Using Spatial Statistical Methods To Determine an Effective Minimum Detectable Activity – 15681

Matthew Knop and Lynn Kidman
Navarro, Las Vegas, Nevada 89030

ABSTRACT

Instrument minimum detectable activity (MDA) is generally established by collecting multiple measurements over time of a single representative background location to estimate a probability distribution of instrument readings. A reading that is considered to be greater than would be expected from the background probability distribution (i.e., a low-probability event) is deemed to be sufficiently unlikely that it must be attributable to added radioactivity and is therefore used as the instrument MDA. An effective MDA can be determined from spatially distributed data that are much lower than the instrument MDA. This has been accomplished using a large number of spatially distributed data points (from a radiation survey) where patterns of elevated survey readings can be distinguished from background at levels. Given the variation of the instrument response and the relatively low regulatory action levels, it is vital to establish MDA thresholds that are well below corrective action levels. Processing of radiological survey data with the use of histogram binning and geostatistics in enabling pattern recognition can be used to establish effective field MDAs that are much lower than instrument MDAs.

INTRODUCTION

Remediation of historical surface radiological contamination at the Nevada National Security Site (NNSS) (previously known as the Nevada Test Site) is addressed through the *Federal Facility Agreement and Consent Order* (FFACO) [1]. A risk-based corrective action strategy is employed that evaluates site closure actions based on the dose received by an individual. Total effective dose (TED) is calculated at all sample locations to determine whether corrective action is necessary. If corrective action is necessary, the spatial extent of the contamination must be determined. The method used at the NNSS to define the extent of contamination is to establish a spatial correlation between TED and corresponding values from radiation surveys. The spatial patterns of radioactivity from the radiation surveys can then be used to delineate the area that exceeds a given dose or corrective action level (e.g., 25 mrem). It is necessary to evaluate the overall survey characteristics to determine whether the survey data can effectively be used to characterize the contamination distribution at contamination levels below corrective action levels. The radiological survey must be able to distinguish contamination from background at well below the dose action level.

MARSSIM (Multi-Agency Radiation Survey and Site Investigation Manual) prescribes a method for determining the minimum detectable activity (MDA) for a given instrument using a well-characterized control source. Furthermore, terrestrial survey instrumentation is evaluated for sensitivity in terms of measurable activity before a survey is conducted with an instrument [2]. While this MDA evaluates the capability of an instrument to detect contaminants relative to action levels, it is also important to evaluate the effective MDA of a survey over variable field conditions. This effective MDA for a particular survey can be characterized through statistical analysis of field data. One can use statistical methods and Geographic Information Systems (GIS) to identify the elevated response levels associated with contamination even in study areas with varied level of background. The lowest responses that are recognizably different than the uncontaminated signal response distribution can be considered the effective MDA for the survey. Clustering of data points is indicative of non-normal distribution, which is indicative of contaminated field conditions. The value at which clustering is first detected can be compared to the dose isopleth (action level) to determine whether the effective MDA survey supports the regulatory dose decision.

Because the method characterizes the sensor response in actual field conditions, the methods presented address many of the variables that have been identified as difficult to control using traditional methods. Difficult-to-control variables such as instrument scatter, site variability, sensor geometry, and other factors that affect the ability to measure specific activity are inherent in the data themselves and are therefore addressed via this approach.

These methods present an alternative way to evaluate whether or not an instrument is capable of characterizing contamination at a level representing the dose received by an individual given a certain exposure scenario. Because this calculated dose is correlated with a radiological survey value, one can estimate the activity associated with the dose action level.

RADIOLOGICAL SURVEYS, DATA ACQUISITION

Geostatistical methods of dealing with large volumes of data have been used in the geosciences for years, but applying these methods to radioactive contamination is relatively novel [3]. Given that the study sites are located at the NNSS and are sometimes several square miles in size, these methods of analyzing large volumes of data are understandably useful.

TERRESTRIAL RADIOLOGICAL SURVEYS

Terrestrial radiological surveys are conducted to characterize an area of contamination. Survey data are generated by connecting the radiological detection instrument to a Trimble GeoXT GPS unit and collecting data at 1-second intervals. The instruments are configured to use Wide Area Augmentation System (WAAS) correction methodology, as there are no available base stations within 80 kilometers with which to perform post-processing differential corrections. This produces survey data accurate to 1 meter or better.

The surveys are conducted by walking, using a side mount on a utility task vehicle (UTV) or pulling a trailer behind a UTV. The radiological detection instrumentation used includes a variety of plastic scintillators and sodium iodide detectors (e.g., FIDLER, PRM 470, Detector Array Rack Towed). The instrument used for a particular survey depends upon the characteristics of the contaminant plume. Every effort is made to maintain a consistent geometry throughout the survey.

The first step in conducting a survey is to collect a daily reference distribution at a site that is selected to roughly represent uncontaminated field conditions in the vicinity of the contaminated area. Data are collected at this reference site every day for each data collection setup (specific instrument/Global Positioning System [GPS] survey hardware setup). This allows for characterization of the daily variability of field conditions that affect the overall survey, such as fluctuation in cosmic and radon contributions.

Optimally, these daily reference surveys should be performed over the same area and should cover terrain that is representative of site conditions. However, due to variability with regard to GPS timing, survey geometry, and other factors, it was determined that it was not possible to collect data from the same points on a day-to-day basis performing a reference traverse. Therefore, a single site roughly representative of site conditions was chosen for a reference location, and the radiological survey equipment was set up at this single site for a minimum of five minutes every day. The resultant histogram showed a normal distribution that shifted with day-to-day variations in background but remained descriptively similar with regard to statistical measures of kurtosis and skewness.

The instrument counts are processed to present the instrument readings relative to that particular day's reference (expressed in multiples of background [MOB]). This normalization of the survey data enables one to analyze data collected over multiple days in conjunction with one another.

Each day's survey is stored with both raw data and with respect to the daily reference signature. The data are collected and stored in an ArcGIS geodatabase for subsequent analysis.

INTERPRETATION OF RAW CHARACTERIZATION DATA

The terrestrial radiological survey data are presented in units relative to reference as MOB (or multiples of reference). The results can be presented using a variety of symbols, sizes, and colors. The object is to present the data in the most meaningful way possible that is easy to interpret. While several software packages offer canned statistical color ramping of the results, these are tools for investigating the data that rarely produce optimal results.

What is important is not the MDA of the instrument but the ability of the instrument to detect/map contamination levels that are above local background. Large areas often have significant variability in background signature due to differences in geology and soil type. This background variation is plainly evident at the NNSS in a post-processed aerial radiological survey [4].

Using a GIS [ArcGIS (ESRI)] as a tool to help interpret the data presents several opportunities to analyze and/or accentuate the spatial variability inherent in the radiological survey data.

The ability of the visual cortex to recognize spatial patterns is well documented [5, 6]. Some have suggested that the human brain's pattern recognition capability surpasses other statistical quantitative methods. Tufte [7] suggests that the patterns and trends in point data are often only discernable if presented in data-rich illustrations, figures, or maps. ArcGIS provides the user with a variety of tools that can be used in a dynamic manner to graphically depict the radiological survey data. The objective is to maximize color contrast to present the variation in the data in a manner so that it can be detected and interpreted with the human eye.

Initial presentations of the data can be easily performed with a number of statistical methods to optimize histogram groupings. No one method works "best" in all situations. Reliance on statistical methods to perform histogram slicing may not produce the best results but may serve to provide a good initial starting point. Figure 1 presents a ¼ standard deviation histogram slice applied with a two-tone end-member color scheme. This numerical statistic method is biased by the amount of non-contaminated data points in the sample set and therefore does not optimize the ability to visually interpret the variation in the data. This is true with other statistical methods such as equal interval, quantile, or geometric statistical approaches.

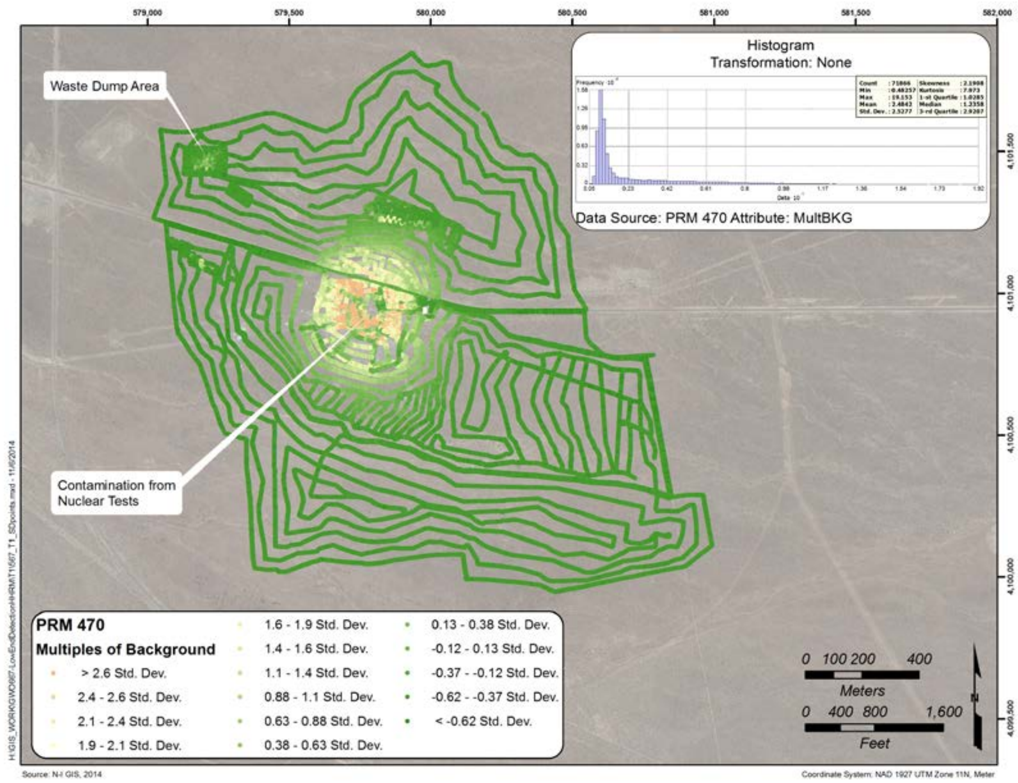

Fig. 1. PRM 470 survey data binned by standard deviation [8].

One can use initial statistical binnings as a starting point in analyzing the data. A way to investigate the instrument response variation in conjunction with the spatial variability is to analyze regular value bins of the histogram. The data presented in Figure 2 have been binned into 120 equal interval bins. Visually investigating the spatial distribution of each of the bins can reveal much information with regard to the ability of the instrument to detect contamination above background levels.

Incrementally moving from lowest to higher normalized instrument response values will reveal the spatial nature of the contamination relative to a spatially variable background. Where a given level exhibits clustering, one can consider that the survey is effectively identifying discrete combinations of location and radiological response.

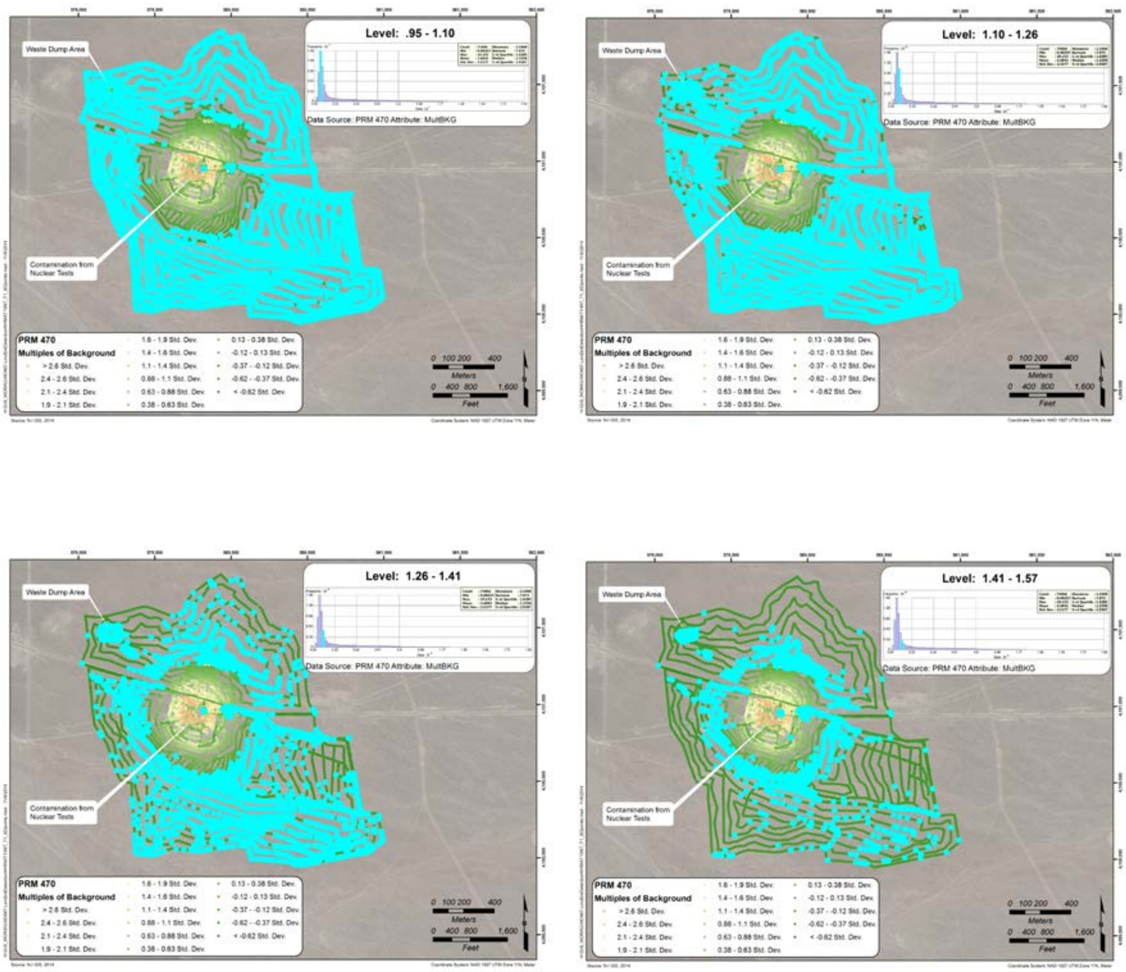

Fig. 2. PRM 470 survey data binned into equal intervals [8].

The 1.10 – 1.26 MOB level exhibits a slight clustering tendency in the Waste Dump area. This area has a lower background associated with it than the rest of the site [4]. The 1.26 – 1.41 MOB level produces a distinct cluster around the Waste Dump and exhibits a clustering tendency around the area affected by nuclear tests. Therefore, one could say that the ability of the survey to detect contamination distinctly different from background could be as low as 1.26. It may be higher in the area of nuclear test contamination (~ 1.4 MOB) due to the fluctuation in local background.

KRIGING/INTERPOLATION USING GEOSTATISTICS

There are numerous ways to interpolate data, and it is important to evaluate the available methodologies and compare them to a particular task. Methods that use kriging and other predictive interpolators create surfaces that have areas where values are based on mathematical assumptions of trends in the data. For the purposes of radiological characterization, the goal is not to create estimates of values that may be greater than the particular instrument sensitivity as is produced by a predictive interpolator, but to maintain as much of the original data as possible. An exact linear interpolator such as a TIN (triangular interpolated network) does not effectively model non-constant rates of instrument response over distance. The value of any TIN surface between data points is calculated linearly based on its nearest neighbors. An

Inverse Distance Weighted interpolator is an exact deterministic interpolator that does not smooth the data. It will consider values beyond nearest neighbors but will decrease a point's value contribution to a given interpolated point based on the square of its distance. It is important to consider these points that are at a distance from the interpolated point because the data show a more or less logarithmic increase in instrument response as one approaches ground zero.

To best define the neighborhood that should be used in interpolation, one should look at the data themselves to analyze what kind of spatial variability is inherent in the original data. Two methods for looking at the spatial dependence between two points are the semivariogram and the variogram. Both depict the influence of distance on the variability between points.

In a semivariogram, one looks for the distance at which a change in variance becomes constant. In Figure 3, one can see that as the distance between two point pairs increases to beyond about 150 to 200 meters, the variance is no longer related to distance.

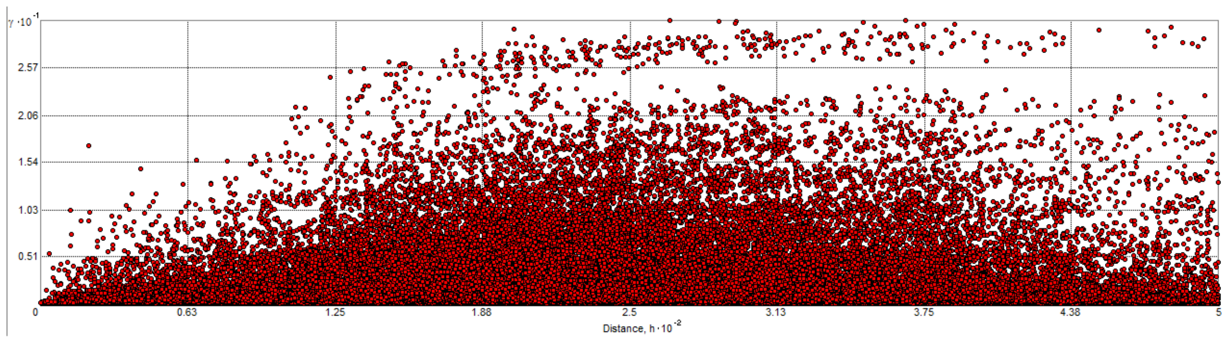

Fig. 3. Semivariogram of Cabriolet site.

Similarly, an investigation of the variograms of radiological survey data for two sites shows a declining influence of distance on value after about 100 to 200 meters. In Figure 4, a steady slope indicates that the rate of increase in variance remains constant or is not dependent upon distance.

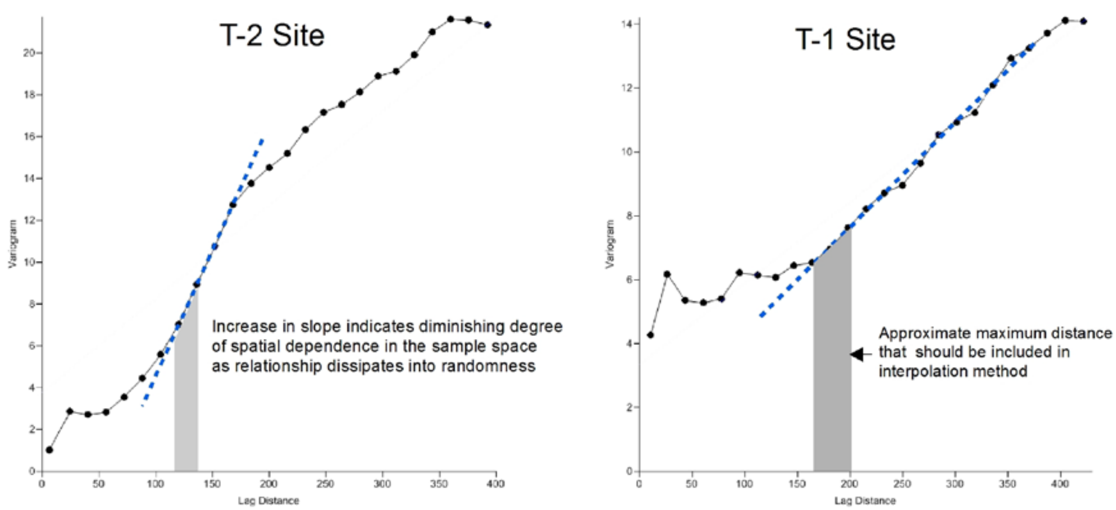

Fig. 4. Variogram, T-1.

Based on the analysis of terrestrial radiological surveys over a number of sites across the NNSS, variograms and semivariograms of the data indicate that the maximum separation distance where one reading has an influence on another is approximately 200 meters. This value is used to restrain the neighborhood of values used in defining the inverse distance-weighted interpolation. The directional bias (two-dimensional anisotropy) of the contaminant distribution is evaluated and incorporated into the interpolation method as well. The major and minor semiaxes used to define the search neighborhood are modified to approximate the contaminant distribution anisotropy (e.g., elongated in an east–west or a northeast–southwest trend).

Once the data have been used to create a surface, it is important to remember that any displays of the data should consider the original tenet discussed earlier. That is, the goal is not only to be able to present the dynamic range of the data themselves, but also to be able to differentiate between areas that represent uncontaminated zones of instrument scatter and field background conditions versus areas that represent an emergent signal pattern above this background noise.

The tools available to the use in ArcGIS (ESRI) for displaying surfaces are more robust than those available for displaying cardinal point values. One can use any of the default contrast-enhancing tools, such as histogram stretches, or a number of statistical binning algorithms; or one can perform custom histograms stretches to accentuate the distribution patterns in discrete ranges. All of these tools can contribute to creating useful displays and interpretations of the data.

An effective visual display of the data can be facilitated by stretching the dynamic range of the display at the survey values that represent the low end of contaminant discrimination from background and at the decision level itself.

In Figure 5, the display has been optimized to show survey-level variation at responses near the field contamination discrimination level. Scatter, speckles, or noise is present at lower levels, indicating that those values are representative of instrument variability in the presence of background. At some level of contamination, the instrument response is no longer masked by the instrument variability. Note that in the area of lowest response (light blue shades), the areas display a mottled pattern. This is indicative of a condition where the sensor variability at or near background conditions is producing an irregular pattern. As the level goes from 1.1 MOB to 1.3 MOB, the clustering tendency is stronger, and the mottling effect is no longer dominating the shape of the cluster. This is indicative of the survey detecting contamination (increase in sensor response) at responses significantly different the normal background response. The effective field MDA, as illustrated by sensor response data patterns, is in the 1.2 – 1.3 MOB level range at this particular site.

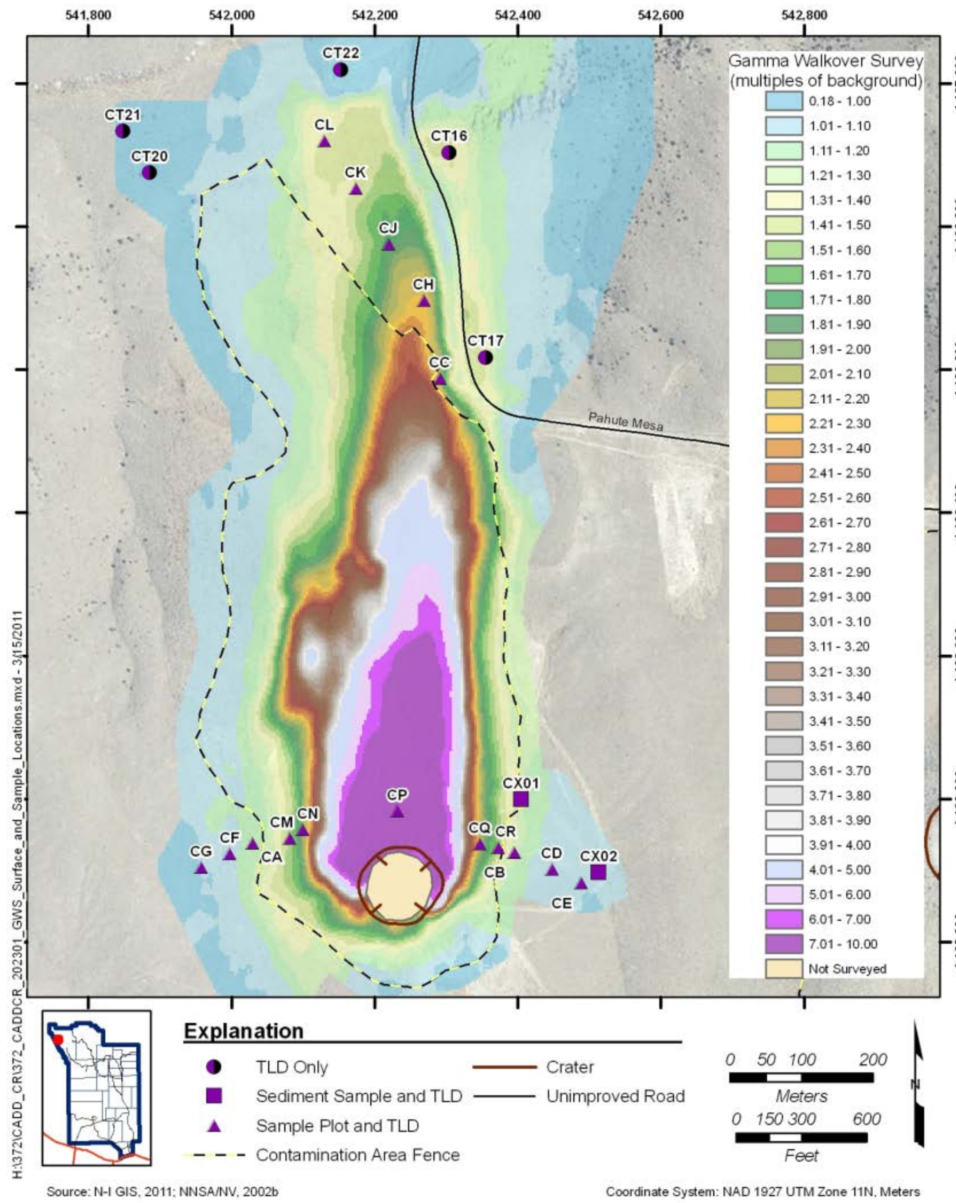

Fig. 5. Noise and clustering in a contaminated area [8, 9].

Determining Extent of Contamination

The functional reason for creating a surface is to derive the area of contamination that is above the given decision level. In order to do so, ordered pairs of survey value versus dose are used. The survey surface value at each sample location is extracted using ArcGIS. These values are correlated to total effective dose. The strength of this relationship is affected by the number of samples and the total variance amongst samples.

A correlation is established between total effective dose and multiples of reference using a calculated 95% lower confidence limit as described in EPA unified guidance (Figure 6).

$$LCL_{1-\alpha} = \hat{x}_0 - \sqrt{2s_e^2 \cdot F_{1-2\alpha,2,n-2} \cdot \left[\frac{1}{n} + \frac{(t_0 - \bar{t})^2}{(n-1)\cdot s_t^2}\right]}$$

$$UCL_{1-\alpha} = \hat{x}_0 + \sqrt{2s_e^2 \cdot F_{1-2\alpha,2,n-2} \cdot \left[\frac{1}{n} + \frac{(t_0 - \bar{t})^2}{(n-1)\cdot s_t^2}\right]}$$

Fig. 6. Equations 21.24 and 21.25 from EPA *Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities: Unified Guidance*, March 2009 [9].

One can then calculate the survey surface value associated with the decision level. Using GIS, it is then a simple task to extract an isopleth representing the regulatory dose limit (Figure 7).

This method provides a way to quantify the area where decisions regarding remediation are necessary.

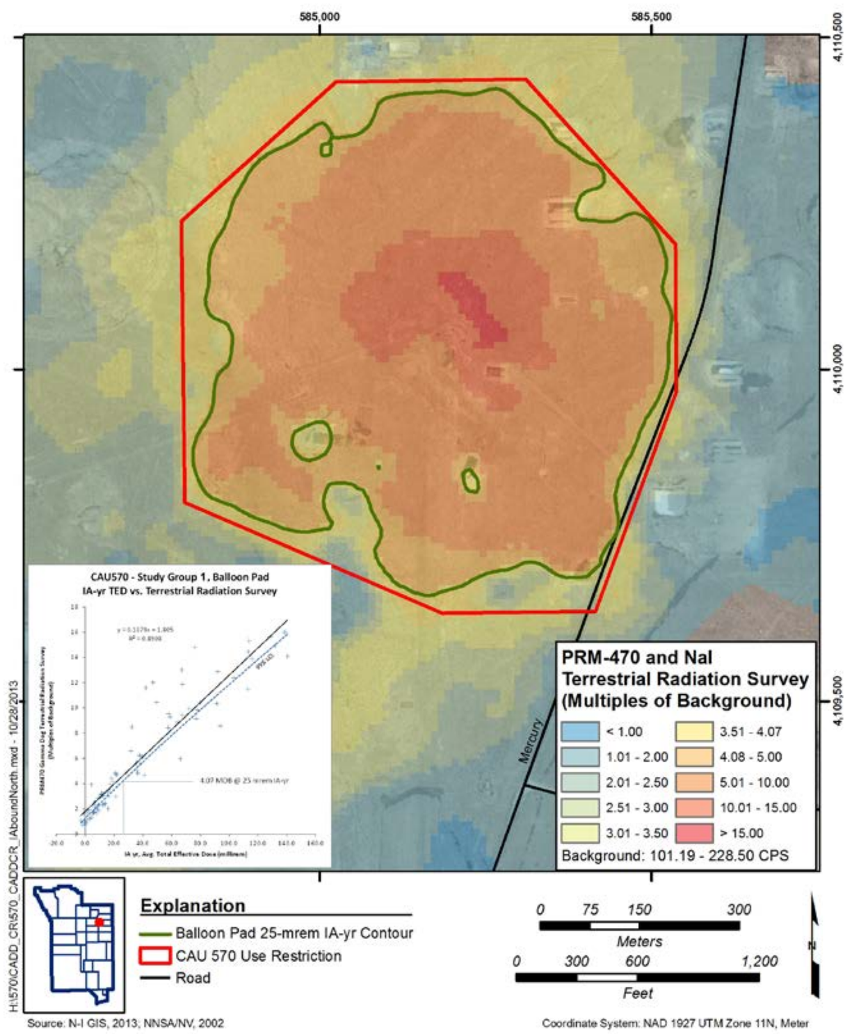

Fig. 7. Contamination level isopleth at specific dose [8, 10].

CONCLUSION

The ability to produce a large number of sensor readings in a very short time enables one to characterize large areas at reasonable cost. These survey results can be displayed and analyzed using a GIS to manipulate the data. These manipulations can reveal low-end contamination detectability levels in the presence of background even where local background is somewhat variable. Visual pattern recognition is a powerful tool that can greatly enhance the interpretability of survey results. The human brain's ability to recognize patterns inherent in the data is a powerful analysis tool.

This methodology takes in to account the inherent variability of the study area rather than assuming homogeneity. The advantage of applying GIS geostatistics to these types of large surveys (large in number of data points) is that the classification of impacted areas for survey planning can be evaluated readily and effectively. This could lead to substantial cost savings in optimizing survey planning.

REFERENCES

1. *Federal Facility Agreement and Consent Order.* (1996, as amended March 2010). Agreed to by the State of Nevada; U.S. Department of Energy, Environmental Management; U.S. Department of Defense; and U.S. Department of Energy, Legacy Management. Appendix VI, which contains the Soils Sites Strategy, was last modified June 2014, Revision No. 5.

2. Abelquist, E.W., W.S. Brown, G.E. Powers, and A.M. Huffert. (1998). *Minimum Detectable Concentrations with Typical Radiation Survey Instruments for Various Contaminants and Field Conditions*, NUREG-1507. Washington, DC: U.S. Nuclear Regulatory Commission, Office of Nuclear Regulatory Research.

3. Desnoyers, Y., and D. Dubot. (2014). "Characterization of Radioactive Contamination Using Geostatistics." *Engineering International*. May.

4. National Security Technologies, LLC. (2011). Written communication regarding the Data Note "Re-processing of Legacy NT94 Data To Estimate Site-Wide Natural Activity" that includes text, figures, and tables, 2 February. Las Vegas, NV.

5. Margolis, H. (1987). *Patterns, Thinking and Cognition: A Theory of Judgment*. Chicago, IL: University of Chicago Press.

6. "Pattern Recognition Primer" from the Wolfram Demonstrations Project. Available at http://demonstrations.wolfram.com/PatternRecognitionPrimer.

7. Tufte, E.R. (1983). *The Visual Display of Quantitative Information* (2nd ed.). Cheshire, CT: Graphics Press, ISBN 0-9613921-4-2.

8. Navarro-Intera Geographic Information Systems. (2012, 2013, 2014). ESRI ArcGIS Software.

9. U.S. Environmental Protection Agency. (2009). *Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities: Unified Guidance*, EPA 530/R-09-007. Washington, DC: Office of Resource Conservation and Recovery.

10. U.S. Department of Energy, National Nuclear Security Administration Nevada Operations Office. (2002). *Nevada Test Site Orthophoto Site Atlas*, DOE/NV/11718--604. Aerial photos acquired Summer 1998. Prepared by Bechtel Nevada. Las Vegas, NV.