15 Sept 2000

KRIGING OVERVIEW - ASARCO OMAHA FACILITY SUPERFUND SITE

Overview

Sample point coverages were created from subsets of four text files of sample data that were originally provided to us within an ArcView project file for the Omaha Lead study. The data represented Commercial/Industrial, Daycare, EBLs, and Residential properties, respectively. The point locational coordinates within these files were provided in Decimal Degrees format, but there were no metadata or other documentation files provided with the ArcView project file that would help us discern which projection was being applied for the spatial views. Therefore, we contacted previous users of the data set until we were able to talk with Mike Toolen of Zambrana Co., one of the data set originators, who confirmed for us that the ArcView projection parameters were Nebraska State Plane, Fipszone 2600, NAD83 datum. We prepared new point coverages according to this projection and then set forth to perform the point-to-grid kriging using ArcInfo GRID software.

The point coverage was processed to remove missing values characterized by the value '888888'. The parameters we chose for the first kriging run were as follows: ordinary kriging, spherical function, 12-point sample method, 90-meter output grid cell size. We had no *a priori* basis for setting the output cell size, but adopted 90 m as being a fine enough resolution to capture any significant changes in the Pb gradient within the immediate study area.

We produced seven plots as a result of this exercise: two base maps on 17" x 22" white paper and five overlay kriged grid maps on 17" x 22" clear mylar sheets. One of the two base maps contains a color-coded view of the Pb sample point locations, whereas the other base map has the points left off so that the kriged-grid overlays containing their own sample point locations will be visible. The five kriged-grid overlays represent All Properties Combined, Commercial/Industrial, Daycare, EBLs, and Residential properties, respectively, using a common color scheme shown in the legend at the base of each overlay sheet. Color shading changed from green to red tones at a concentration of 400 mg/Kg which was considered a base threshold for significant lead content.

When viewing the overlays and base maps, the user will quickly observe that the range of kriged-grid values (up to 1400 mg/Kg or so) is much narrower than that of the actual Pb samples (up to nearly 5000 mg/Kg in one set). It should be noted that the very highest Pb values are few indeed, and are outlying with respect to the magnitudes of the other samples in the data set. Also, there is a natural tendency for the kriging process to "smooth" extremes in peaks and valleys within a data set, yielding a surface that is markedly less coarse than the input data set.

Utilizing the Kriging Model

The kriging model can be used to optimize the locations of additional samples. Several examples can be seen on the "All Properties" overlay. Exiting the center ring and extending to the east/southeast is a red/contaminant aberration of the kriging model. This is created by a high concentration sample point located outside the eastern region of view. Taking more samples in between the area where the model is being distorted and the outlying sample point will result in a more accurate definition of the contaminated area. This is an example of the adage that a hotspot is define by the cold areas surrounding it. This sort of extension error can be seen on the top of the same overlay where the kriging model is distorted due to insufficient samples resulting in a larger contaminated area than actually exists.

Population Partitioning

The method of population partitioning first divides the sample of observed lead concentrations into two or more homogeneous subpopulations. Population partitioning then models the probability of observed lead concentrations as a mixture of normal distributions. For Omaha Lead (daycare) data, this resulted in seven subpopulations:

Subpopulation 1: observed lead inside the range 0 - 50 mg/kg
Subpopulation 2: observed lead inside the range 50 - 60 mg/kg
Subpopulation 3: observed lead inside the range 60 - 200 mg/kg
Subpopulation 4: observed lead inside the range 200 - 500 mg/kg
Subpopulation 5: observed lead inside the range 500 - 600 mg/kg
Subpopulation 6: observed lead inside the range 600 - 800 mg/kg
Subpopulation 7: observed lead above 800 mg/kg.

Figure 1 shows the subpopulations spatially. Figures 2 through 8 show the results of the Kolmogorov-Smirnov test of normality for each of the 7 subpopulations. Descriptive statistics and 90% confidence intervals for mean lead concentrations for each of these subpopulations are shown in Table 1. Subpopulations 1 through 4 have mean lead concentrations less than 400 mg/kg, and subpopulations 5 through 7 have mean lead concentrations greater than 400 mg/kg (all with 90% confidence). This table can be used to decide where to collect future samples, as well as how many to collect.

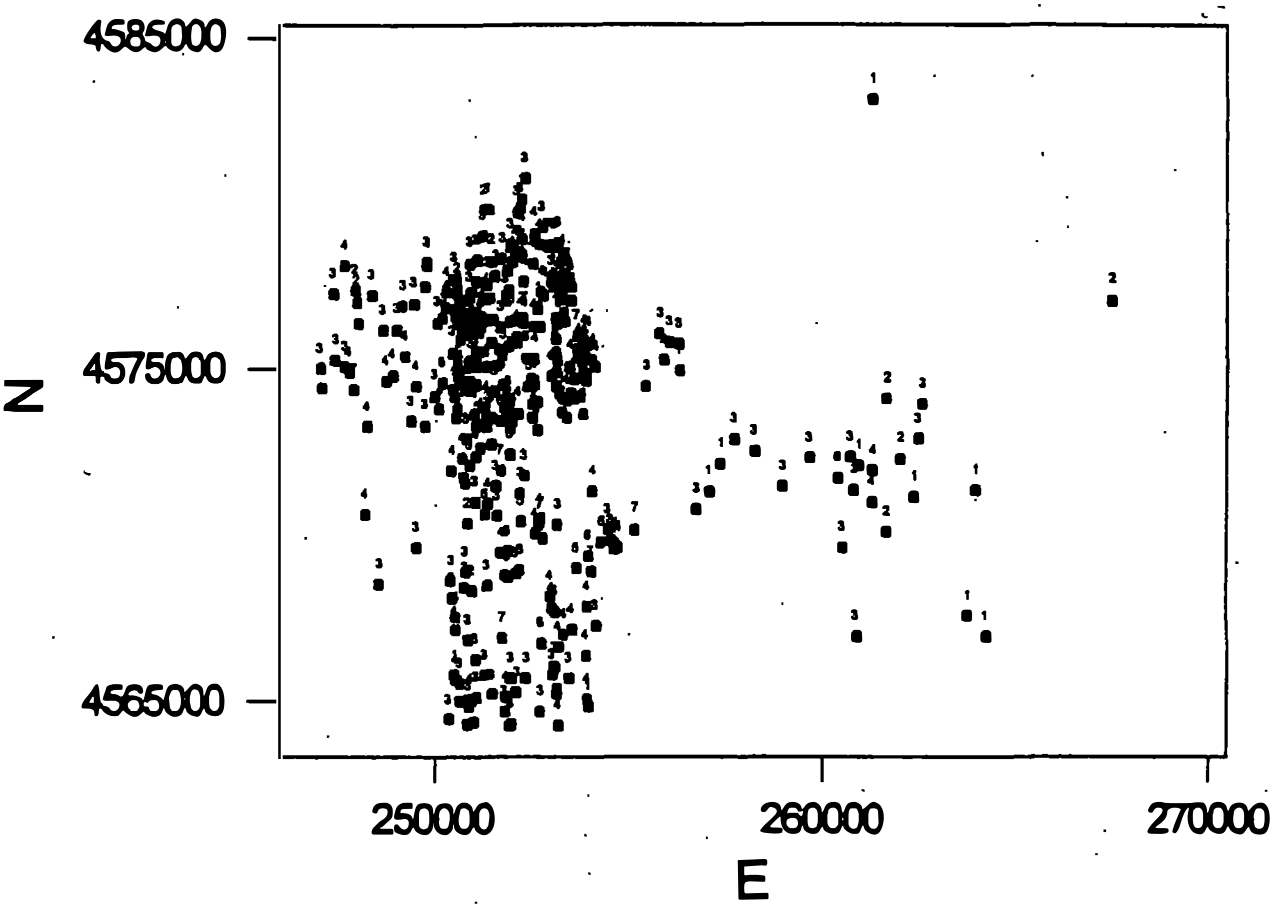

Figure 1: population partitioning results for lead data (daycare facilities - 1 = lowest lead concentrations,...., 7= highest measured lead concentrations)

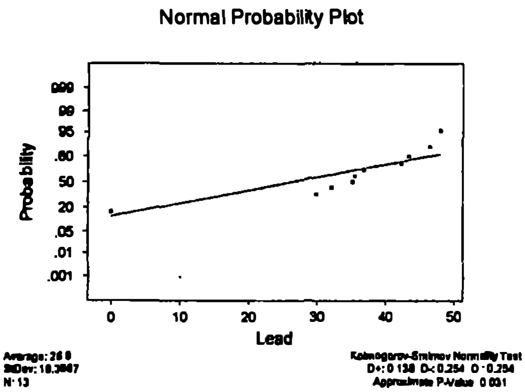

Figure 2: KS Test of normality for subpopulation 1 (P = 0.031)

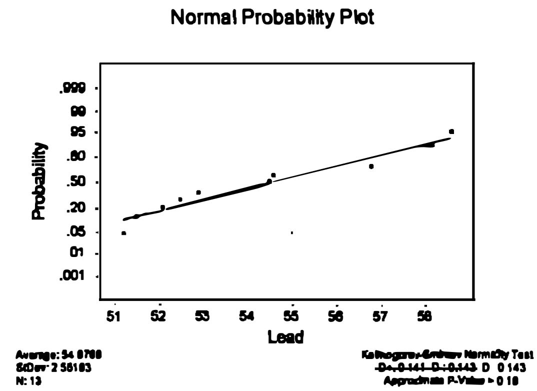

Figure 3: KS Test of normality for subpopulation 3 (P > 0.15)

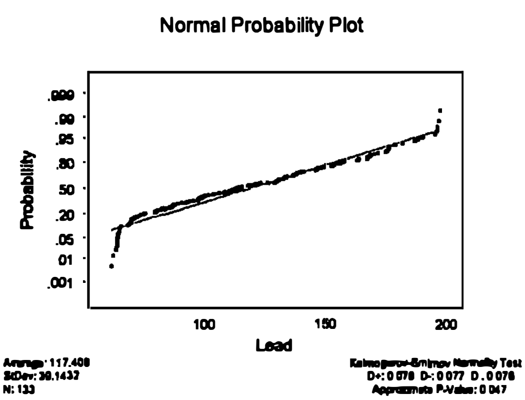

Figure 4: KS Test of normality for subpopulation 3 (P = 0.047)

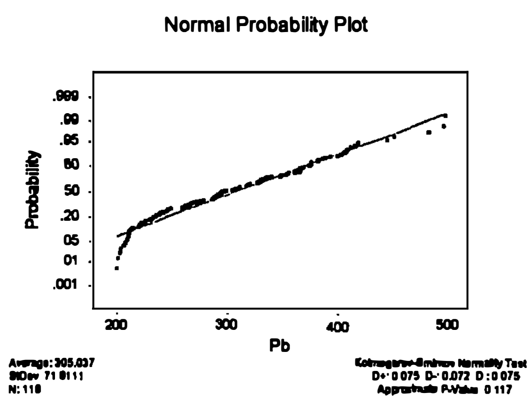

Figure 5: KS Test of normality for subpopulation 4 (P = 0.117)

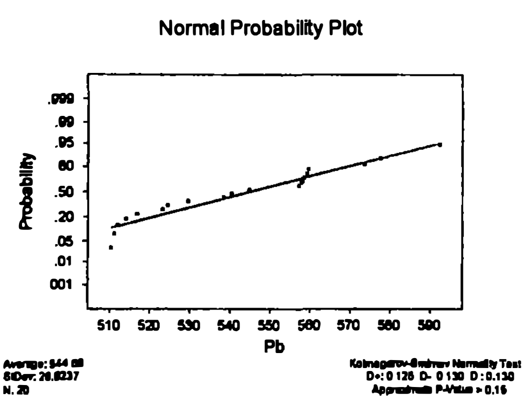

Figure 6: KS Test of normality for subpopulation 5 (P > 0.15)

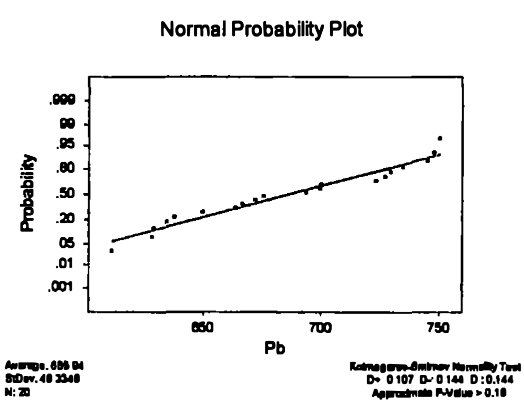

Figure 7: KS Test of normality for subpopulation 6 (P > 0.15)

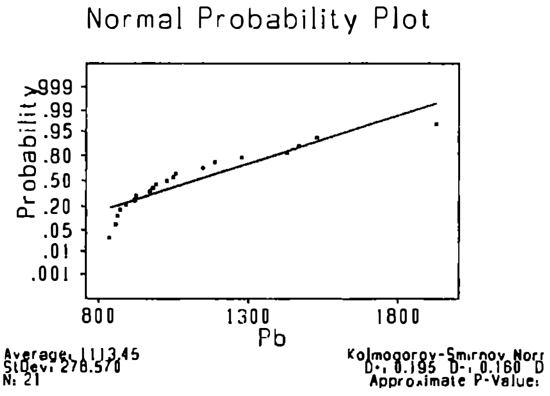

Figure 8: KS Test of normality for subpopulation 7 (P = 0.04)

TABLE 1: Descriptive statistics and 90% confidence Intervals for Omaha Lead by the Subpopulations

SubPop	N	Mean	Median	sd	L90	U90
1	13	26.90	35.2	19.390	15.178	38.62
2	13	54.68	54.5	2.562	53.130	56.22
3	133	117.41	113.0	39.140	110.704	124.12
4	115	305.04	295.0	71.810	291.767	318.31
5	20	544.68	542.6	26.820	532.122	557.24
6	20	685.90	684.6	45.300	664.760	707.04
7	21	1113.40	1029.6	278.600	986.573	1240.23