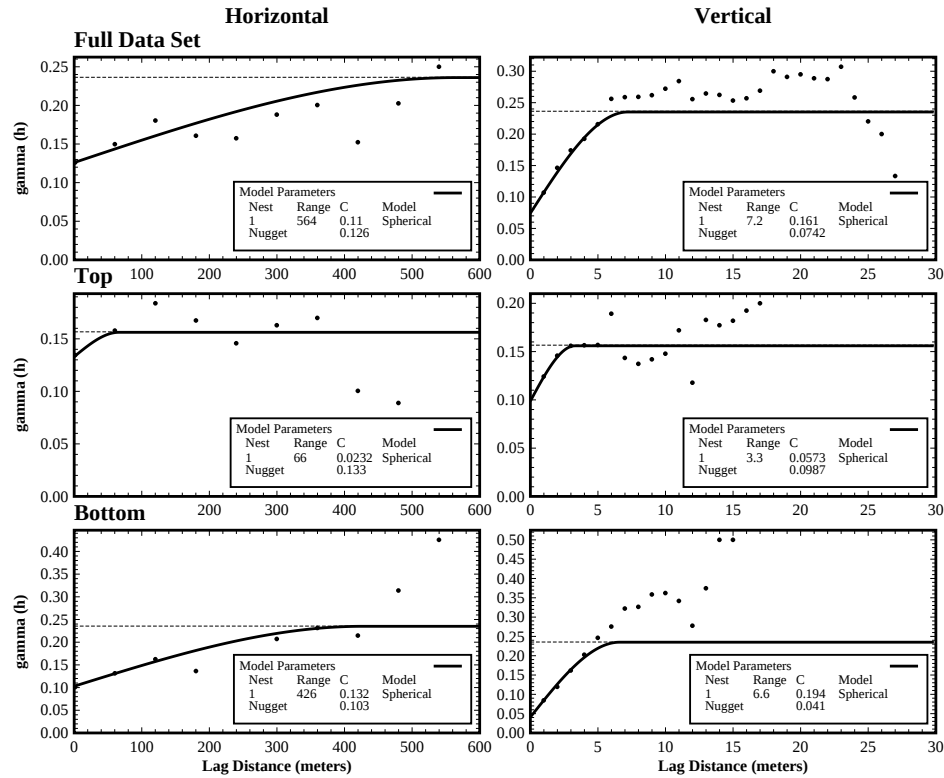


FIGURE 5-10. Sub-sampled experimental and Model indicator semivariograms for SH-SS/SS threshold (well data) of the Yorkshire data set.

methods consistently, correctly, identified about 5700 of the 9000 cells) misclassifications in 80 of the 100 pairs. When ten equally spaced wells (Figure 5.6c) were used with the same semivariogram models, instead of the original seven wells (Figure 5.6b), realizations generated with the two zone approach had fewer misclassifications in 91 of the 100 pairs (new semivariogram models may have improved results even more). The reduced number of misclassifications indicate that modeling the site with two zones yields better results. However more work is required to set up the model (additional semivariogram calculations, data entry, zone definition).

In addition to improved accuracy, the modeler has more flexibility in fine tuning the model solutions for each zone. The results can be dramatic in one zone without affecting the other. A sample two-zone realization is shown in Figure 5.13a. Realizations resulting from independently decreasing the nuggets in the top and bottom zones are shown in Figures 5.13b and 5.13c respectively. In both cases, the continuity is increased and the random scatter is reduced, without

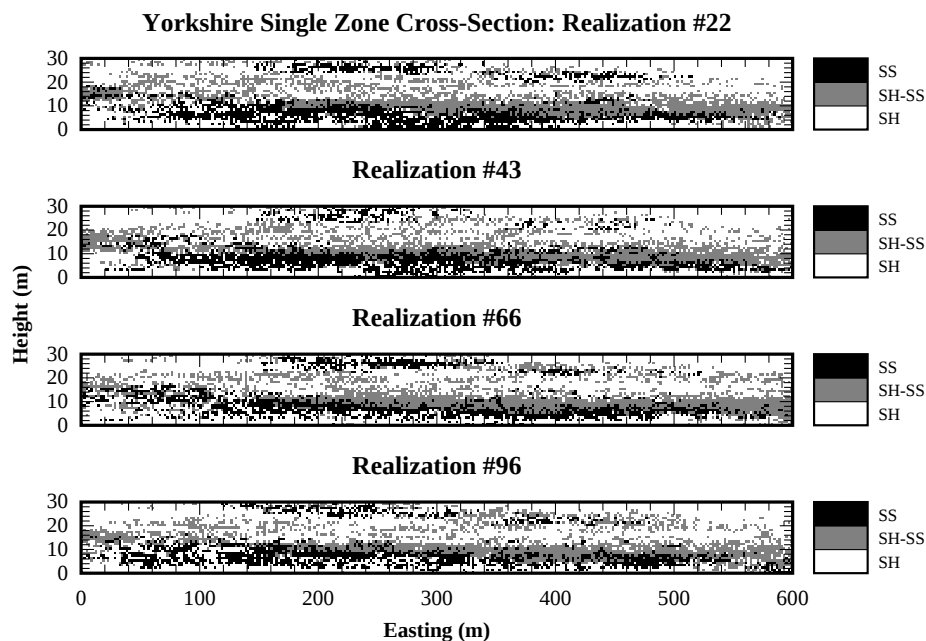


FIGURE 5-11. Realizations from single-zone simulation series.

affecting the alternate zone. If these same changes were made on a single zone realization, they would effect the entire grid, and improvement in one portion of the grid would have to be balanced against the degradation of the other zone.

5.3.3: Colorado School of Mines Survey Field Example

Unlike the Yorkshire, England example, field geology is rarely known in such detail. Usually only a limited amount of data are available at a site, typically far less than would be desired. At the CSM survey field, located on the West edge of Golden, Colorado (Figure 5.14), hard and soft data were collected to investigate the use of soft data for reducing uncertainty associated with ground water flow models. The site contains core and chip data from eighteen boreholes; borehole geophysical logs; and eight (Figure 5.15 and 5.16), two-dimensional, cross-hole, tomographic sections. Even though the site crosses two geologic formations, the site was initially simulated as a single zone. This was done because a zonal simulation tool was not available. In this study, zonal kriging is used in combination with directional semivariograms and class indicator simulation to model the CSM survey field.

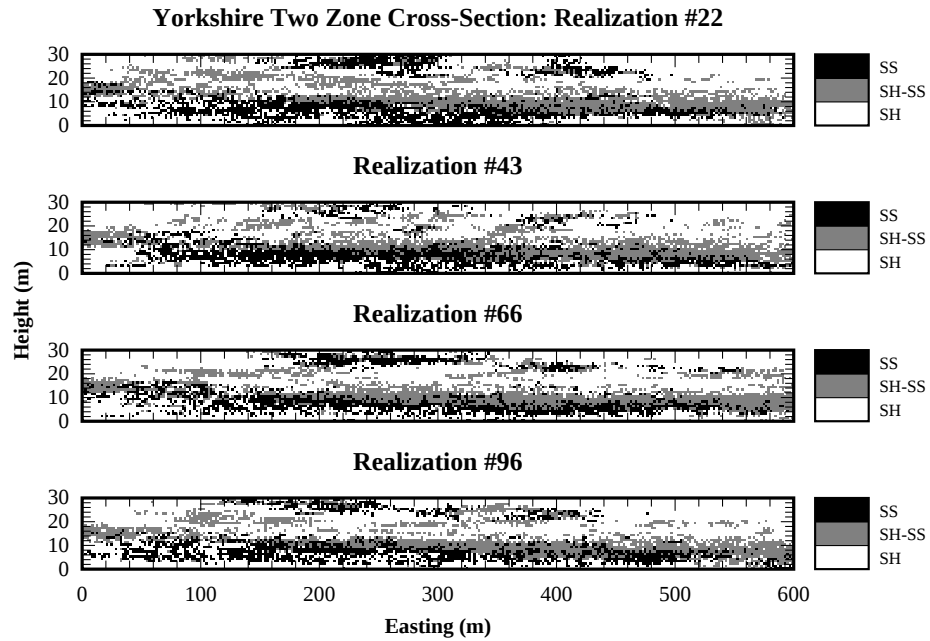


FIGURE 5-12. Realizations from two-zone simulation series.

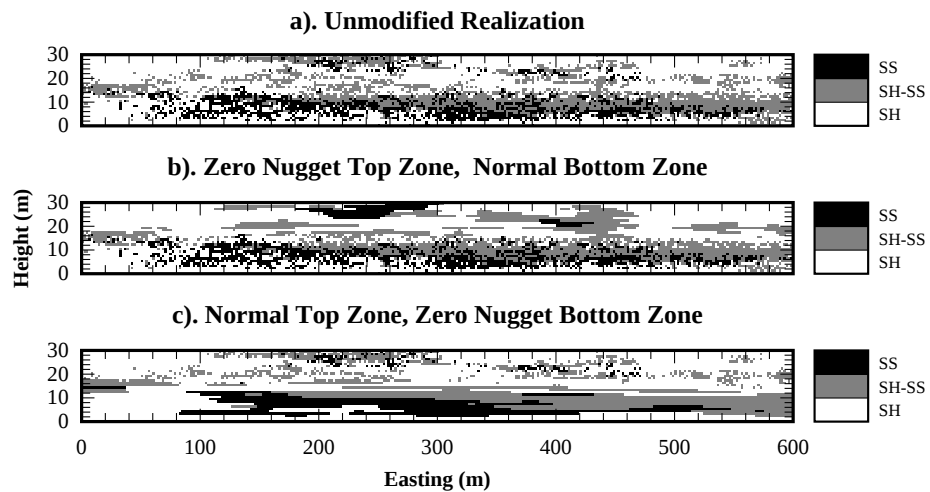


FIGURE 5-13. Impact of altering the semivariogram nugget independently in the top and bottom zones of a section: a) original simulation section; b) reduced nugget in top zone; c) reduced nugget in bottom zone.

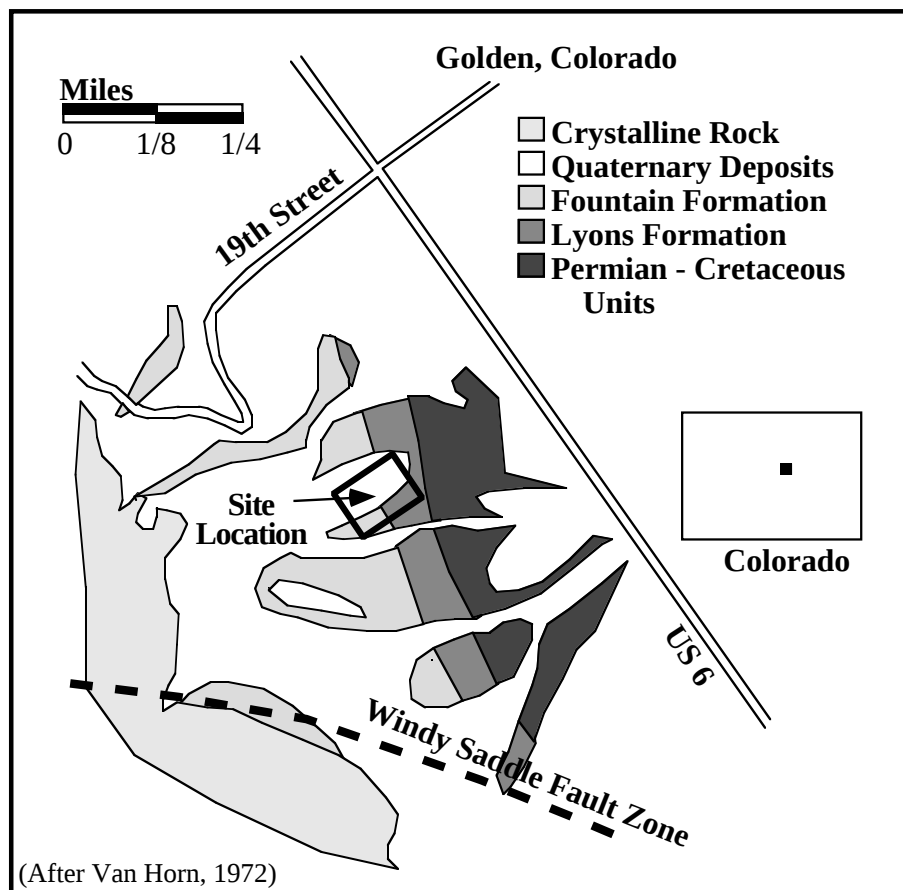


FIGURE 5-14. CSM Survey Field location map.

5.3.3.1: Evidence That Zonation is Required

Zonal kriging is needed because 1) the data populations and 2) the spatial statistics of the data from the two formations are different. The difference in spatial statistics, applied to kriging, is the most important reason for using zonal kriging. The data are divided into two data sets along the formation contact between the Fountain and Lyons Formations. The contact dips approximately 40° ENE. The contact is clearly defined on the seismic tomogram cross-sections (Figure 5.17). The full data set is shown in Figure 5.16 with all data converted to one of eight indicator values. The indicator values represent discrete sonic velocity ranges; hydraulic conductivity's (K) were

estimated from field and laboratory tests, or inferred from lithologic character and sonic velocity. Later optimal values of hydraulic conductivity were estimated using inverse flow modeling:

Indicator	Sonic Velocity (ft/sec)	Initial K Estimates (ft/day)	Optimized K Estimates (ft/day)	
			Hard Only	Hard/Soft
1	> 10870	0.0011	0.010	0.0020
2	10000 - 10870	0.0011	0.0063	0.00079
3	9050 - 10000	0.0025	0.63	0.050
4	8550 - 9050	0.0043	0.079	1.6x10 ⁻⁵
5	8050 - 8550	0.040	0.025	2.5x10 ⁻⁶
6	7250 - 8050	0.0043	NO	NO
7	6060 - 7250	0.40	0.016	0.0040
8	< 6060	7.8	NO	NO

NO = Not Optimized

Data sets, for the Fountain and Lyons Formations are shown in Figure 5.18. Examination of the full data set does not readily reveal sub-populations (Figure 5.19), but independent examination of data from the two formations reveals their differences. Indicators #4 through #7 have a high frequency in the Fountain Formation. In the Lyons Formation, indicators #1 through #4 are more frequent. Indicators #5 and #6 are poorly represented, and indicator #8 occurs exclusively in the Lyons Formation. The difference in the distributions can be explained by the materials the indicator represent:

Indicator	Material Description
1	Conglomerate (Lyons Formation)
2	Fine to coarse sandstone with conglomerate lenses
3	Fluvial sandstone (Lyons Formation)
4	Very fine to very coarse sandstone
5	No core recovered. Moderately consolidated with low-moderate clay
6	Two materials: 1) silty sandstone, and 2) poorly sorted sandstone with siltstone and conglomerate lenses
7	No core recovered. Poorly consolidated, low clay material
8	No core recovered. Well fractured area of any material type.

The spatial statistics of the indicators also differ between the two formations. Semivariograms were calculated for three principle axes (North-South, East-West, and Vertical) for each indicator threshold. The semivariogram model results are summarized in Table 5.1a-c. For all the indicators (except #8; which does not occur in Fountain Formation) there were significant differences between

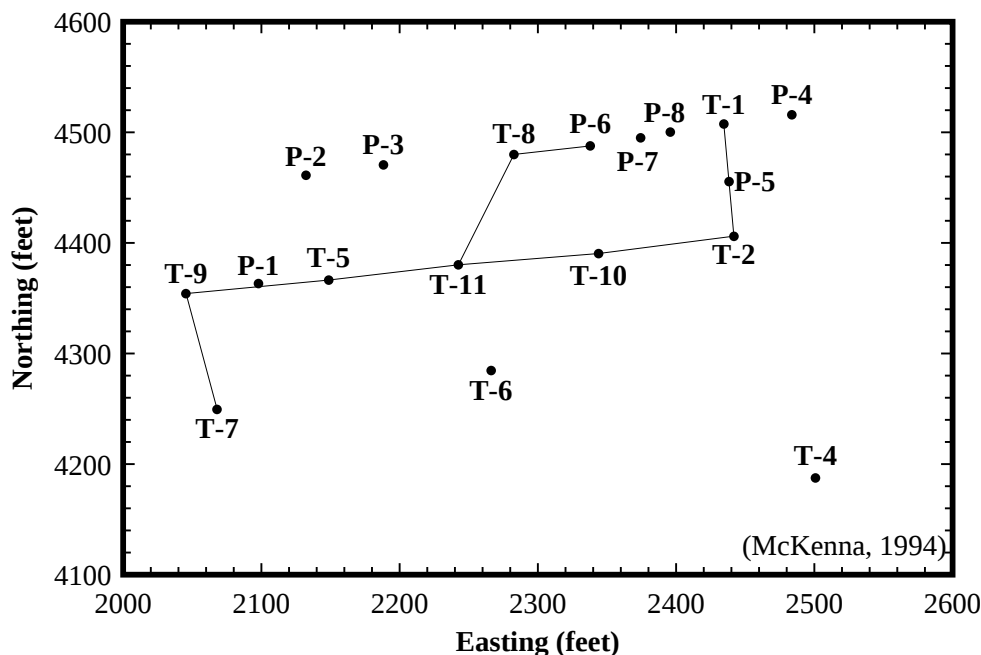


FIGURE 5-15. CSM Survey Field site map. Dots represent borehole locations. Solid lines identify location of tomography surveys.

the full data set, the Fountain Formation, and the Lyons Formation semivariogram models. For class 4, Fountain and Lyons Formation semivariograms, in the North-South direction the ranges are 72 ft. and 18 ft. respectively. In the East-West direction they are; 27 ft. and 48 ft. respectively. Not only are the ranges for the same indicator substantially different, but the principle anisotropy directions are 90° apart. Consequently it was concluded that two distinct zones are present at the site.

5.3.3.2: Grid and Model Definition

Single and two zone simulations were conducted. The regular model grid was defined as 80 columns representing 1600 feet in the X direction, 60 rows representing 1200 feet in the Y direction, and 72 layers representing 144 feet in the Z direction. The search ellipsoids for locating data were identical for both zones. The differences between the models were defined by 1) the semivariogram models used (Table 5.1), 2) the zone definitions (Figure 5.20), and 3) the quality of the soft data (Table 5.2). The first two differences have already been defined. The final difference though is less obvious, because the same data are used, however the calculation of the misclassification probabilities, p_1 and p_2 for Type-A soft data differ for the threshold and class simulations. In this example, the single zone model is simulated using thresholds, and as such, the

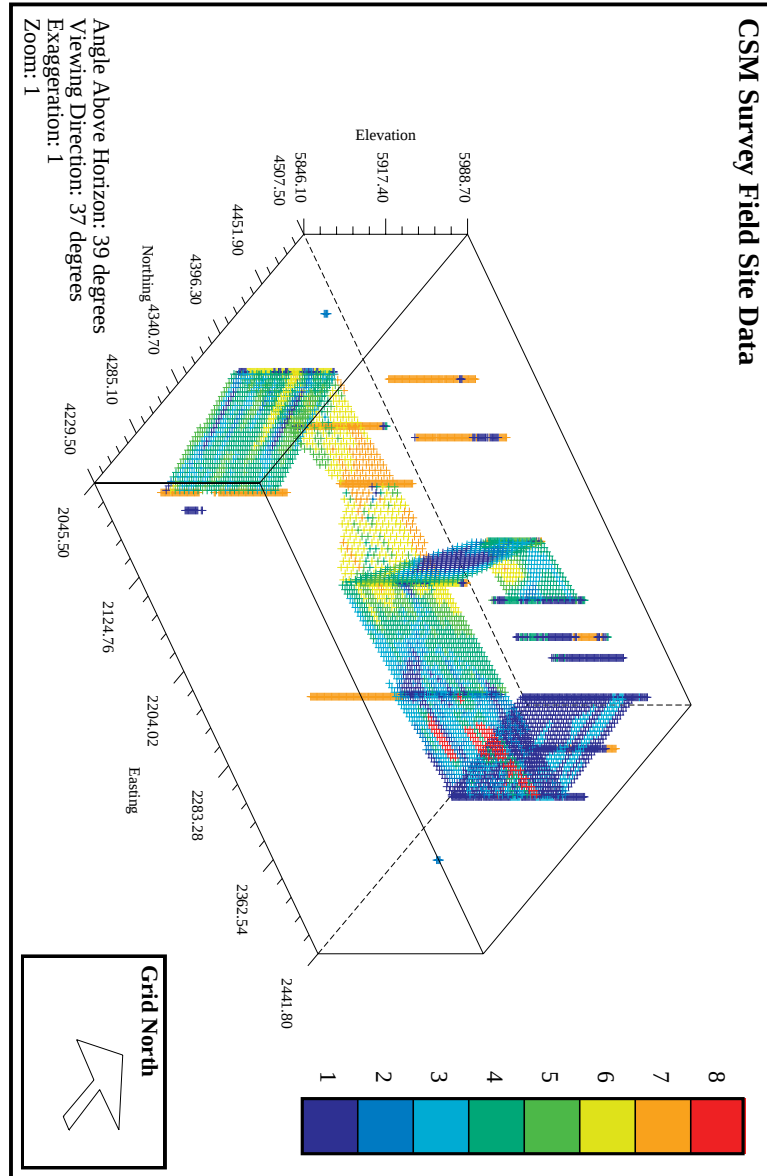


FIGURE 5-16. CSM Survey Field borehole and tomography data converted to indicators.

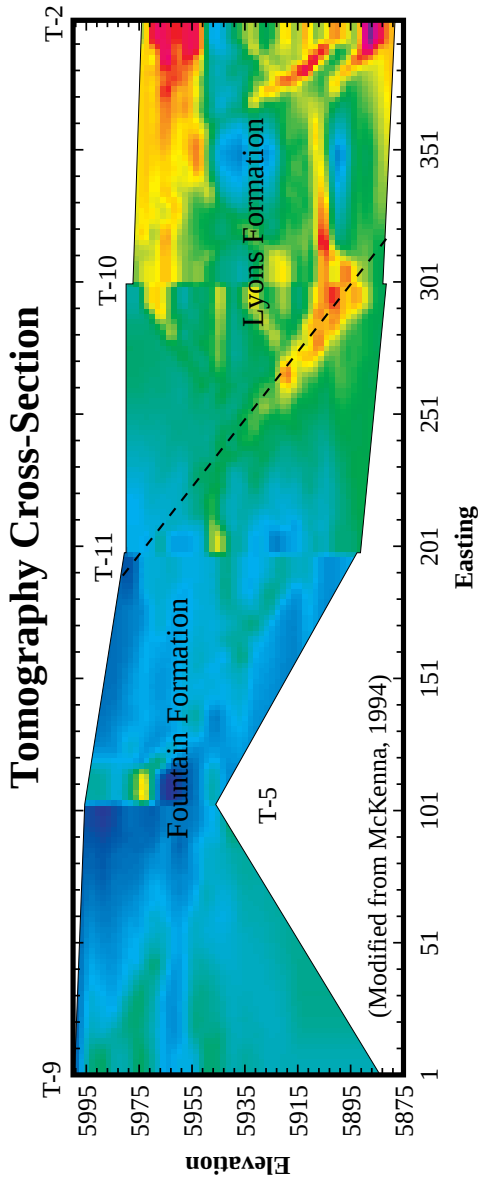


FIGURE 5-17. Tomographic cross-section at CSM Survey Field (viewing North-West). Dashed line is approximate location of Fountain / Lyons Formations contact.

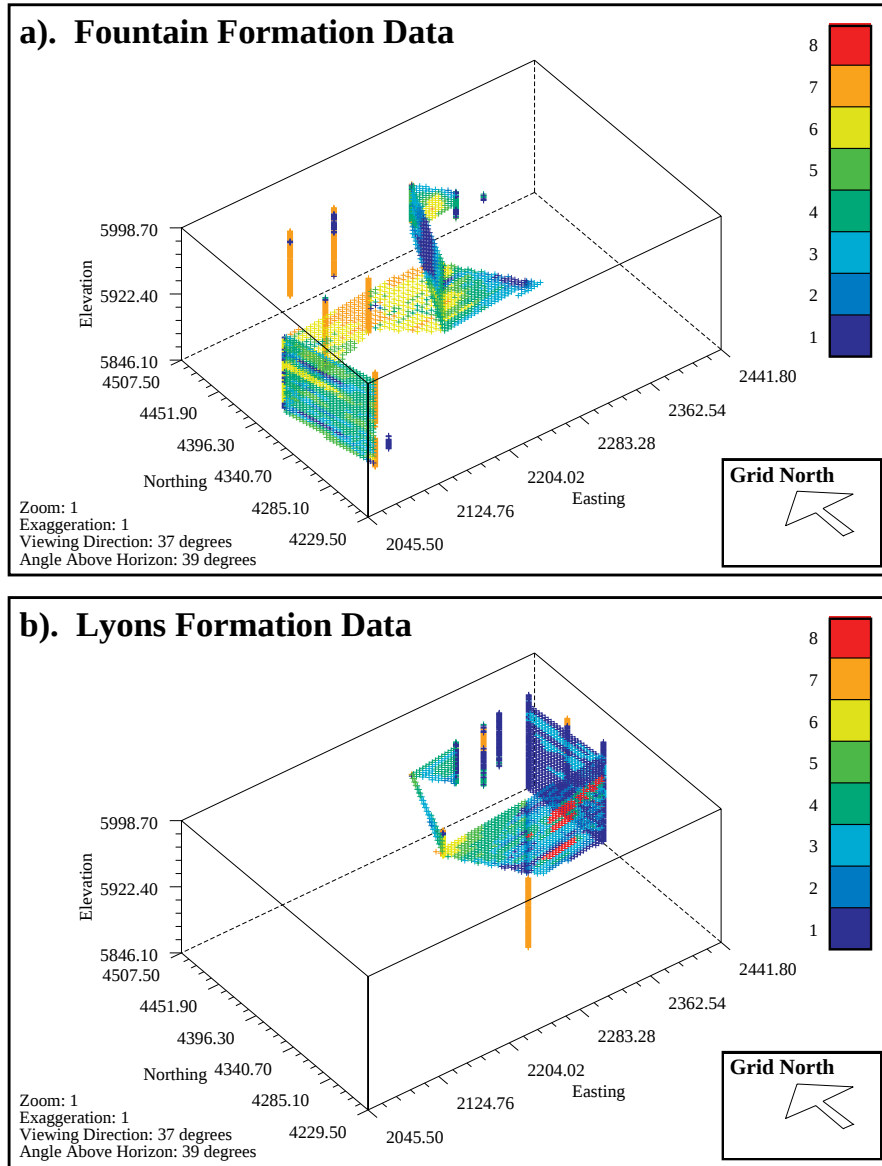


FIGURE 5-18. CSM Survey Field hard and soft data distributions (converted to indicators) for the Fountain (a) and Lyons (b) Formations.

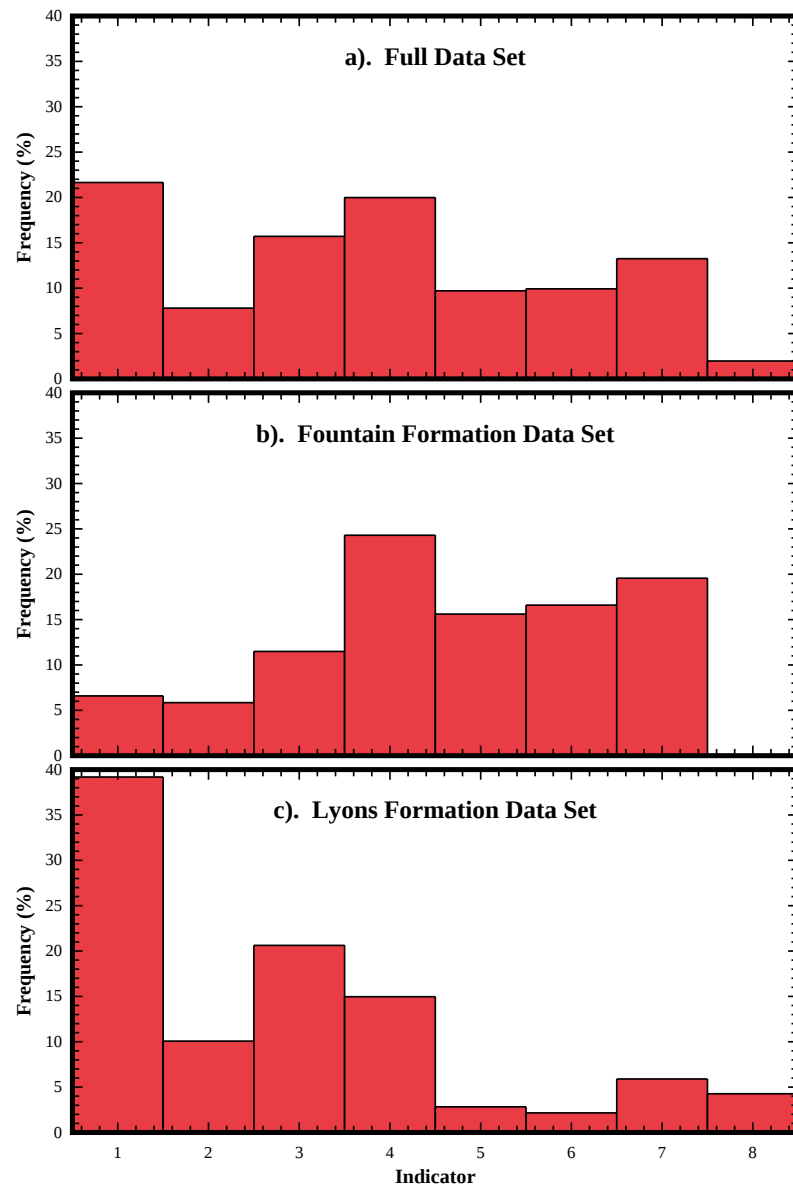


FIGURE 5-19. Distribution of hard and soft (Type-A only) data for full data set and for Fountain and Lyons Formations regions.

a). Single Zone - Threshold - All Models Spherical

Threshold	East-West		North-South		Vertical		Nugget
	Range	Sill	Range	Sill	Range	Sill	
1.5	81.0	0.118	155.6	0.118	81.0	0.0623	0.0516
2.5	93.0	0.207	126.0	0.207	54.0	0.0061	0.0
3.5	75.0	0.169	174.0	0.246	21.0	0.0468	0.0
	282.0	0.0783					
4.5	90.0	0.0831	15.0	0.149	3.0	0.0405	0.0
	204.0	0.145	99.0	0.0765	47.0	0.0358	
5.5	132.0	0.189	12.0	0.158	36.0	0.0692	0.0
			48.0	0.0308			
6.5	78.0	0.129	15.0	0.129	93.0	0.0284	0.0
7.5	61.5	0.0208	18.5	0.0208	36.9	0.0079	0.0

NOTE: Multi-nested models require two rows.

b). Fountain Formation - Zone 1/2 - All Models Spherical

Class	East-West		North-South		Vertical		Nugget
	Range	Sill	Range	Sill	Range	Sill	
1	18.0	0.0353	59.9	0.0353	69.0	0.0353	0.0265
2	15.0	0.0256	30.0	0.0256	21.0	0.0139	0.0297
					48.0	0.0115	
3	18.0	0.0215	15.0	0.0256	30.0	0.0256	0.0297
	60.0	0.0229					
4	27.0	0.111	72.0	0.111	72.0	0.111	0.0253
5	15.0	0.0500	57.0	0.100	27.0	0.100	0.0317
	156.0	0.500					
6	9.0	0.103	36.0	0.137	78.0	0.137	0.000
	60.0	0.0353					
7	66.0	0.137	27.0	0.137	44.0	0.0786	0.0214
8	74.0	0.142	125.0	0.142	74.0	0.0799	0.0100

NOTE: Multi-nested models require two rows.

TABLE 5.1 a,b. CSM Survey Field threshold (single-zone) and class (two-zone: Fountain and Lyons Formation) semivariogram models.

c). Lyons Formation - Zone 2/2 - All Models Spherical

Class	East-West		North-South		Vertical		Nugget
	Range	Sill	Range	Sill	Range	Sill	
1	75.0	0.160	114.0	0.160	75.0	0.160	0.0776
2	54.0	0.0902	12.0	0.0902	42.0	0.0583	0.000
3	84.0	0.148	84.0	0.148	69.0	0.148	0.0141
4	48.1	0.126	18.0	0.126	75.1	0.126	0.000
5	27.0	0.0275	18.0	0.0275	45.0	0.0275	0.000
6	15.0	0.0211	15.0	0.0211	5.0	0.0211	0.000
7	12.0	0.0468	12.0	0.0468	63.0	0.0468	0.00856
8	66.0	0.0410	15.0	0.0410	63.0	0.0410	0.000

NOTE: Multi-nested models require two rows.

TABLE 5.1 c. CSM Survey Field threshold (single-zone) and class (two-zone: Fountain and Lyons Formation) semivariogram models.

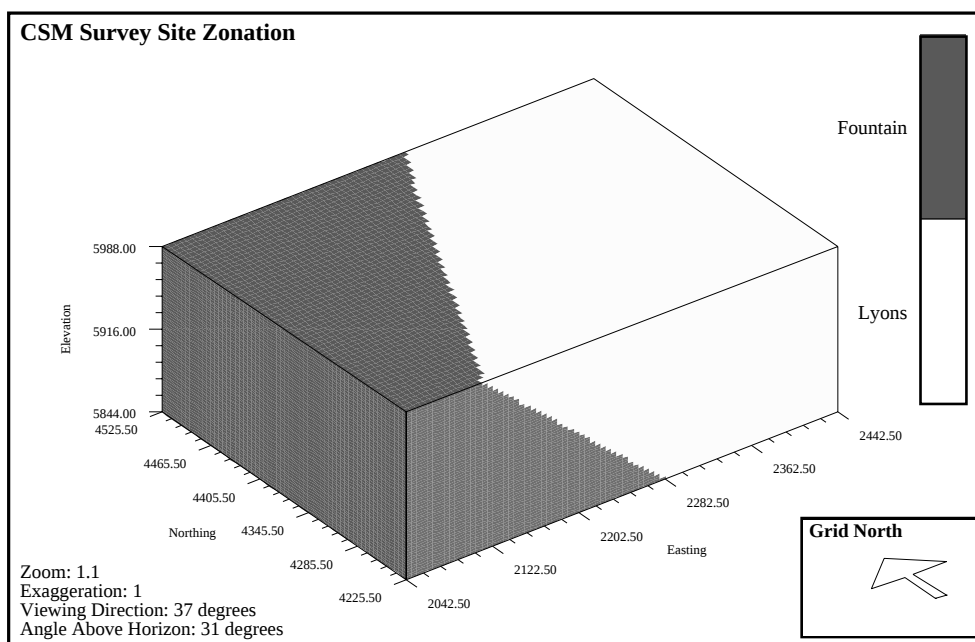


FIGURE 5-20. CSM Survey Field zone definition.

Threshold	Cumulative Probability		
	Hard	Soft	Difference
1.5	0.1638	0.2306	0.0668
2.5	0.2370	0.3120	0.0750
3.5	0.2706	0.5031	0.2325
4.5	0.3693	0.7306	0.3613
5.5	0.3886	0.8491	0.4605
6.5	0.5066	0.9429	0.4363
7.5	1.0000	0.9748	0.0252

Class	Individual Probability		
	Hard	Soft	Difference
1	0.0018	0.0843	0.1317
2	0.0148	0.0710	0.0524
3	0.0000	0.1479	0.1400
4	0.1107	0.2810	0.1552
5	0.0314	0.1919	0.1502
6	0.1697	0.1649	0.0137
7	0.6716	0.0589	0.6159
8	0.0000	0.0000	0.0000

Class	Individual Probability		
	Hard	Soft	Difference
1	0.3628	0.3936	0.0308
2	0.1451	0.0888	0.0563
3	0.0748	0.2415	0.1667
4	0.0839	0.1673	0.0834
5	0.0045	0.0347	0.0302
6	0.0544	0.0128	0.0416
7	0.2744	0.0012	0.2732
8	0.0000	0.0541	0.0541

TABLE 5.2. CSM Survey Field threshold (single-zone) and class (two-zone: Fountain and Lyons Formation) hard and soft data (Type-A only) sample data distributions.

p_1 - p_2 values are calculated based on a value being above or below a particular threshold (threshold #3, Figure 5.21). The p_1 - p_2 values used in the single-zone, threshold simulations are:

Threshold Velocity (ft/sec)	Threshold	P_1	P_2	$P_1 - P_2$
6060	7	0.00	0.00	0.00
7250	6	0.56	0.04	0.52
8050	5	0.58	0.05	0.53
8550	4	0.63	0.10	0.63
9050	3	0.84	0.17	0.67
10000	2	0.90	0.17	0.73
10870	1	0.91	0.15	0.74

In the two-zone case, class simulation is performed and the p_1 - p_2 values are calculated based on a value being between two thresholds (class #3, Figure 5.22). Because classes are more restrictive, the probability of misclassification is higher and the p_1 - p_2 values will be lower. This suggests that the soft data (Type-A) are less useful at reducing the uncertainty in the two-zone model. As will be shown, the two-zone simulations produce smaller uncertainties. The p_1 - p_2 values used in the two-zone, class simulations are:

Velocity Range (ft/sec)	Class	P_1	P_2	$P_1 - P_2$
< 6060	8	0.00	0.00	0.00
6060 - 7250	7	0.56	0.00	0.56
7250 - 8050	6	0.39	0.04	0.35
8050 - 8550	5	0.25	0.12	0.13
8550 - 9050	4	0.45	0.18	0.27
9050 - 10000	3	0.26	0.13	0.13
10000 - 10870	2	0.38	0.05	0.33
> 10870	1	0.84	0.00	0.84

Note, the less than 6060 ft/sec velocity measurements are only observed in the tomography cross sections. Hard data are not available to for calibration.

In addition to the differences in the p_1 - p_2 values, there are significant differences in the prior hard data CDF's and soft data CDF's. These differences also affect the realization calculations and are summarized in Table 5.2.

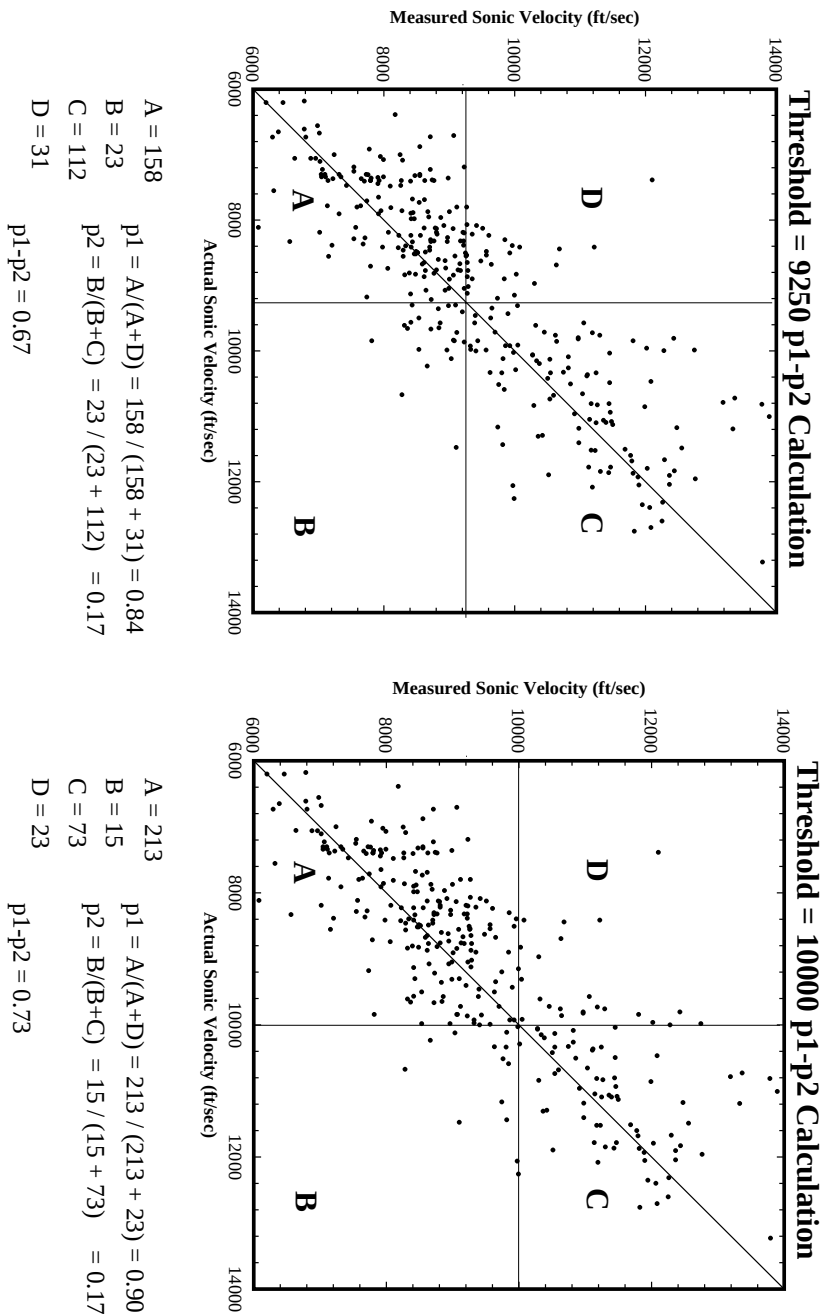


FIGURE 5-21. Graphical method for calculating p_1 and p_2 values for a specific threshold (After Alabert, 1987). Data from CSM Survey Field.

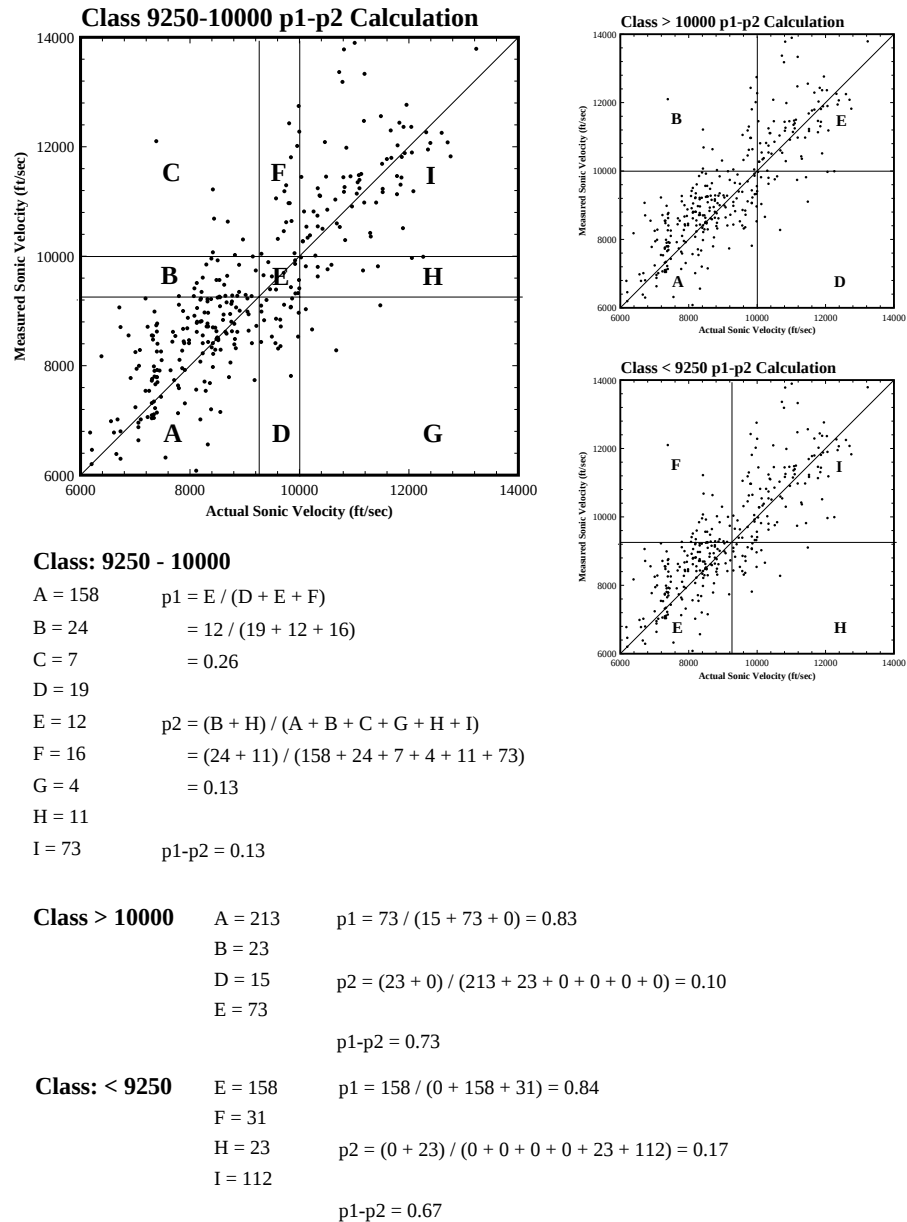


FIGURE 5-22. Graphical method for calculating p_1 and p_2 values for a specific class. Data from CSM Survey Field. Note the similarities between the *Class > 10000* and the *Class < 9250* diagrams, and the diagrams in Figure 5.21. They are fundamentally identical. This is always the case for the first and last class and threshold p_1 - p_2 calculations.

5.3.3.3: *Realizations and Indicator Populations*

To demonstrate the differences in simulations with and without zones, fifty conditional simulations were calculated for each model assumption (single and two-zone). When evaluating results, it is reasonable to expect that the calculated data distribution should approximately reproduce the original data sample population. Several sets of realizations shown here (Figures 5.23 through 5.28) demonstrate that modeling with two zones yields significantly different results than modeling with a single zone. With a single zone, the character of individual indicators is fairly uniform throughout the site. The Fountain Formation should primarily exhibit indicators 4-7, but the realization indicator populations poorly reproduced the initial data population when a single zone was used (Figures 5.24a, 5.26a, and 5.28a vs. Figure 5.19a). Using two-zones, the Lyons Formation is dominated by indicators 1-4 and 7, and the indicator populations for each two-zone model reasonably reproduce the field data distributions (Figures 5.24b-d, 5.26b-d, and 5.28b-d vs. Figure 5.19a-c). This is not proof, but it strongly suggests that the two-zone realizations are better approximations than the single-zone realizations.

5.3.3.4: *Minimizing Uncertainty*

Another approach to comparing the single and two-zone realizations, is to evaluate which series produces the smallest uncertainty. This is done visually and graphically in two steps. First, the probability that a particular indicator will occur at any specific location is calculated; and second, based on these calculations, the maximum probability any individual indicator will occur in a particular cell is calculated. The first series of maps are useful for evaluating where a particular indicator is likely to occur, and the second series is useful for identifying areas of the site where the modeler can be relatively certain or uncertain about the model results.

The maps in Figures 5.29 and 5.30 were developed by combining the results of 50 realizations into individual indicator probability maps. The difference resulting from using a single zone and two zone model was most pronounced with indicators #1 and #6. From the maps, it can be seen that the single zone realizations do not identify a transition of indicator frequency across the site; the uncertainties are fairly uniform except immediately near conditioning data. In the two-zone realizations, there are distinct differences. There is a high frequency of indicator #1 in the Lyons Formation (indicator #1 identifies two conglomerate Lyons Formation facies), consequently most cells have a relatively high probability of being indicator #1. Indicator #6, is rare in the Lyons Formation, thus its relative probability of occurrence to the Fountain Formation is low.

Ideally zonal kriging will yield more definitive results. The maps in Figure 5.31 indicate the maximum probability any particular indicator will occur in each cell. Visually, for the single-zone simulations, the only areas of low uncertainty are near the hard and soft conditioning data; in the two-zone model though, the improved definition of indicator #1, has reduced much of the uncertainty (green areas indicate much lower uncertainty than blue areas) in the Lyons Formation (right). Comparing the histograms from the single and two-zone simulations (Figures 5.32a and 5.32b), the two-zone model have fewer low probability cells, and substantially more mid-probability and high probability cells. The two-zone model exhibits a bi-modal distribution of uncertainty, due to separate populations from the two formation zones (Figures 5.32c and 5.32d).

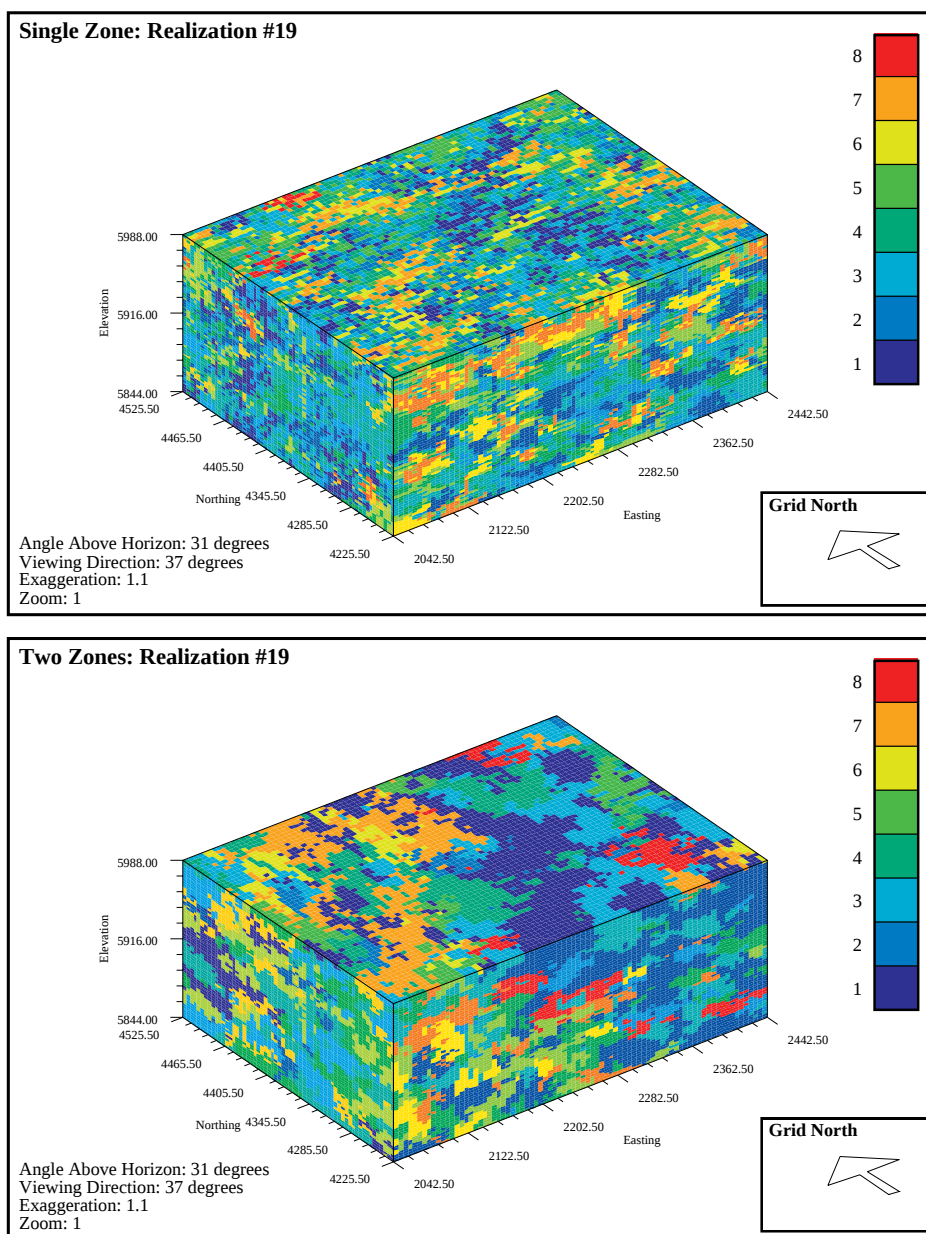


FIGURE 5-23. Single-zone and two-zone realization pair #19.

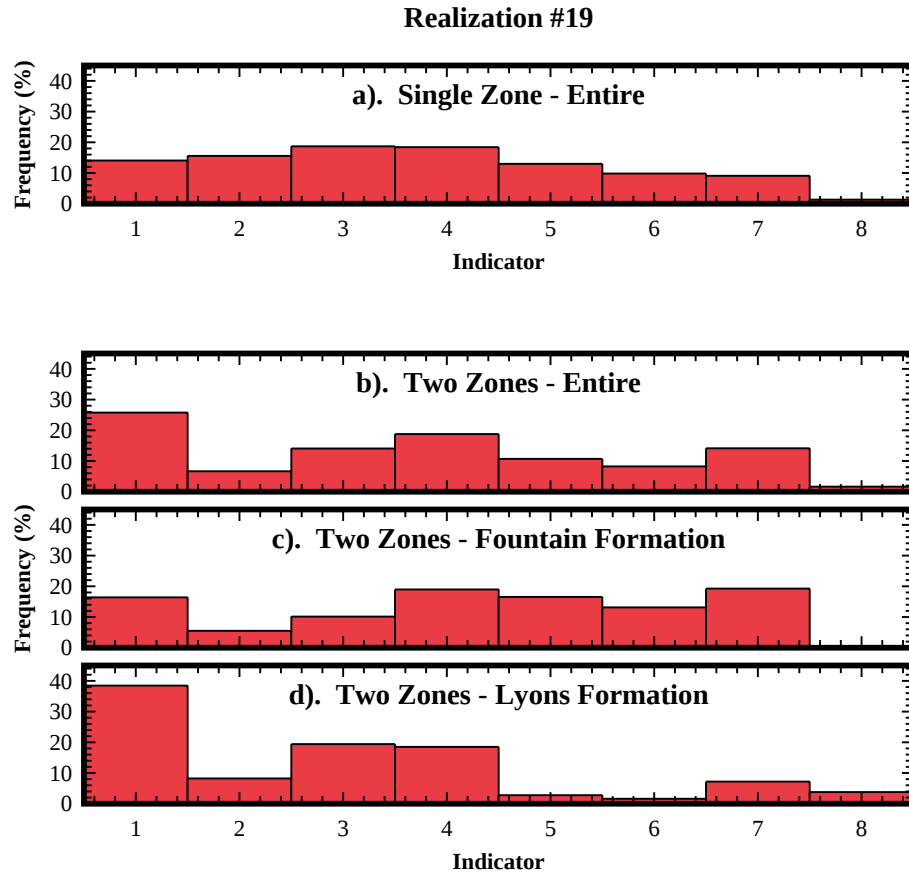


FIGURE 5-24. Distribution of indicators a) in the single-zone realization #19 (Figure 5.23a) and b-d) in two-zone realization #19 (Figure 5.23b). The single-zone realization poorly reproduces original data distribution (Figure 5.19a), whereas the two-zone realization reasonably reproduces the full and individual formation distributions (Figure 5.19a-c).

From these histograms (Figure 5.32) we can see that most of the model uncertainty reduction is due to the improvement in the definition of the Lyons Formation. The Fountain Formation uncertainty is slightly less than occurs when the entire site is modeled with a single-zone. These histograms imply that there is less uncertainty in the two-zone model which suggests the realizations have improved. Without further exploration or groundwater flow modeling, however, it is not possible to confirm this conclusion.

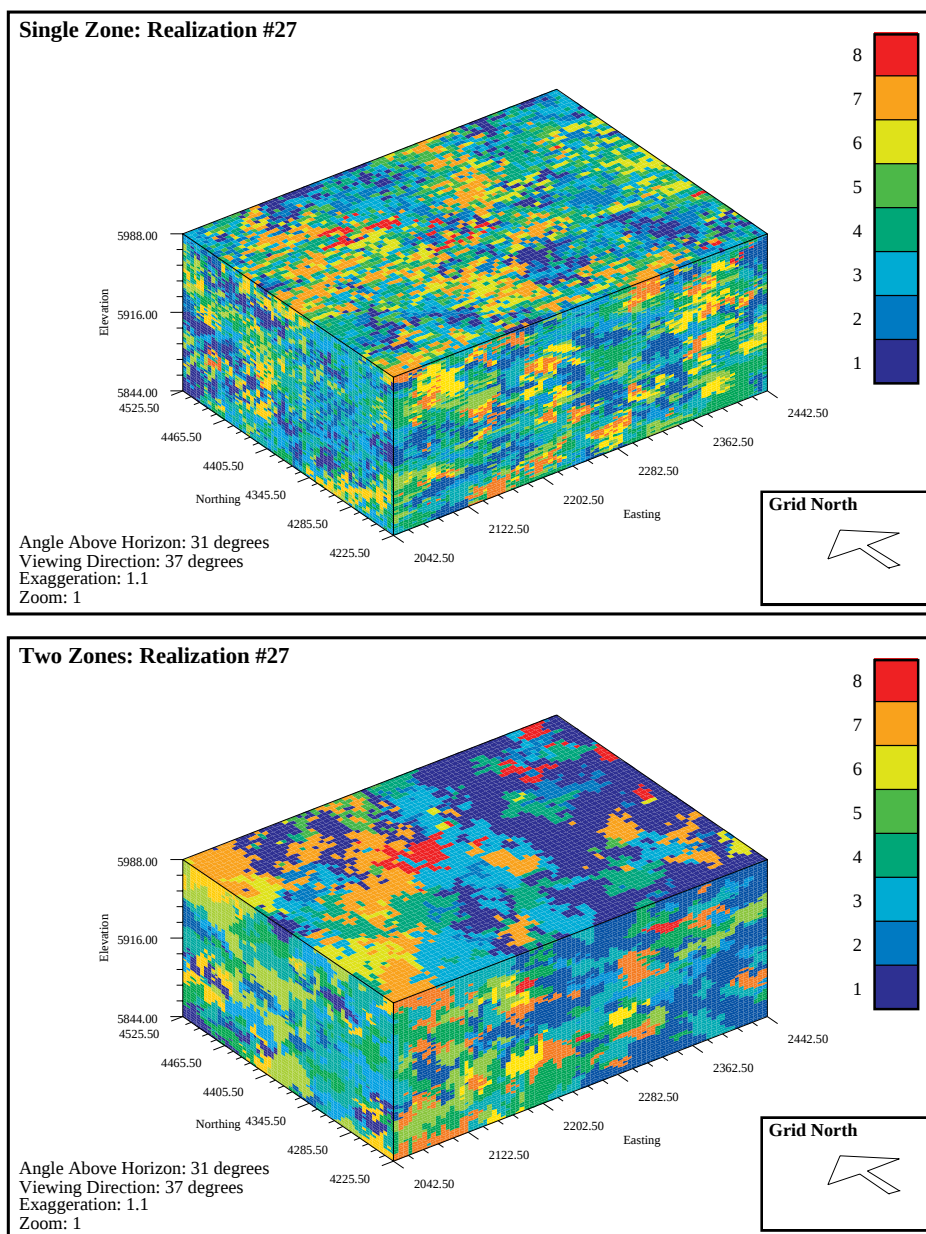


FIGURE 5-25. Single-zone and two-zone realization pair #27.

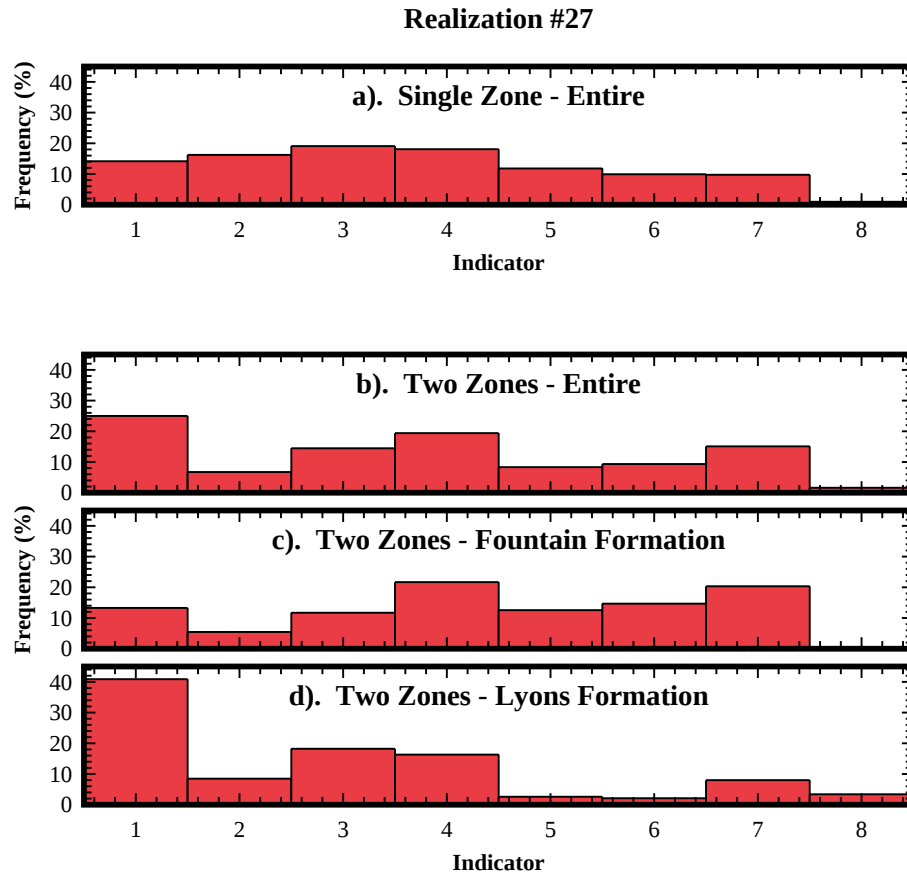


FIGURE 5-26. Distribution of indicators a) in the single-zone realization #27 (Figure 5.25a) and b-d) in the two-zone realization #27 (Figure 5.25b). The single-zone realization poorly reproduces original data distribution (Figure 5.19a), whereas the two-zone realization reasonably reproduces the full and individual formation distributions (Figure 5.19a-c).

5.4: Steps to Determine if Zonal Kriging is Appropriate

Before using zonal kriging, it is important to determine whether the statistics of the data suggest that zonation is appropriate. Several items to consider are whether:

- a visual display of the field geologic data suggests distinct zones.
- the full data set exhibits a bi/multi-modal population frequency distribution.
- there are statistical differences between the populations in each suspected zone.

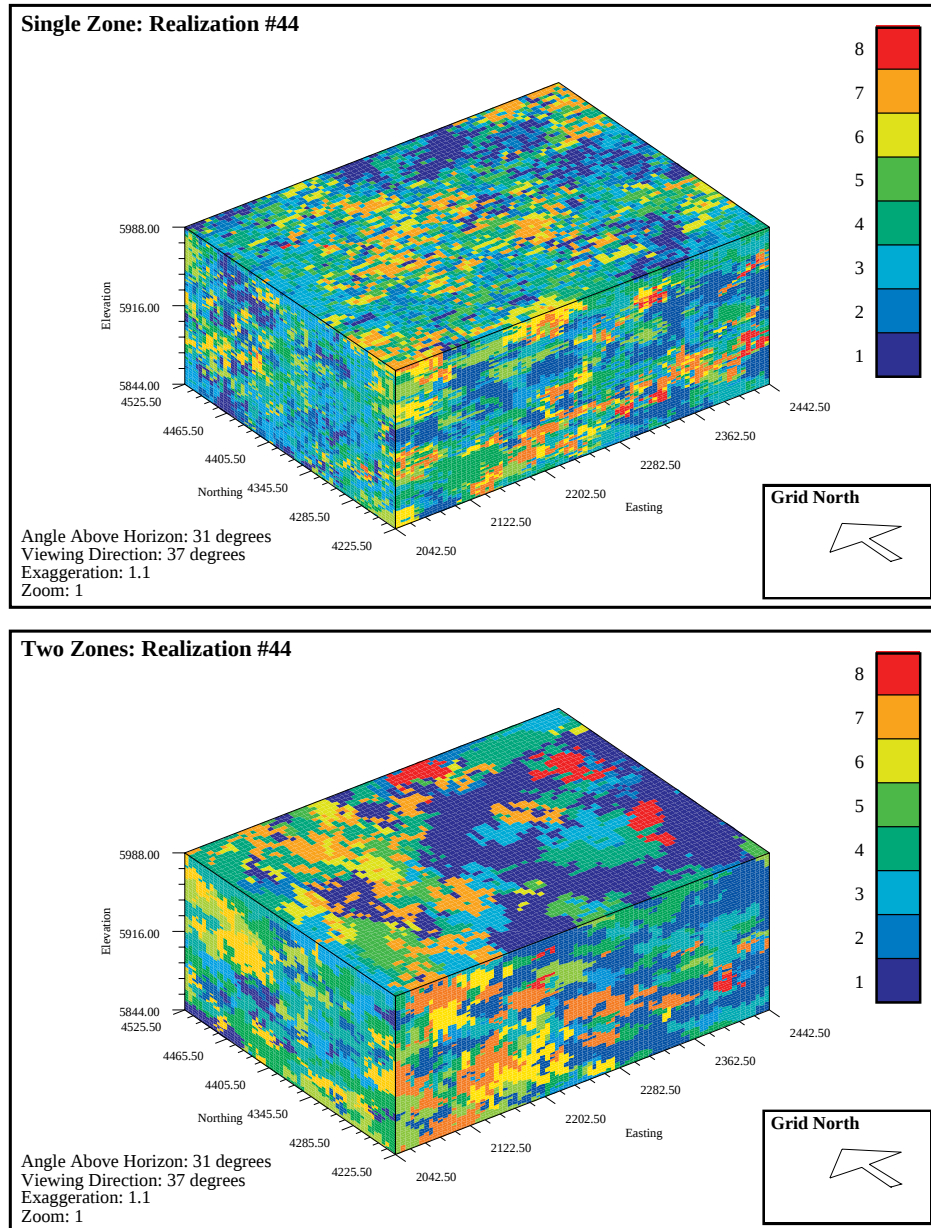


FIGURE 5-27. Single-zone and two-zone realization pair #44.

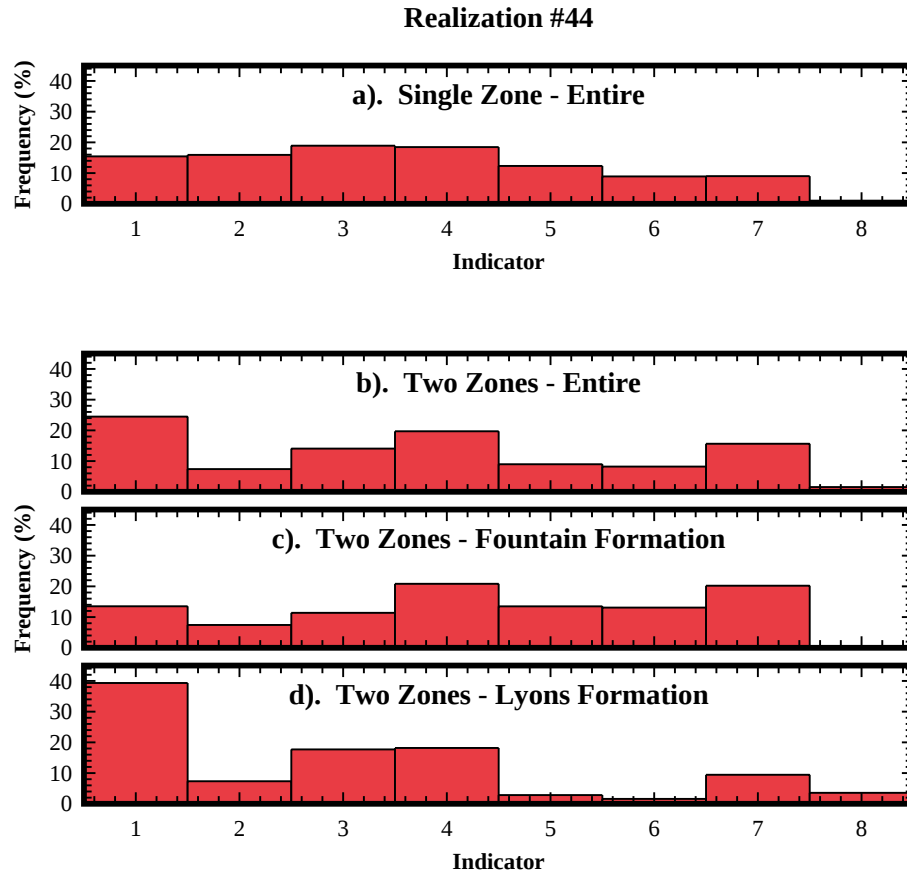


FIGURE 5-28. Distribution of indicators a) in the single-zone realization #44 (Figure 5.27a) and b-d) in the two-zone realization #44 (Figure 5.27b). The single-zone realization poorly reproduces original data distribution (Figure 5.19a), whereas the two-zone realization reasonably reproduces the full and individual formation distributions (Figure 5.19a-c).

- the frequency distribution between the populations in each of the suspected zones varies significantly.
- the spatial statistics (semivariogram models) vary significantly between zones.

If these conditions exist, consider using zonal kriging.

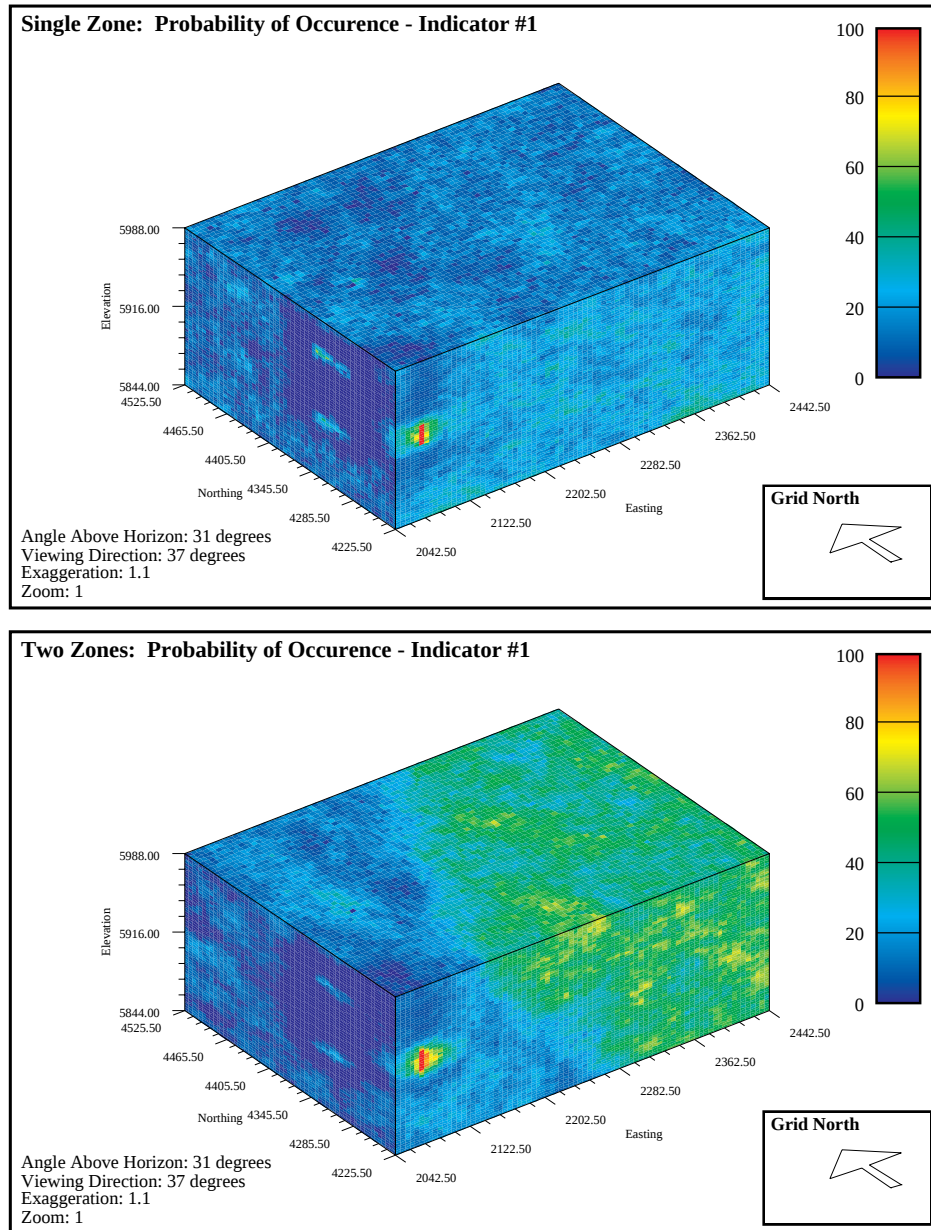


FIGURE 5-29. Maximum probability of occurrence of indicator #1 for single-zone and two-zone realizations. In the single-zone model, the uncertainty is fairly consistent throughout the site except near hard and soft data locations. There are significant differences between the zones in the two-zone map. The Lyons Formation has a much higher percentage of Indicator #1, and this is reflected in the maximum probability map.

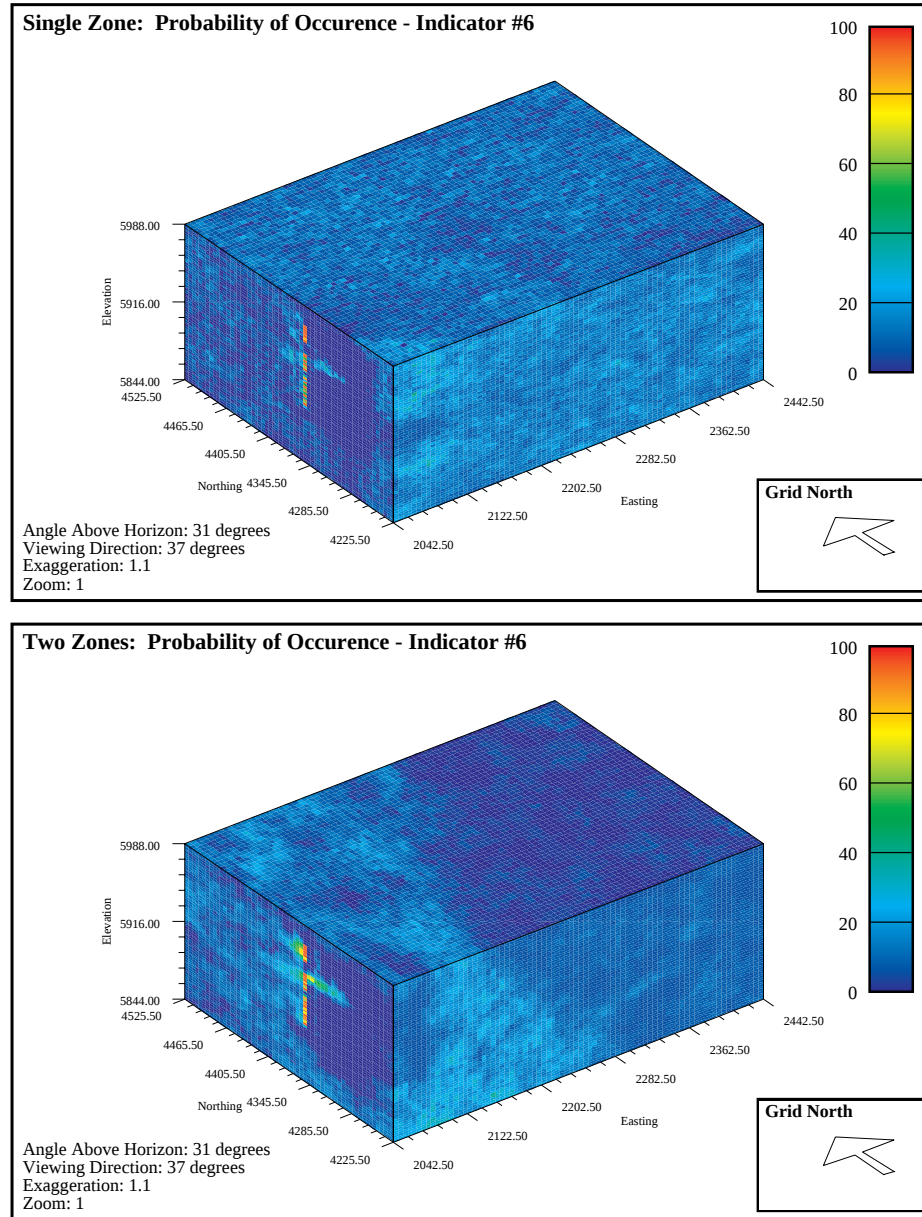


FIGURE 5-30. Maximum probability of occurrence of indicator #6 for single-zone and two-zone realizations. Note, in the single-zone model, the uncertainty is fairly consistent throughout the site except near hard and soft data locations. There are significant differences between the zones in the two-zone map. The Lyons Formation has a much lower percentage of Indicator #6, and this is reflected in the maximum probability map.

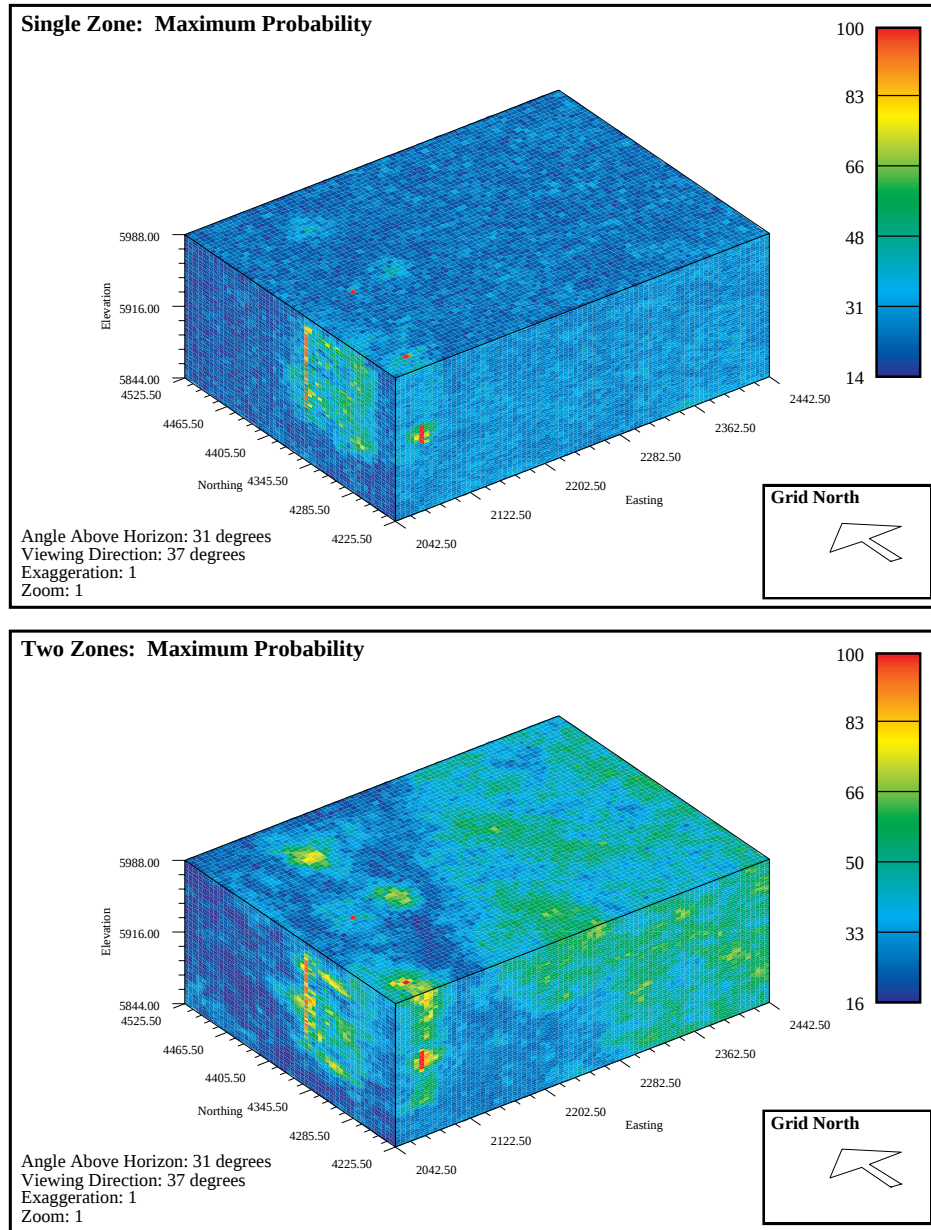


FIGURE 5-31. Maximum probability of occurrence of any indicator. In the single-zone model, uncertainty is fairly uniform across the site except near hard and soft data locations. In the two-zone model, the Lyons Formation exhibits significantly lower uncertainty. These maps are useful for identifying the spatial distribution of uncertainty.

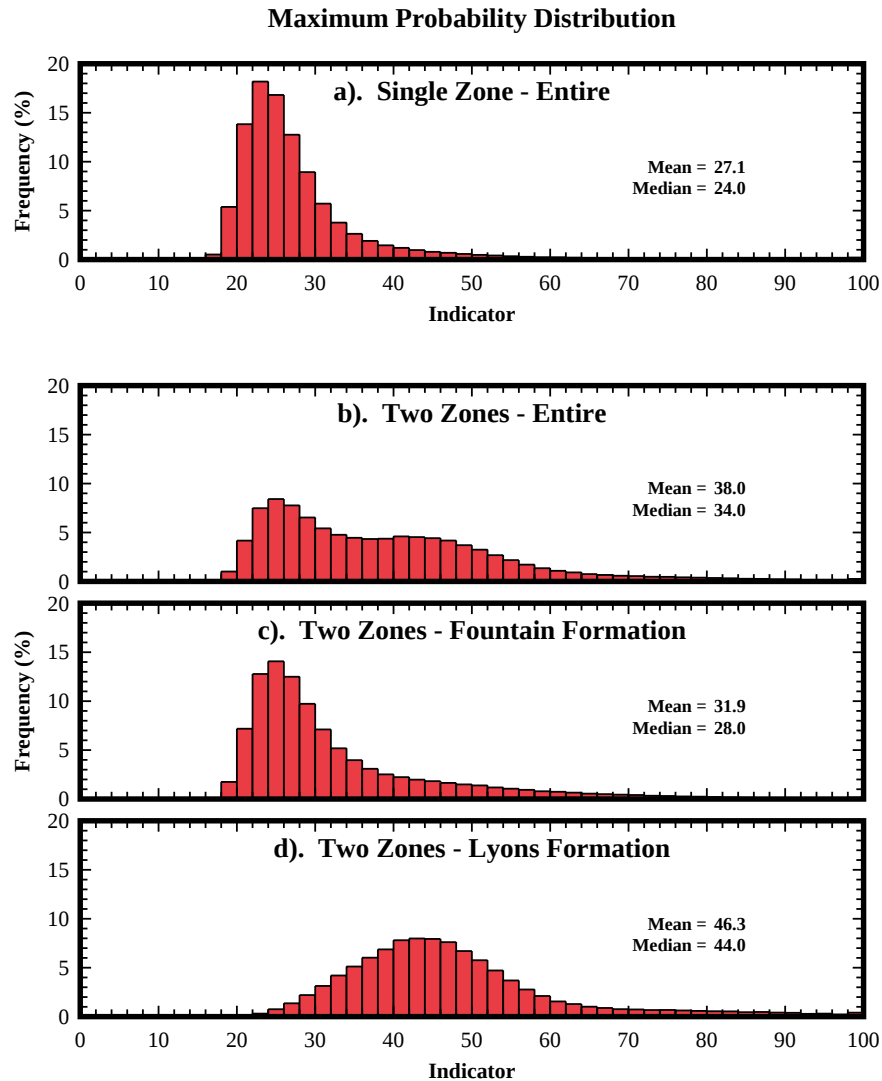


FIGURE 5-32. Distribution of the maximum probability of occurrence of any indicator. The two-zone model (b) has fewer low probability cells and more mid-probability cells than the single-zone model (a). This implies the two-zone model has less uncertainty, and is therefore a better solution. The two-zone histogram also has a bi-modal distribution (b), and the populations may be separated by formation (c and d). Most of the model improvement comes from improved definition of the Lyons Formation.

5.5: Conclusions

Through a series of examples, it has been shown that zonal kriging can yield significantly different results than those obtained using SK or MCIS alone. At sites where the assumption of stationarity is not valid, correctly applied zonal kriging produces realizations that more accurately represent site conditions with greater certainty. The technique requires additional data processing to define the model, and unusual boundary effects may occur when sample data are sparse or are located at spacings near the range of the semivariogram models. These shortcomings, however are offset by the increased certainty, improved accuracy, and modeling flexibility.

