Environmental Security Technology Certification Program

ESTCP

Anomaly Density Mapping: Lowry AGGR Site Demonstration Report

ESTCP Project # 200325

Final Report
09/30/2006

Sean A. McKenna
Sandia National Laboratories

Brent Pulsipher and Brett Matzke
Pacific Northwest National Laboratory
PNNL Clearance # PNNL-16106

Distribution Statement A: Approved for Public Release, Distribution is Unlimited

Table of Contents

1 Introduction 1
1.1 Data Set Provided for Analysis 2
1.2 Modifications to Data Set 2
2 Full Data Set Analysis 5
2.1 Density Estimation (CM1) 5
2.1.1 Spatial Averaging of Transect Data 5
2.1.2 Variogram Analysis 6
2.1.3 Estimated Anomaly Density 7
2.2 Density Estimation (CM2) 8
2.2.1 Determination of Background Anomaly Density 8
2.2.2 Residual Variogram 10
2.2.3 Residual Estimates 11
2.2.4 Estimated Anomaly Density 12
2.3 Target Area Boundary Delineation (CM1) 13
2.3.1 Distribution Analysis 13
2.3.2 Density Threshold Contour 16
2.4 Target Area Boundary Delineation (CM2) 17
3 Jackknife Validation 19
3.1 Density Estimation (CM1) 19
3.1.1 Variogram Analysis 20
3.1.2 Estimated Anomaly Density 20
3.2 Density Estimation (CM2) 21
3.2.1 Determination of Background Anomaly Density 21
3.2.2 Residual Variogram 22
3.2.3 Residual Estimation 23
3.2.4 Anomaly Density Estimation 24
3.3 Density Estimation Evaluation (CM1 and CM2) 25
3.4 Target Area Boundary Delineation (CM1) 26
3.4.1 Distribution Analysis 26
3.4.2 Density Threshold Contour 32
3.5 Target Area Boundary Delineation (CM2) 34
4 Conclusions 36

Figures

Figure 1. Aerial view of the Air-to-Ground Gunnery Range (AGGR). The left-hand view shows the AGGR in its entirety, while the right-hand view zooms in on the area sampled in Phase 1 sampling. 1
Figure 2. Original geophysical anomaly locations (top) and the same locations after rotation and translation to the local coordinate system (bottom). The magenta line in the bottom image is the site boundary used in the anomaly density estimation. 4
Figure 3. Average anomaly density per cell (left image) and per acre (right image for each 150m^2 estimation cell that is intersected by a sampling transect. The X and Y scales are in meters. 6
Figure 4. Omnidirectional variogram for the anomaly density values at the estimation cell scale. This variogram is calculated using all anomaly data and CM1 (no trend). 6
Figure 5. Estimated number of anomalies per 150m^2 cell (left) and per acre (right) created using all transect data under CM1. 7
Figure 6. Average values of cell density for each geophysical transect as a function of the Y coordinate (northing) of each transect. The red line is the cubic function fit to the averaged data. 8
Figure 7. Trend model residual values as a function of Y (Northing) coordinate. 9
Figure 8. Spatial distribution of residuals between the trend model and the number of anomalies per 150m^2 cell. 10
Figure 9. Experimental and model variograms of the residuals between measured densities and the background trend model (CM2) using all transect data. The variograms calculated and modeled in the east-west and north-south directions are shown in the top and bottom images respectively. 11
Figure 10. Estimation of residual values between the trend model and the number of anomalies in each 150m^2 cell. 12
Figure 11 Estimated number of anomalies created by taking into account the trend (CM2) per 150m^2 cell (left) and per acre (right). 12
Figure 12. Frequencies of site densities for different window sizes. 13
Figure 13. Flagged Areas for 45.72m (300 ft.) diameter window under CM1, 70 anomalies/acre background density, and 95% confidence an area is greater than background density. "Flagged" coordinates along transects are identified with a small box. 14
Figure 14. Possible target area boundaries under CM1. 15
Figure 15. Definition of target areas (red) and background areas (grey) based on the estimated anomalies per acre under CM1 as shown in Figure 5. 16
Figure 16. Histogram and normal probability plot of the residual data as determined from all transect data 17
Figure 17. Delineation of target areas using the residuals about the background density trend (CM2). The image on the left is for a threshold of -3.5 and the image on the right is for a threshold of -2.5. 18
Figure 18. Average anomaly density (per cell) for each 150m^2 estimation cell that is intersected by a sampling transect. 19
Figure 19. Omnidirectional variogram for the anomaly density values from one-half of the transects at the estimation cell scale for the CM1 model. 20

Figure 20. Estimated number of anomalies per 150m^2 cell (left) and per acre (right) using one-half the transects under CM1. 21
Figure 21. Average values of cell density for each geophysical transect as a function of the Y coordinate (northing) of each transect. The red line is the cubic function fit to the averaged data from half of the data as contained in set 1. 22
Figure 22. Experimental variograms and models for one-half the detrended anomaly count data (CM2) at the estimation cell scale. The north-south variogram is shown in the top and the east-west variogram in the bottom. 23
Figure 23. Estimation of residual values calculated using half of the transects as contained in set 1 between the trend model and the number of anomalies in each 150m^2 cell. 24
Figure 24. Estimated number of anomalies created by from one-half of the transect data taking into account the trend (CM2) per 150m^2 cell (left) and per acre (right). 24
Figure 25. Comparison of measured and estimated anomaly density at the Set 2 measurement locations. Results are shown for the CM1 (left image) and CM2 (right image) estimations. 26
Figure 26. Set 1 – every other transect. Frequencies of site densities for different window sizes. 27
Figure 27. Set 1 – every other transect. Flagged Areas for 91.44m (300 ft.) diameter window, 60 anomalies/acre background density, and 95% confidence an area is greater than background density. "Flagged" coordinates along transects are identified with a small box. Results are for CM1. 28
Figure 28. Set 2 – every other transect. Frequencies of site densities for different window sizes. 29
Figure 29. Set 2 – every other transect. Flagged Areas for 91.44m (300 ft.) diameter window, 65 anomalies/acre background density, and 95% confidence an area is greater than background density. "Flagged" coordinates along transects are identified with a small box. Results are for CM1 30
Figure 30. Set 1 – every other transect. Possible target area boundaries under CM1. ... 31
Figure 31. Set 2 – every other transect. Possible target area boundaries under CM1. ... 32
Figure 32. Definition of target areas (red) and background areas (grey) based on the estimated anomalies per acre under CM1 as shown in Figure 20. 33
Figure 33. Histogram and normal probability plot of the residual data as determined from the Set 1 half of the transect data 34
Figure 34. Delineation of target areas using the residuals about the background density trend (CM2). The image on the left is for a threshold of -3.5 and the image on the right is for a threshold of -2.5 36

Tables

Table 1. Results of target area boundary delineation. 37

1 Introduction

The aim of this report is to show the methods and results of using the tools developed by Pacific Northwest National Laboratory (PNNL) and Sandia National Laboratory (SNL) for mapping the geophysical anomaly density at the Former Lowry Bombing and Gunnery Range (FLBGR) site located east of Denver, Colorado. PNNL and SNL have developed the tools under SERDP to support detection, delineation, density estimation, and mapping of unknown target areas where UXO are most likely. Some of the methods have been deployed via the Visual Sample Plan (VSP) software.

The specific area analyzed is within the Air-to-Ground Gunnery Range (AGGR), which is approximately 3 square miles in size. In Phase 1 sampling of the AGGR, approximately 1/3 of the site was sampled. A three-coil EM-61 instrument was used to collect data along 20 transects running approximately east-to-west between two creek beds. Transects were 3 meters wide, and spaced approximately 200 feet apart from their centers. Figure 1 shows an aerial view of the AGGR with transect locations marked on the map.

PNNL and SNL performed two different sets of analyses. In the first analysis, the full data set is used to create maps showing the anomaly density and the probability of any area in the site being above a threshold background value. In the second analysis, the data set is split into two data sets so that one data set contains the odd-numbered transects, and the other contains the even-numbered transects. This split of the data set allows for a jackknife analysis of the data where one data set is used to recreate the estimates made with the full data set, and the other data set is used to evaluate those estimates.

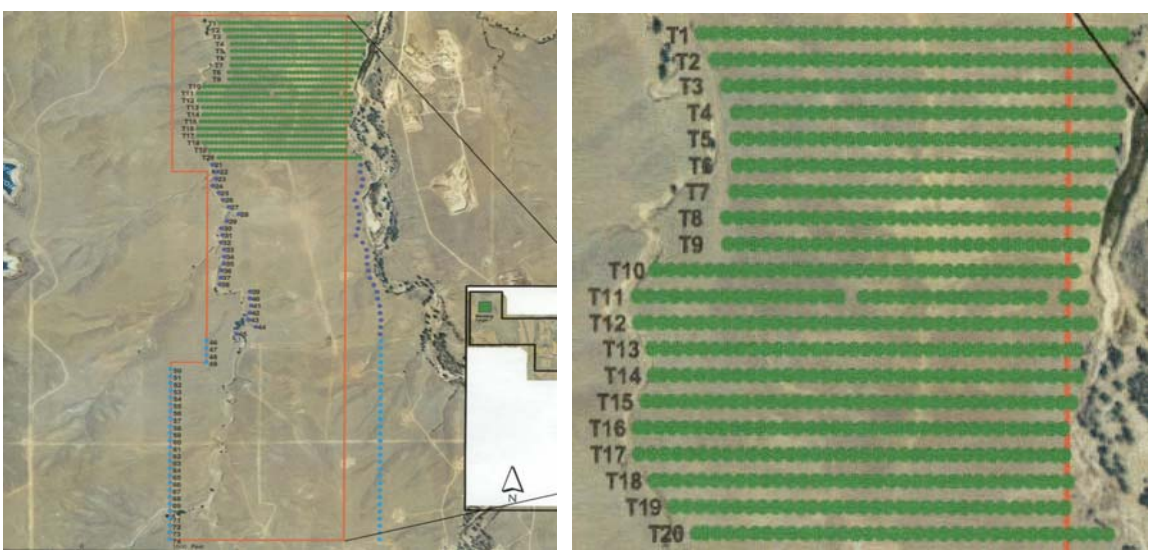


Figure 1. Aerial view of the Air-to-Ground Gunnery Range (AGGR). The left-hand view shows the AGGR in its entirety, while the right-hand view zooms in on the area sampled in Phase 1 sampling.

Additionally, two different conceptual models of the spatial variation of the anomaly density are investigated. These two models are: 1) the gradual increase of the anomaly

density from north to south is due to a large target area with elevated densities near the center of the target area at the southern edge of the site and 2) that the majority of the increase in anomaly density from north to south is not related to the presence of a target area, but is due to some unrelated, perhaps natural, increase in the density of conductive material in the soil. Under conceptual model 1 (CM1) the goal is to identify what portion of the site is above a chosen threshold density and, depending on the threshold, this target delineation may delineate a large-portion of the site. Under conceptual model 2 (CM2) the same goal of target delineation is pursued, yet the delineated area is expected to be much more subtle in that it will be a relatively small area superimposed on the larger low-frequency trend in the background anomaly density.

1.1 Data Set Provided for Analysis

The data used for this analysis were transmitted via email from Bart Hoekstra of Sky Research on March 17, 2005 in the *AGGR_Phase1_Targets_1_edt.csv* file. This file contains the easting and northing locations of the geophysical "targets" as surveyed at the AGGR site using a three-coil EM-61 instrument. Here, the term "target" is reserved for locations that represent a region of the site, generally 10's to 100's of square meters with an anomaly density that is significantly above the background anomaly density. The individual point locations identified by the geophysical survey as representative of surficial or subsurface metallic objects are referred to here as "anomalies". Accurate estimation of the spatial distribution of the density of these anomalies across the site and identifying and delineating high density areas are the goals of this work.

Other information regarding the geophysical anomalies at the AGGR site include the amplitude of the geophysical anomaly from the EM instrument and flags indicating whether or not the anomaly was detected by the center coil or one of the outside coils and a code defining the origin of the anomaly in the database. For these analyses, the anomalies used are either termed as "Original Pick in the Database" or "Pick Added". The relative location of each anomaly, center or outside coil, is not taken into account in these analyses and all anomalies are treated equally in the analysis. The amplitudes are not used in this analysis either, only the easting and northing coordinates of each anomaly are used.

1.2 Modifications to Data Set

PNNL and SNL performed several modifications to the AGGR data set to facilitate the analyses described in this report.

1) Transects are modified so that they run in a straight line. In some analyses, each anomaly identified in the geophysical survey is assigned to a transect. There are a total of 20 transects and these are numbered from 1 to 20, south to north, at the site (Figure 1).
2) The locations of the transects were rotated counter-clockwise by 0.31 degrees to make the average alignment of the transects directly east-west. Figure 2 compares the original (top image) and rotated (bottom image) anomaly locations.
3) For some analyses, the anomaly locations were translated from the original coordinate system to a new local coordinate system (bottom image, Figure 2).

The origin (0,0) of the new coordinate system is also the point about which the coordinates were rotated (531,000, 4,385,200). In other analyses, coordinates were multiplied by a constant so that (x,y) coordinates would be in feet.

4) The spatial domain for the analysis is limited by a polygon with irregular shape to take into account the variation in the end points of the different transects. This polygon is shown in the lower image of Figure 2.

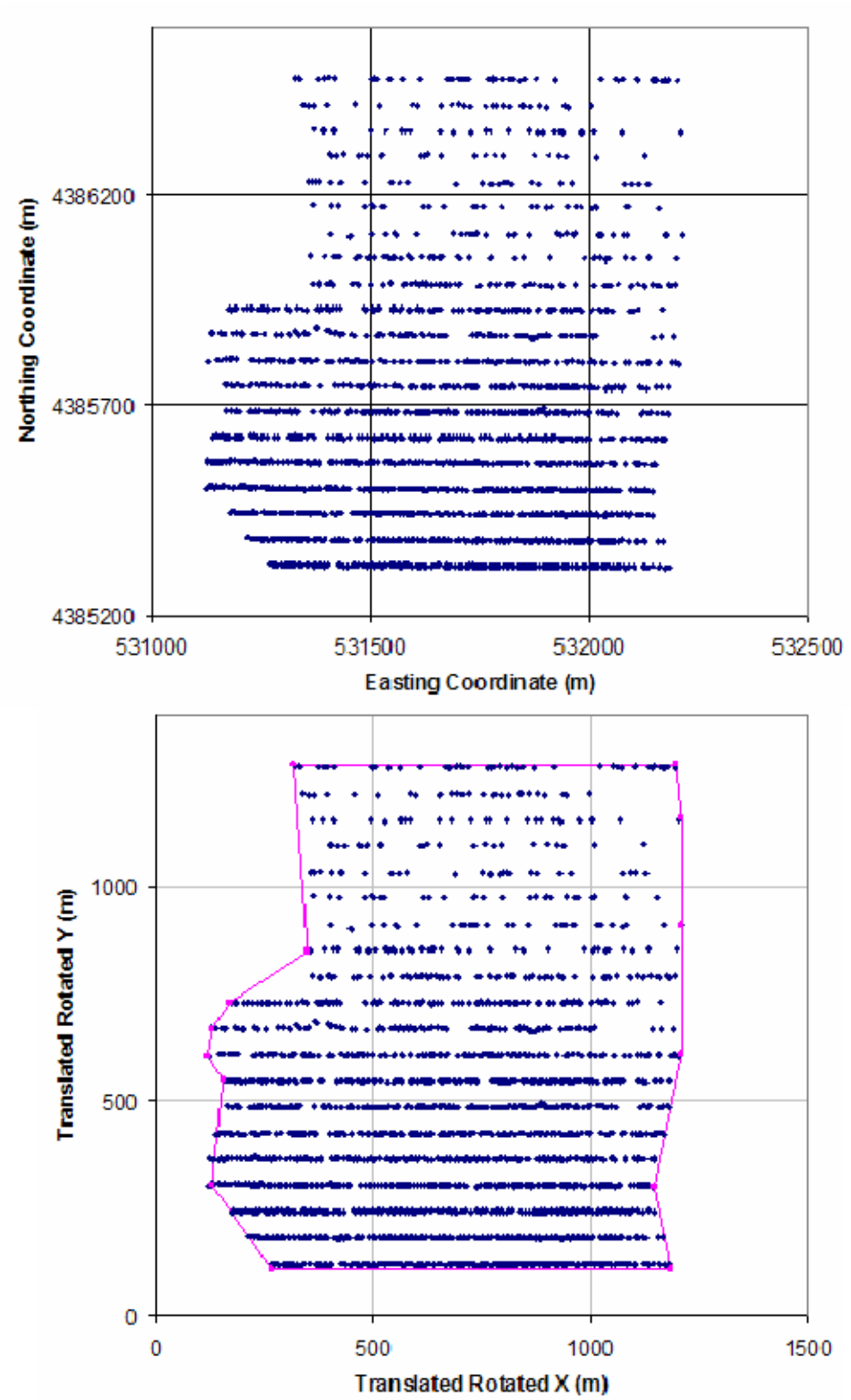


Figure 2. Original geophysical anomaly locations (top) and the same locations after rotation and translation to the local coordinate system (bottom). The magenta line in the bottom image is the site boundary used in the anomaly density estimation.

2 Full Data Set Analysis

Several different analyses are completed using the full anomaly data set. In some analyses, a uniform background density is assumed across the site (CM1). In other analyses a trend function for the anomaly density is determined and used to detrend the data prior to creating the anomaly density estimates (CM2). Detrending is done to make the assumption of a spatially stationary model used in the geostatistical analysis valid. The two sets of estimates under the different conceptual models are compared. The target area delineation is done two ways for both of the conceptual models. The first approach is to simply draw the appropriate contour interval on the estimated anomaly density map and define everything above the threshold as being within the target area. Knowledge of the true location of the contour line defining the target boundary is uncertain. This approach to boundary delineation does not acknowledge this uncertainty. Indicator kriging of the probability of any location on the site being within a target area boundary does allow for the user to define an acceptable level of uncertainty when defining the target area boundary.

2.1 Density Estimation (CM1)

The raw transect data are averaged onto sample and then estimation cells using properties of the Poisson distribution. These spatially averaged data are then used to calculate variograms and create anomaly density estimates across the site area.

2.1.1 Spatial Averaging of Transect Data

The estimates of anomaly density are done on a regular grid where each estimate is the average of a Poisson distribution within each grid cell. The average value completely defines the Poisson distribution for that cell. The estimation cells used here are 30m long (east-west) and 5m wide (north-south) so that each cell has an area of 150m^2. There are a total of 7469 estimation cells inside the site domain for a total area of 1.12E+06 m^2. Each geophysical transect is 3m wide and is discretized into 30m segments (area of each segment is 90m^2). A moving window that is three segments (90m) long is used to calculate the average anomaly density for the center cell in the window. This average density value is assigned to the 30x3m sampling cell at the center of the window being moved along the transect. The average density for this 90m^2 transect cell is assumed to be the appropriate density for the corresponding 30x5m estimation cell, and these data are used in the analyses below. Spatially averaged values calculated from the moving window along the transect are used here to provide more representative samples than are acquired directly by the geophysical surveying. The inherent assumption is that the average value calculated in the moving window along the transect is representative of the true density value that would be measured if the transects were as wide as the estimation cells, i.e., 5m wide transects instead of 3m wide transects. The locations and values of the averaged density values are shown in Figure 3.

The averaged anomaly density values show the significant increase in density from the north to the south across the site (Figure 3). The transects have a variable spacing between 55 and 65 meters (180-210 feet).

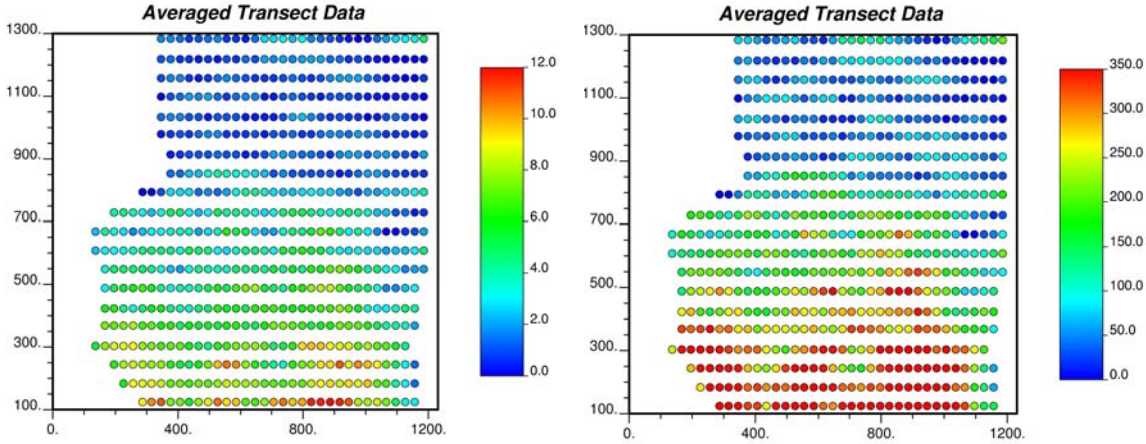


Figure 3. Average anomaly density per cell (left image) and per acre (right image for each 150m^2 estimation cell that is intersected by a sampling transect. The X and Y scales are in meters.

2.1.2 Variogram Analysis

The data shown in Figure 3 are transformed from anomalies/acre to anomalies/estimation cell for analysis of variograms and the following spatial estimations. A single variogram considering spatial correlation in all orientations at once (omnidirectional) is calculated using the cell average anomaly densities and is shown in Figure 4. This variogram is fit with a Gaussian model having a nugget value of 2.0, a sill of 20.0 and a range of 750m.

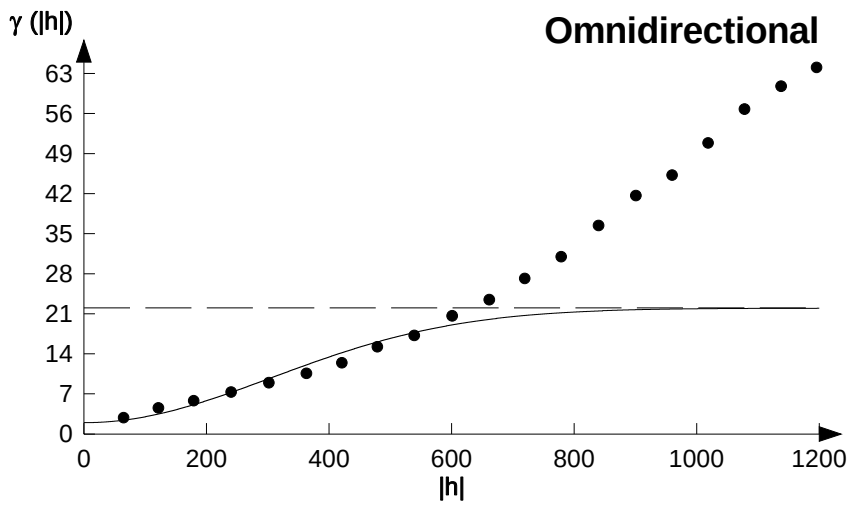


Figure 4. Omnidirectional variogram for the anomaly density values at the estimation cell scale. This variogram is calculated using all anomaly data and CM1 (no trend).

The most notable feature of the variogram shown in Figure 4 is that the γ values of the experimental points continue to increase with increasing values of the separation distance, *h*. These values go well above the theoretical sill value that is equal to the variance of the anomaly density data set (horizontal dashed line in Figure 4). This continual increase in variance with increasing separation distance is indicative of a trend

in the sampled values. For this data set, the continually increasing variogram is due to the strong north to south increase in the density values (see Figure 3). This behavior in the variogram indicates that the assumption of second-order stationarity invoked in most geostatistical estimation studies does not hold for this data set.

The degree to which kriging produces accurate estimates in the presence of strongly non-stationary data is case-specific. The use of the ordinary kriging with a fairly dense data set in the presence of a strong spatial trend can produce accurate estimates and that approach is used here as conceptual model 1 (*CM1*). In addition to this approach, the data are used to identify a trend model, each data point is then subtracted from this trend, ordinary kriging is used to estimate the spatial distribution of the residuals, and the trend model is added back to the estimated residual values to produce the final estimates of the anomaly density. This explicit consideration of a trend in the background anomaly density is labeled conceptual model 2 (*CM2*). This second approach uses the stationarity of the residuals about the trend model to produce accurate estimates of the anomaly density.

2.1.3 Estimated Anomaly Density

The ordinary kriging algorithm is used to create estimates of the number of anomalies within each of the 30x5 meter cells. The raw data exhibiting the strong spatial trend of increasing density from north to south are used directly.

Estimation results are shown with two different color scales in Figure 5. The image on the left shows the estimated average number of anomalies within each 150m^2 cell. The image on the right shows the same data after scaling them to represent the expected number of anomalies per acre.

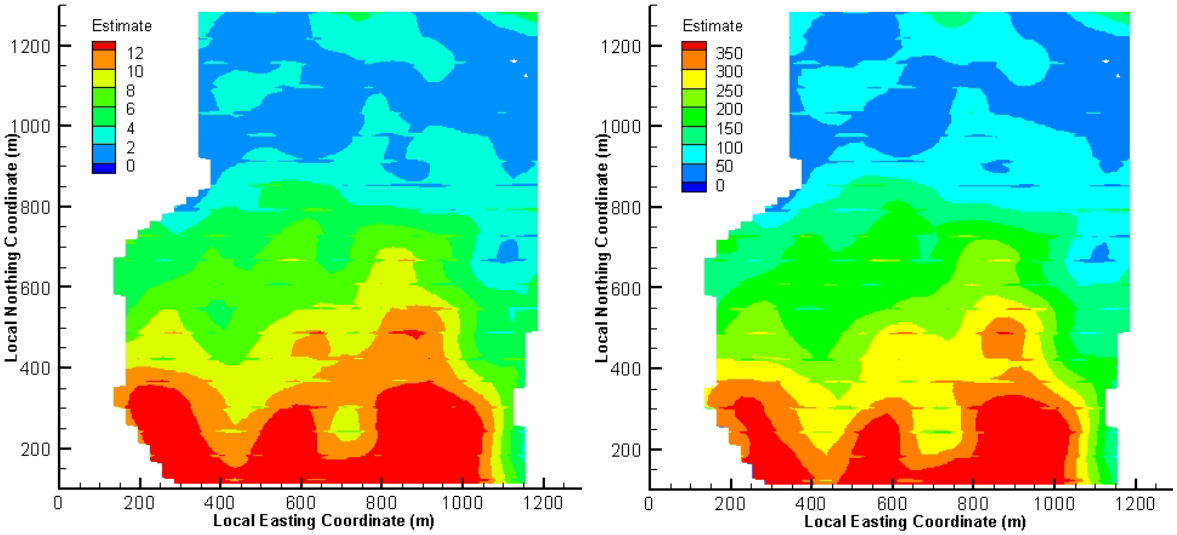


Figure 5. Estimated number of anomalies per 150m^2 cell (left) and per acre (right) created using all transect data under CM1.

2.2 Density Estimation (CM2)

In CM2, the large-scale trend in the anomaly density across the site is not assumed to be due to the presence of a target area, but to some other factor. The spatially averaged data as calculated for CM1 above are used to identify a large scale trend in the anomaly densities across the site and residual values between this large-scale trend and the measured data are used in the variogram modeling and the spatial estimation. The final anomaly density estimate is the sum of the background trend model and the spatially varying residual estimates.

2.2.1 Determination of Background Anomaly Density

One goal of this analysis is to map the edge of a target area. In order to do this, it is necessary to define the background density above which a region would be classified as a target area. This definition of the density threshold is especially difficult at the AGGR site where a low frequency trend in the density anomaly is present. The average anomaly density per estimation cell (150m^2) was calculated along each transect and these average values are graphed in Figure 6 as a function of the northing coordinate of the center of each transect.

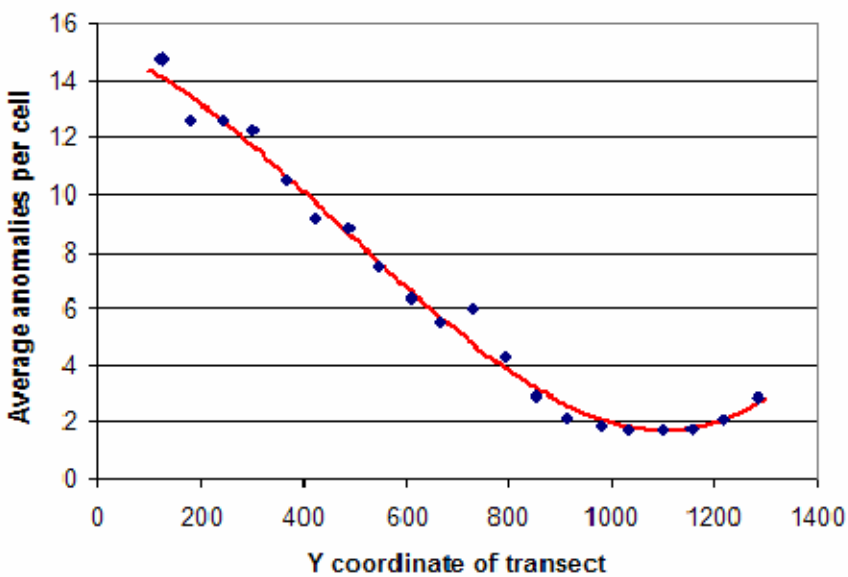


Figure 6. Average values of cell density for each geophysical transect as a function of the Y coordinate (northing) of each transect. The red line is the cubic function fit to the averaged data.

The strong increase in anomaly density to the south of the site could be due to a large target area centered near the southern edge of the site, or several target areas in close proximity of each other. This could also be due to the transects being collected on the northern edge of a very large target area and the ordnance scrap associated with the target area is causing the increase in anomaly density. Given the data shown in Figure 6 this target area would have to stretch across the entire length of the geophysical transects, roughly 1000m, and the target area would have a diameter greater than 2000 m in the north-south direction. As another explanation, there could be a natural variation in the soils at the site that is causing the increase in anomaly density from north to south. In

this case, the trend model shown in Figure 6 may actually define the "background" anomaly density at the site.

The function shown in Figure 6 is a cubic model fit to the anomaly density data averaged across each transect. The form of this model is:

$$Y = Y_0 + AX + BX^2 + CX^3$$

where Y represents the anomaly density and X represents the northing coordinate. The parameters Y_0, A, B and C were determined to provide the best fit to the averaged data. These values are Y_0 = 15.26; A = -6.83E-03; B = -2.1E-05; C = 1.46E-08 and provide a fit to the data with an r^2 value of 0.989.

This function provides a model of the average variation of anomaly density across the site. The difference between each of the average values in the cells intersected by a geophysical transect and this trend are the residuals between the measured values and the trend model. These residuals are calculated as residual = trend – measured such that in locations where the trend overestimates the measured values, the residuals have positive values and vice versa. The values of the residuals as a function of the northing coordinate are shown in Figure 7.

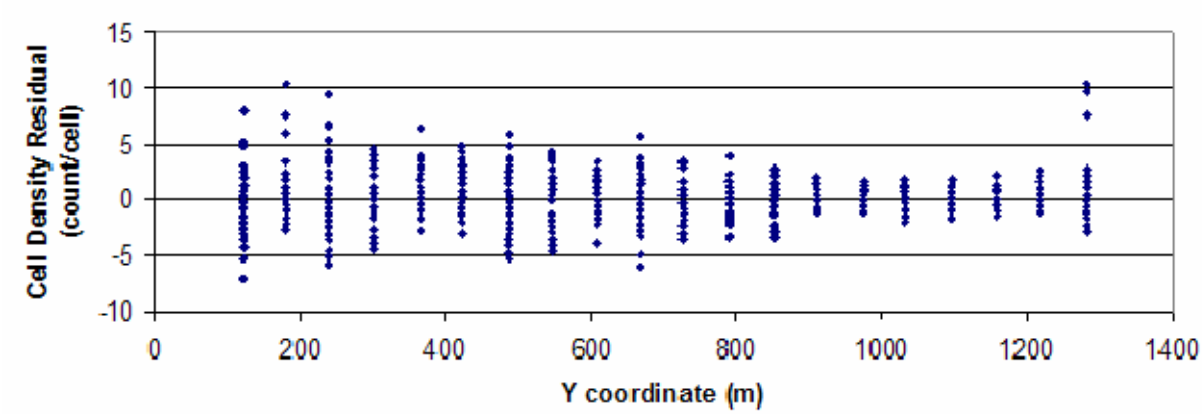


Figure 7. Trend model residual values as a function of Y (Northing) coordinate.

The spatial distribution of the residual data are shown in Figure 8. Of particular interest are the areas of the largest negative residuals near the southern end of the site. Negative residuals indicate locations where the trend model of background anomaly density underestimated the actual measured data – in other words – where the measured anomaly density is higher than expected. If in fact, the trend model represents natural variation in background density, then the areas of large negative residuals may indicate the presence of a target area superimposed on this varying background density.

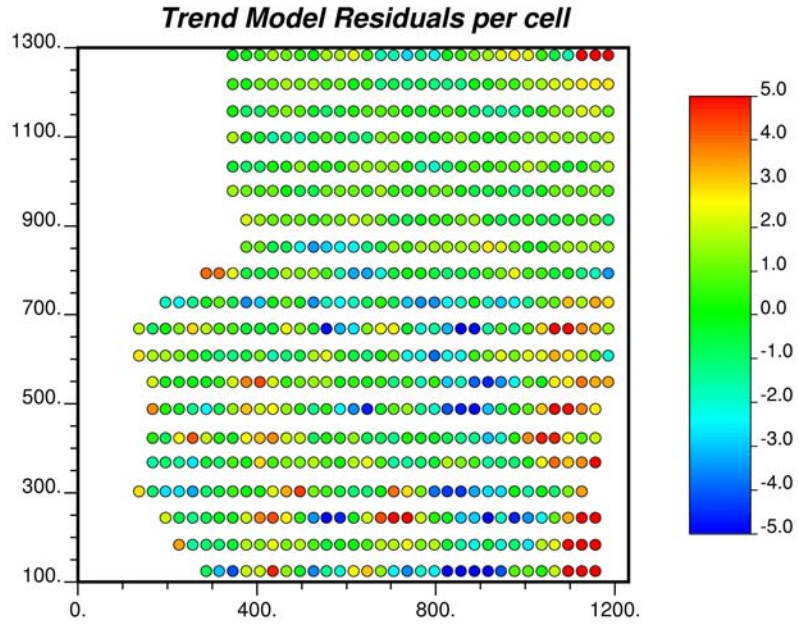


Figure 8. Spatial distribution of residuals between the trend model and the number of anomalies per 150m^2 cell.

2.2.2 Residual Variogram

The residual values determined as the measured values subtracted from the trend model are used as input to variogram calculation and modeling. If the trend model provided a perfect model of the spatial variation of the anomaly densities, then the residuals in the north-south direction would be due solely to random variation about this trend and they would display no spatial correlation. Examination of the bottom image in Figure 9 shows that the residuals have little to no spatial correlation in the north-south direction. The black dots of the experimental variogram show a nearly constant level of variation across all separation distances. However, in the east-west direction (top image of Figure 9), the experimental variogram of the residuals is well-fit by a Gaussian variogram model with a range of 127.5 meters. Additionally, a weak hole-effect with a wavelength of approximately 300m is also noted in the variogram. In summary, the background trend model is able to account for nearly all the spatial correlation of the anomaly density in the north-south direction, but is not able to account for the spatial variation of the anomaly density in the east-west direction.

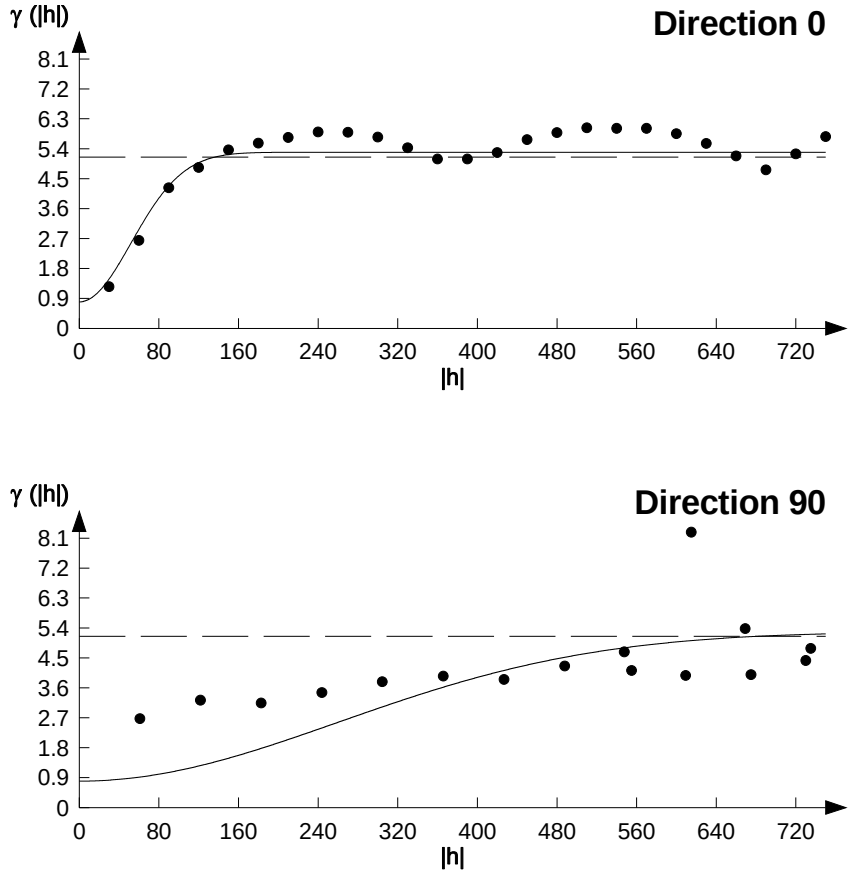


Figure 9. Experimental and model variograms of the residuals between measured densities and the background trend model (CM2) using all transect data. The variograms calculated and modeled in the east-west and north-south directions are shown in the top and bottom images respectively.

2.2.3 Residual Estimates

Ordinary kriging is used to estimate the residual values at all locations in the domain. These estimated values are shown in Figure 10. Of particular interest in Figure 10 are the areas of negative residual estimates. Negative residuals indicate regions where the background trend model is underestimating the actual measured anomaly density. These areas may require additional consideration as potential target areas. The negative residual areas may simply be a result of the low-order trend model not being capable of fitting all of the variation in the observed anomaly density. However, they are the locations where additional sampling is most likely to confirm or deny the existence of CM2.

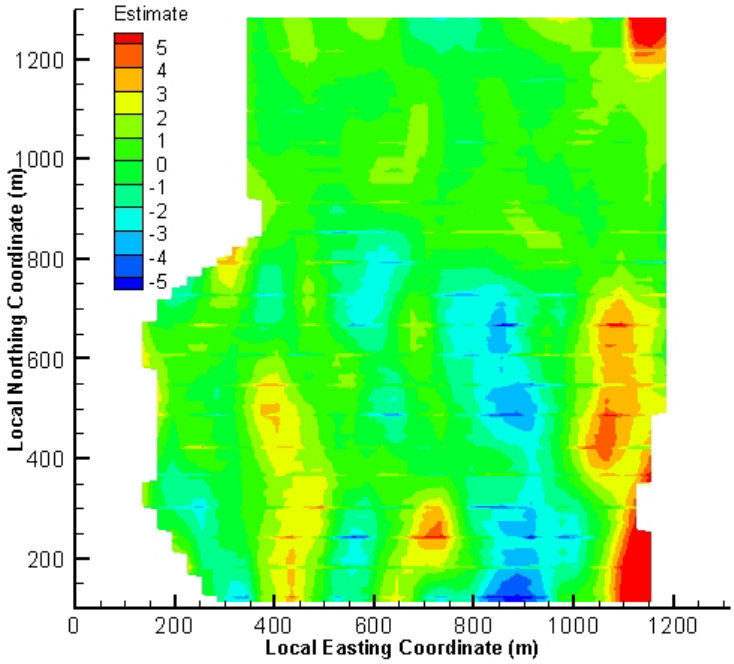


Figure 10. Estimation of residual values between the trend model and the number of anomalies in each 150m^2 cell.

2.2.4 Estimated Anomaly Density

Anomaly density estimation results created by adding the estimated residual field back to the trend model are shown with two different color scales in Figure 11. The image on the left shows the estimated average number of anomalies within each 150m^2 cell. The image on the right shows the same data after scaling them to represent the expected number of anomalies per acre.

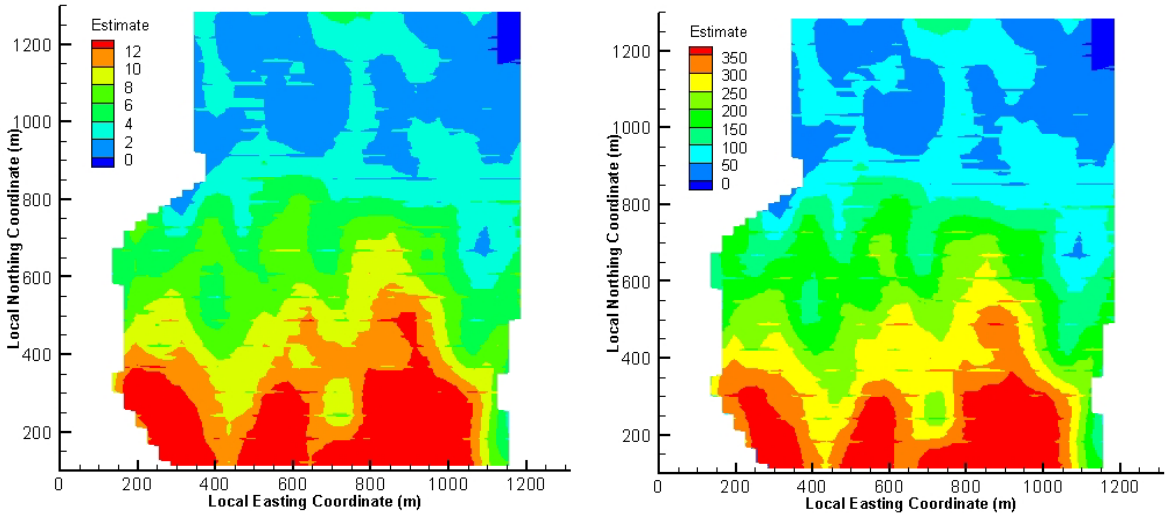


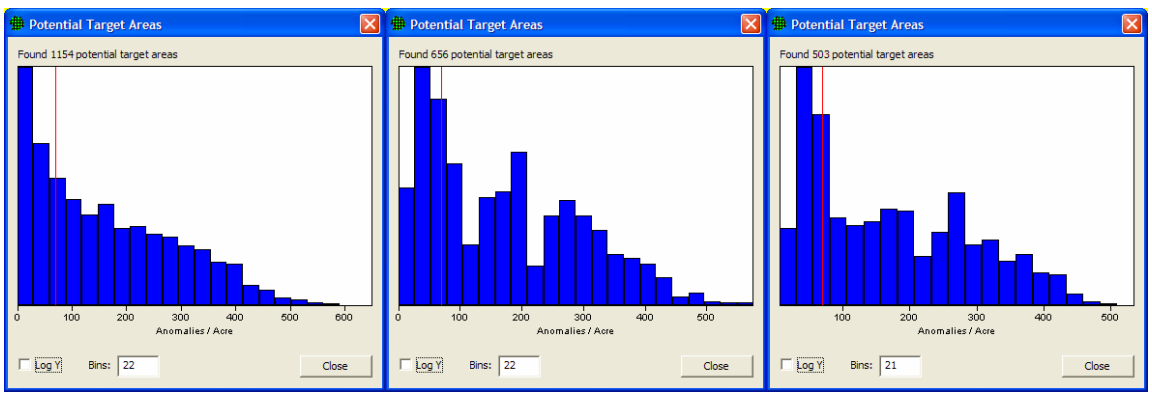
Figure 11 Estimated number of anomalies created by taking into account the trend (CM2) per 150m^2 cell (left) and per acre (right).

2.3 Target Area Boundary Delineation (CM1)

Two approaches to target area boundary delineation are applied to the CM1 data. The first approach examines the distribution of anomaly densities within averaging windows of different sizes and looks for a break in the distribution that signifies the change from background to target area density. The second approach uses the map of anomaly density created through kriging to map the density contour above which densities are considered to be indicative of a target area.

2.3.1 Distribution Analysis

Analysis was conducted using an assumed background density of 70 anomalies per acre. Figure 12 shows the densities across the site for three different window sizes. Windows represent circular areas with centers spaced 1/6 of the window diameter apart along each transect. For each window, VSP computes the density by dividing the number of detected anomalies within the window by the total area traversed by transects.



45.72m (150) ft. diameter 91.44m (300 ft.) diameter 137.16m (450 ft.) diameter

Figure 12. Frequencies of site densities for different window sizes.

Ideally, the histograms in Figure 12 would show high frequencies generally towards the left side of the histogram that represent mostly background areas. As densities increase, we look for abrupt decreases in frequencies that suggest these higher densities are less common. Less common, higher densities are more likely to be potential target areas. Figure 12 shows that a 45.72m (150 ft.) diameter does not distinguish background areas from potential target areas very well. However, the 91.44m (300 ft.) and 137.16m (450 ft.) diameter windows both show many areas with densities under 100 anomalies per acre followed by a drop-off in frequencies as densities increase. In cases where much of the site contains non-background anomalies, it is more difficult to estimate background density, and if available, areas with only background density (no UXO-related material) should be isolated and analyzed to determine a background density estimate.

In Figure 13, VSP "flags" 45.72m (300 ft.) windows where the number of anomalies is significantly greater than a background density of 70 anomalies per acre. This is based on an upper confidence limit for the expected number of traversed and detected anomalies using a Poisson distribution. Since transects are parallel in the data, all windows will contain the same approximate amount of traversed area, 91.44m * 3 m ≈ 274.32m^2. The number of expected anomalies given a background density of 70

anomalies per acre is 274.32m^2 * 70 anomalies/acre * 1 acre / 4,046.86m^2 $\approx$ 4.75 anomalies. For a Poisson distribution with mean 4.75, 9 or more anomalies indicate a window equaling or exceeding the 95th percentile. In Figure 13, windows with 9 or more anomalies have been flagged.

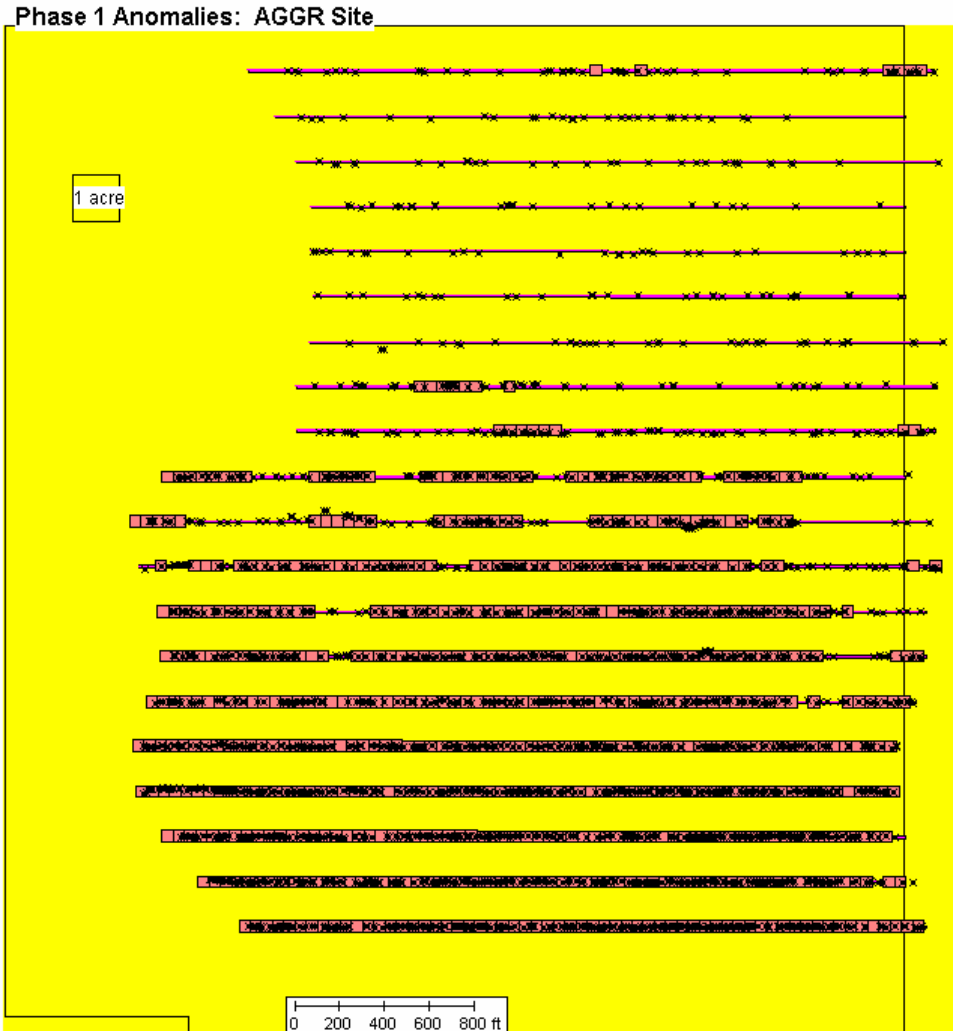


Figure 13. Flagged Areas for 45.72m (300 ft.) diameter window under CM1, 70 anomalies/acre background density, and 95% confidence an area is greater than background density. "Flagged" coordinates along transects are identified with a small box.

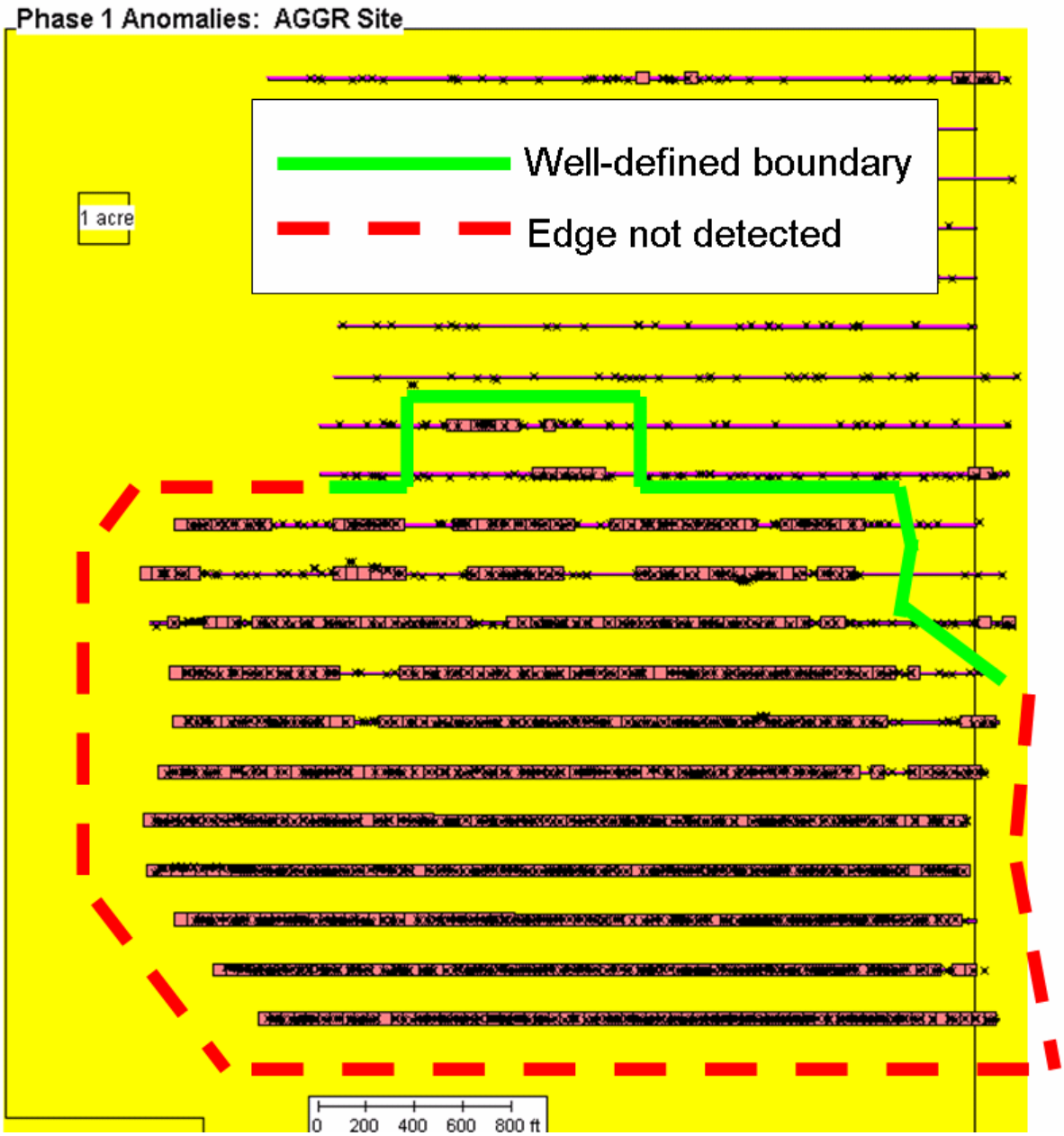


Figure 14. Possible target area boundaries under CM1.

Figure 14 shows the site with suggested boundaries for potential target areas in the Phase 1 area. The solid green line shows a well-defined upper border of an area with high numbers of anomalies compared to an assumed uniform background density. On the left, bottom, and most of the right-hand side of the map, the boundary is unclear since high numbers of anomalies were detected right up to the edge of the transects. This raises the likelihood that the potential target area(s) may extend beyond the creek beds where transects end. The target area also may extend below the lowest transect on the map, and this will be addressed in Phase 2 and Phase 3 of the study.

2.3.2 Density Threshold Contour

Each of the estimated anomaly density values in Figure 5 is the mean of a Poisson distribution defining the expected number of anomalies for that estimation cell. These estimates are compared to the 70 anomaly per acre threshold used in the distribution analysis by determining the 95th percentile cumulative probability of a Poisson distribution with a mean of 70 anomalies per acre. In a similar manner to the distribution analysis above, if the mean of the estimated Poisson distribution is greater than the 95th percentile value of the Poisson distribution with a mean of 70 anomalies per acre, then that cell is defined as being within the target area. Here, we work directly in area units of acres, rather than the area of the traversed transects within the window as done in the distribution analysis. The 95th percentile value of the Poisson distribution with a mean of 70, is 84 anomalies per acre. Thus any estimated anomaly density greater than or equal to 84 is considered part of the target area. This decision process turns the map of estimated anomaly densities into two classes: 1) regions where the estimated anomaly density exceeds the 95th percentile of the background density of 70 anomalies per acre; and 2) regions where the confidence of exceeding the background density is less than 95 percent. These two classes are shown in Figure 15 as the red and grey regions respectively. This classification can be compared qualitatively to the portions of the transects that are flagged in Figure 13. The potential target areas are similar to those identified in the distribution analysis, but slightly larger due to the non-linear scaling of the 95th percentile of the Poisson distribution between units of area traversed and acres.

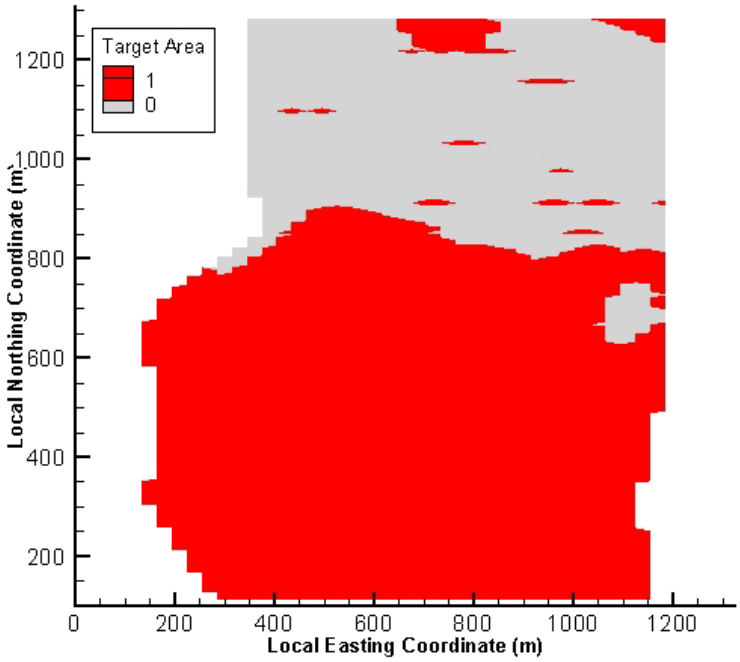


Figure 15. Definition of target areas (red) and background areas (grey) based on the estimated anomalies per acre under CM1 as shown in Figure 5.

2.4 Target Area Boundary Delineation (CM2)

CM2 considers the low-frequency variation in the anomaly density to be caused by some process other than historical use of the site as a munitions range. For target area boundary delineation under the assumption of CM2, the goal is to identify deviations from the spatially varying background anomaly density that would be indicative of a target area. The expression of any target areas above the background density may be very subtle. If the residuals between the background anomaly density as modeled by the cubic function in Figure 6 and the measured anomaly density are truly random, then there will be no spatial clustering, and the regions of the site that are identified as target areas will be spread randomly across the site area.

For the CM2 boundary delineation, it is necessary to define a threshold within residual space to divide background and target areas. The distribution of the residuals is shown in Figure 16, and the proximity of the distribution to a normal distribution, especially at the lower 90 percent of the distribution, is evident. We chose here to define the target areas as those areas where the estimated residual values are the lowest 5 and 10 percent of the residual distribution. This definition requires residual thresholds of -3.5 and -2.5, respectively (Figure 16). Recall that the residual values are defined as the spatially varying background density minus the measured density such that negative residuals indicate areas where the model of spatially varying background density underestimates the measured anomaly density. The thresholds of -3.5 and -2.5 define the locations with the largest underestimation.

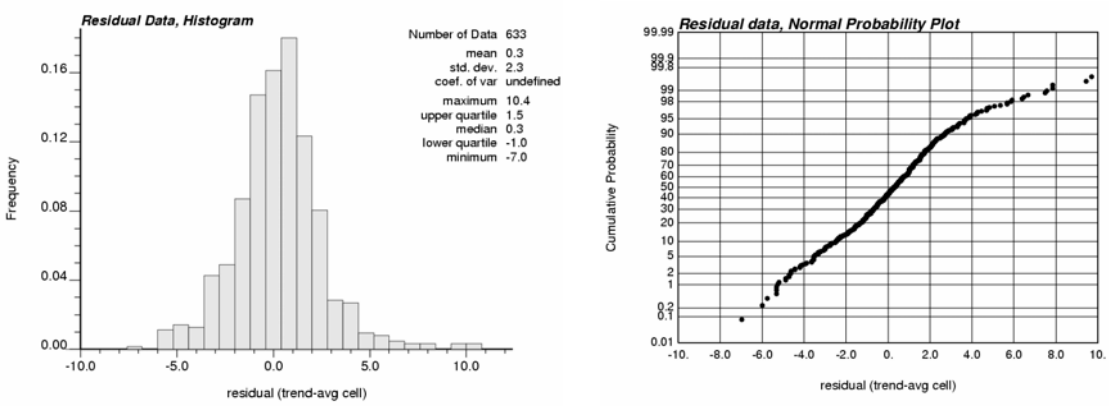


Figure 16. Histogram and normal probability plot of the residual data as determined from all transect data

The two different thresholds are applied to the estimated residual map shown in Figure 10 to produce two different target area maps as shown in Figure 17. The higher the threshold - the larger the amount of target area that is defined. While the defined target areas show some random scatter, there are two larger, spatially contiguous target areas in each map of Figure 17. Specifically, the map highlighting regions where the estimated residual value is in the lowest 10 percent of measured residual values (right image, Figure 17) shows two large connected potential target areas. Each of these areas is roughly 200 m wide by 300-400 meters long. The use of a model of spatially varying background density for detrending the data and the spatial estimation of residuals is able to pull out the subtle expression of what may be target areas from the much larger variation in the background. Additional site knowledge must be applied to determine whether or not CM1 or CM2 is the more plausible conceptual model for this site.

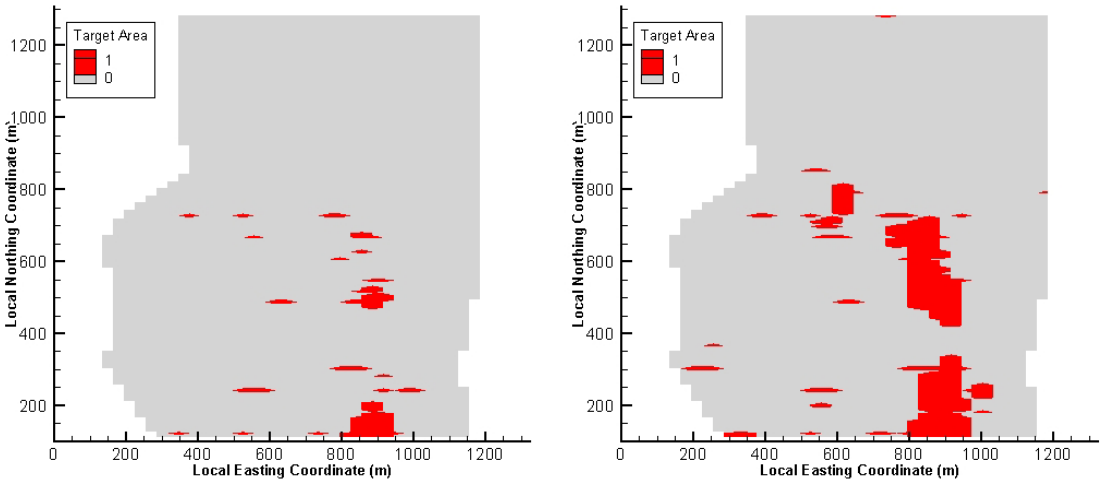


Figure 17. Delineation of target areas using the residuals about the background density trend (CM2). The image on the left is for a threshold of -3.5 and the image on the right is for a threshold of -2.5.

3 Jackknife Validation

In an effort to determine if the same conclusions are reached using transects spaced farther apart, the same data were analyzed using every other transect. The data were split into two groups: (1) Set 1 - every other transect starting with the lowest transect on the map, and (2) Set 2 - every other transect starting with the second lowest transect on the map. Several different analyses are completed using every other transect within the anomaly data set. Results from the two sets are compared to determine if conclusions would have been similar.

3.1 Density Estimation (CM1)

Raw data are created for Set 1 in a similar manner as on the full data set. The only difference is there are fewer transects in this case. The locations and values of the averaged density values are shown in Figure 18. Similar to the full data set, the averaged anomaly density values show significant increase in density from the north to the south across the site. The transects have a variable spacing between 113 and 133 meters (370-430 feet).

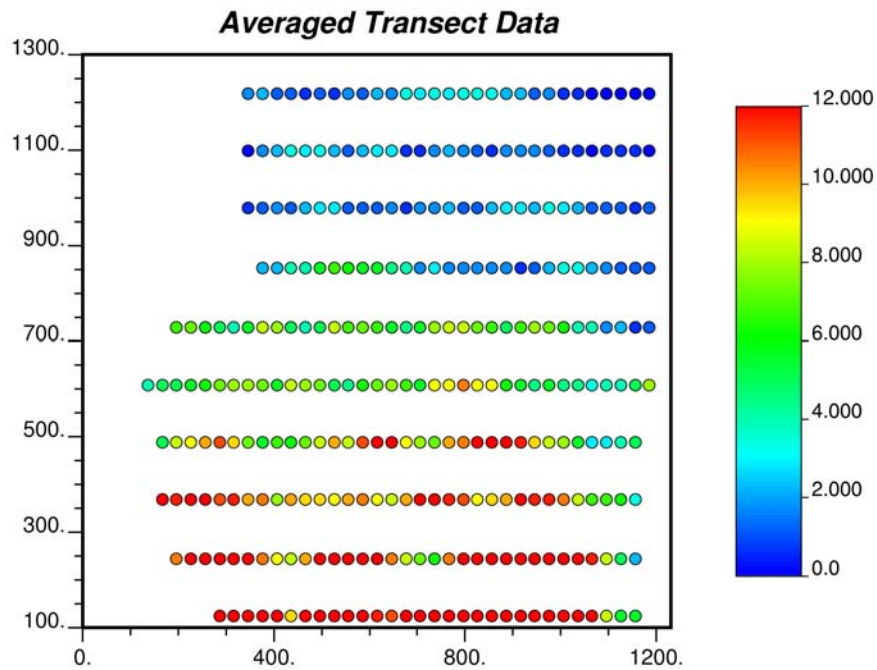


Figure 18. Average anomaly density (per cell) for each 150m^2 estimation cell that is intersected by a sampling transect.

3.1.1 Variogram Analysis

The data shown in Figure 18 are transformed from anomalies/acre to anomalies/estimation cell for analysis of variograms and the following spatial estimations. A single variogram considering spatial correlation in all orientations at once (omnidirectional) is calculated using the cell average anomaly densities and is show in

Figure 19. This variogram is fit with two nested models. A spherical model is used to fit the variogram at small separation distances. This model has a zero nugget, a sill of 4.5 and a range of 120m. At larger separation distances, a Gaussian model is used to fit the experimental variogram. This second model has a sill of 20.5 and a range of 900 meters.

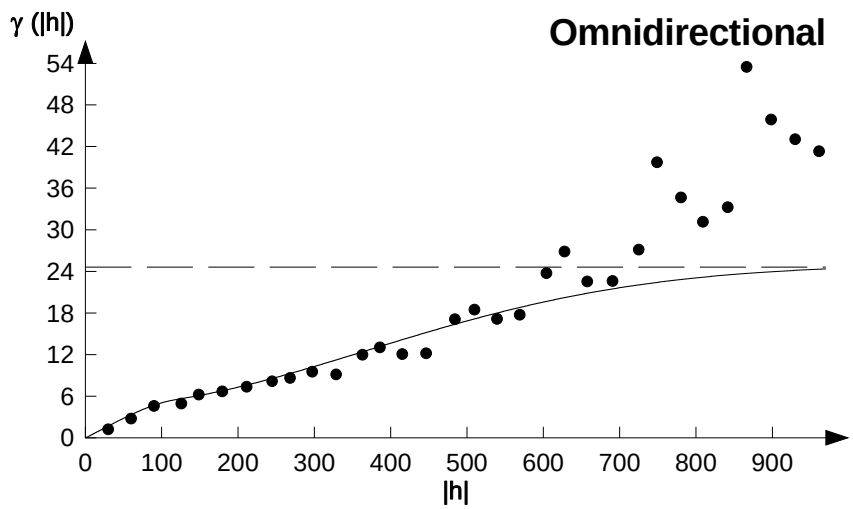


Figure 19. Omnidirectional variogram for the anomaly density values from one-half of the transects at the estimation cell scale for the CM1 model.

When only half of the transects are considered, the behavior of the variogram near the origin is essentially linear, unlike the Gaussian model used with all of the transects. However, a Gaussian model is used to fit the variogram at separation distances beyond 100 meters. The same trend effect as seen in the variogram calculated with all of the data is also present here, although it is less well-defined.

3.1.2 Estimated Anomaly Density

The ordinary kriging algorithm is used to create estimates of the number of anomalies within each of the 30x5 meter cells.

Estimation results are shown with two different color scales in Figure 20. The image on the left shows the estimated average number of anomalies within each 150m^2 cell. The image on the right shows the same data after scaling them to represent the expected number of anomalies per acre.

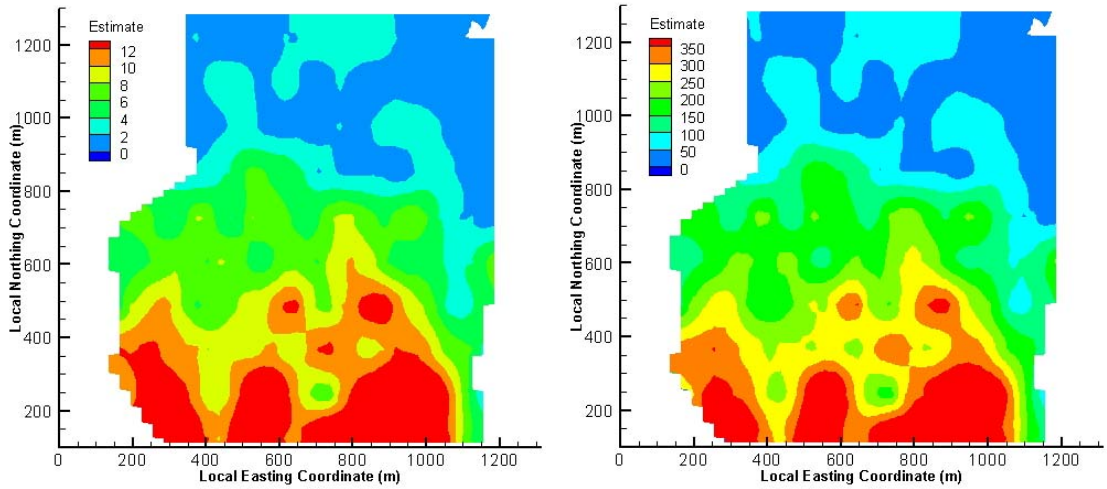


Figure 20. Estimated number of anomalies per 150m^2 cell (left) and per acre (right) using one-half the transects under CM1.

3.2 Density Estimation (CM2)

Similar to the analysis of the full data set, a trend in the background anomaly density is identified and the spatial analysis is completed on the residuals between this trend model and the anomaly densities observed along the transects.

3.2.1 Determination of Background Anomaly Density

Defining the background density above which a region would be classified at the AGGR site is difficult using Set 1 just as it was in the analysis of the full data set due to the anomaly frequency trend running from north to south. The average anomaly density per estimation cell (150m^2) was calculated along each transect and these average values are graphed in Figure 21 as a function of the northing coordinate of the center of each transect.

The same cubic function as used with the full data set is used to define the average background anomaly density of the transects contained in set 1. The parameters Y_0, A, B and C were determined to provide the best fit to the averaged data. These values are Y_0 = 15.97; A = -1.01E-02; B = -1.58E-05; C = 1.21E-08 and provide a fit to the data with an r^2 value of 0.990. The cubic function parameters for this data set have very similar values to those obtained for the full data set.

Similar to the analysis of the full data set, this function provides a model of the average variation of anomaly density across the site. The difference between each of the average values in the cells intersected by a geophysical transect and this trend are the residuals between the measured values and the trend model. These residuals are calculated as residual = trend – measured such that in locations where the trend overestimates the

measured values, the residuals have positive values and vice versa. The values of the residuals as a function of the northing coordinate are shown in Figure 21.

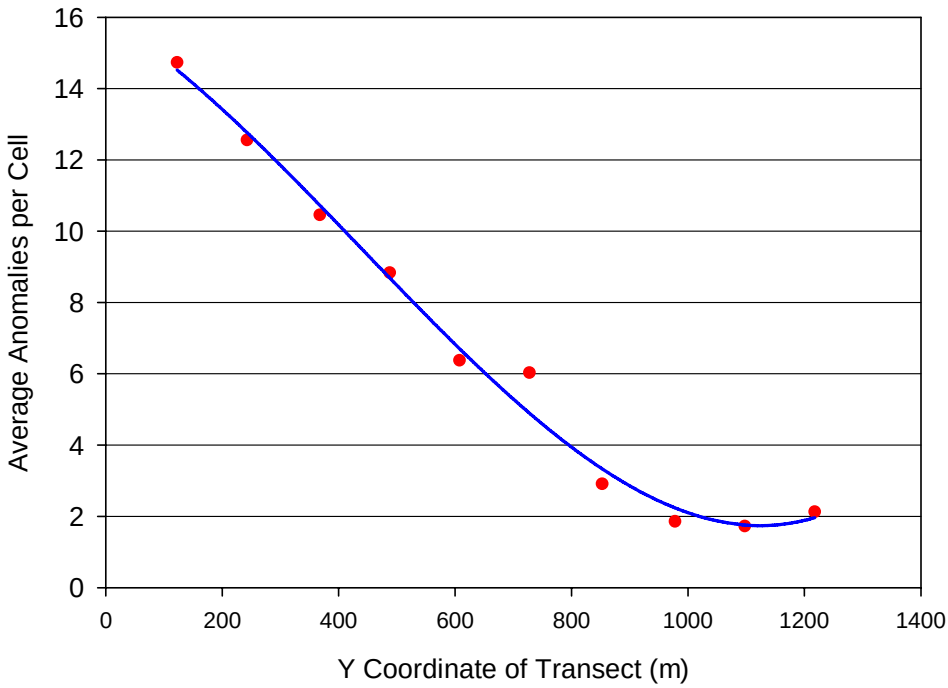


Figure 21. Average values of cell density for each geophysical transect as a function of the Y coordinate (northing) of each transect. The red line is the cubic function fit to the averaged data from half of the data as contained in set 1.

The strong increase in anomaly density to the south of the site is observed and is similar to the trend seen on the full data set. Given the data shown in Figure 21, if the high anomaly densities come from a very large target area, this target area would stretch across the entire length of the geophysical transects, roughly 1000m and the target area would have a diameter greater than 2000m in the north-south direction.

3.2.2 Residual Variogram

A single Gaussian model is used to fit the residual variogram. This model has a strong anisotropy as would be expected after the trend removal. There is essentially infinite correlation in the direction parallel to the trend (north-south) and a well defined spatial correlation in the direction normal to the trend direction (east-west). The variogram model has a nugget value of 0.7 and a sill of 5.25. The range of the variogram in the east-west direction, normal to the trend model, is 130m while in the north-south direction the range is five times larger, 650m. These variograms are shown in Figure 22.

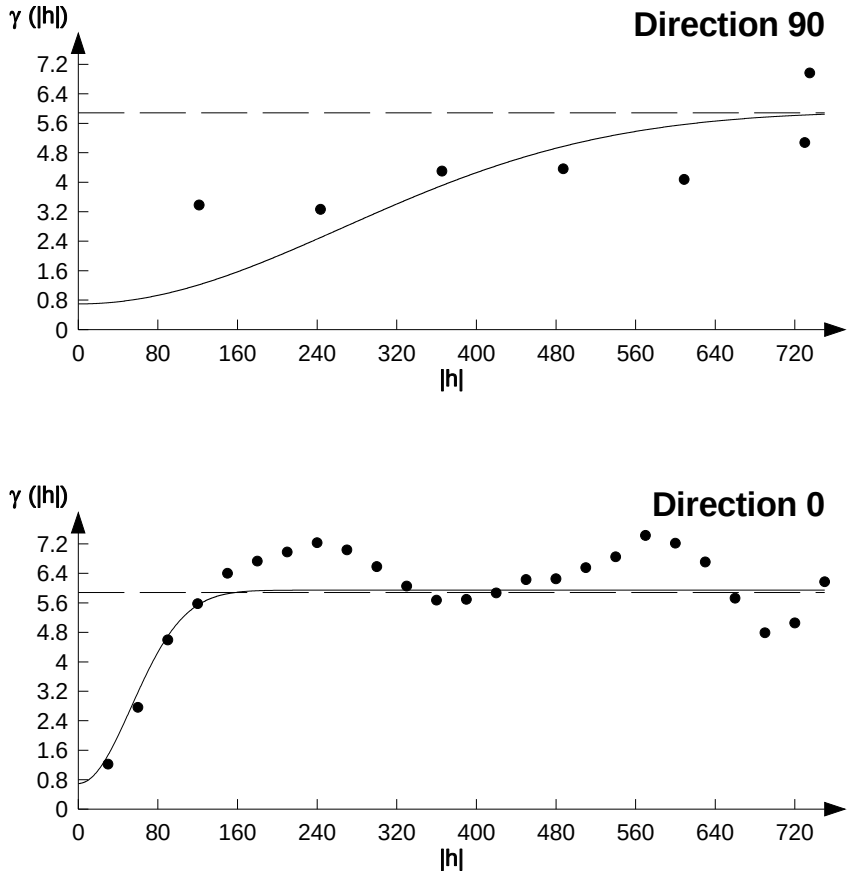


Figure 22. Experimental variograms and models for one-half the detrended anomaly count data (CM2) at the estimation cell scale. The north-south variogram is shown in the top and the east-west variogram in the bottom.

3.2.3 Residual Estimation

Ordinary kriging is used to estimate the residual values at all locations in the domain. These estimated values are shown in Figure 23 and have a similar pattern to the residuals estimated using the full data set. Of particular interest in Figure 23 are the areas of negative residual estimates. Negative residuals indicate regions where the background trend model is underestimating the actual measured anomaly density. These areas may require additional consideration as potential target areas and are examined further in the target area boundary delineation section.

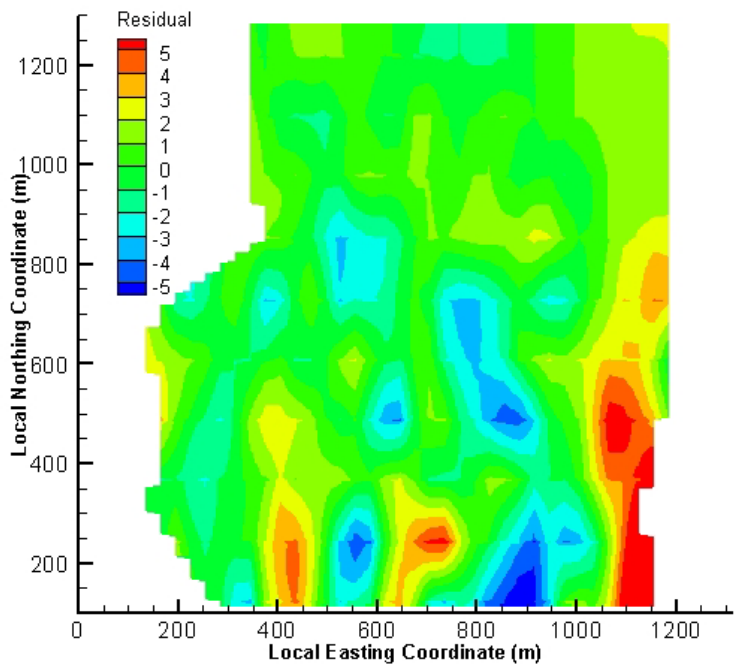


Figure 23. Estimation of residual values calculated using half of the transects as contained in set 1 between the trend model and the number of anomalies in each 150m^2 cell.

3.2.4 Anomaly Density Estimation

Estimation results created by adding the estimated residual field back to the trend model are shown with two different color scales in Figure 24. The image on the left shows the estimated average number of anomalies within each 150m^2 cell. The image on the right shows the same data after scaling them to represent the expected number of anomalies per acre.

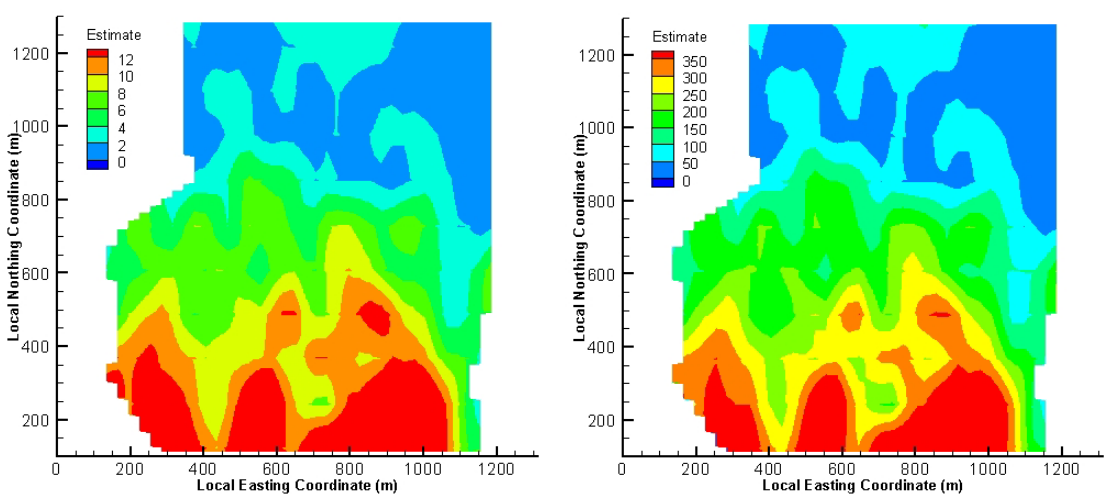


Figure 24. Estimated number of anomalies created by from one-half of the transect data taking into account the trend (CM2) per 150m^2 cell (left) and per acre (right).

3.3 Density Estimation Evaluation (CM1 and CM2)

An advantage of splitting the data set into two halves is that the estimates made with one half of the data can be compared to the measured values in the other half of the data. This jackknife analysis creates anomaly density estimates with the Set 1 transect data and then compares these estimates at the locations of the Set 2 transect data. Estimates of anomaly density per acre as created with both the *CM1* and *CM2* estimation approaches are compared in

Figure 25. The quality of the estimates is defined by the value of the correlation coefficient between the measured and estimated anomaly densities. Both estimation approaches are capable of producing accurate density estimates using only half of the transect data. The correlation coefficient is greater than 0.90 for both the *CM1* and *CM2* approaches.

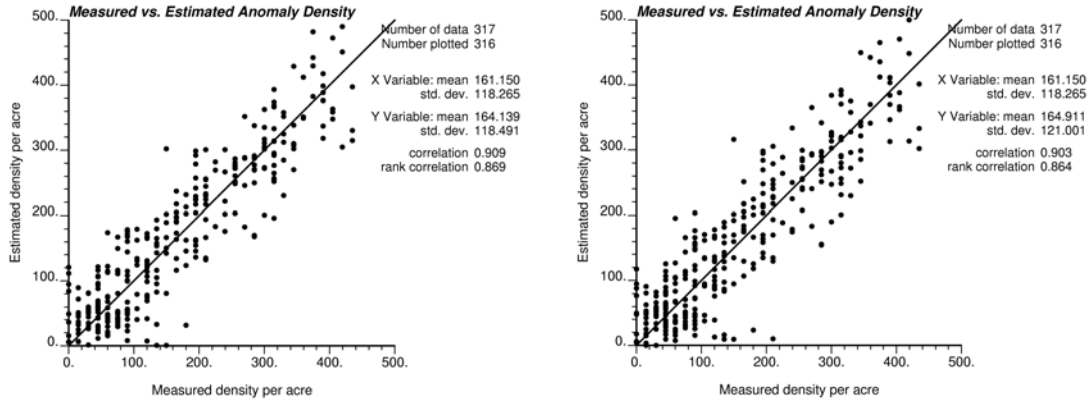


Figure 25. Comparison of measured and estimated anomaly density at the Set 2 measurement locations. Results are shown for the CM1 (left image) and CM2 (right image) estimations.

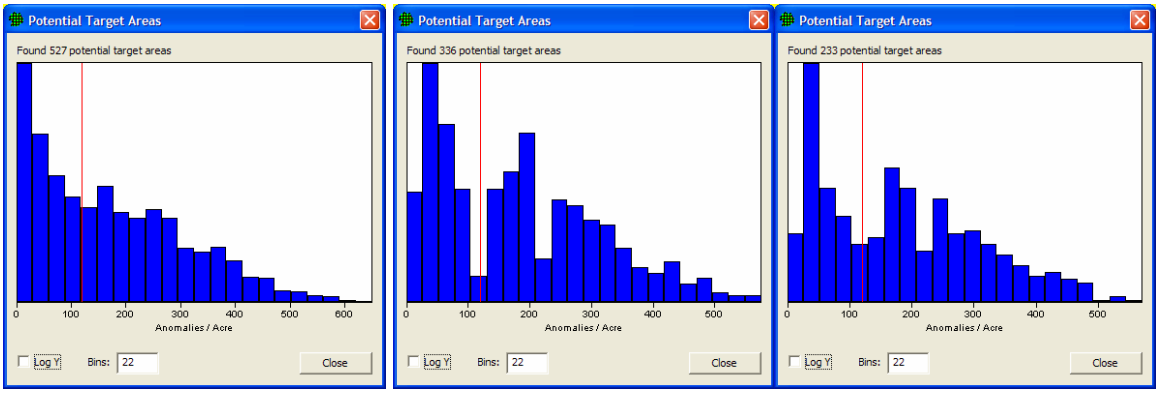
3.4 Target Area Boundary Delineation (CM1)

Similar to the analyses completed with the full data set, two approaches to target area boundary delineation are applied to the CM1 data. The first approach, Distribution Analysis, examines the distribution of anomaly densities within averaging windows of different sizes and looks for a break in the distribution that signifies the change from background to target area density. The Distribution Analysis is completed for both halves of the original transect data, Set 1 and Set 2. The second approach uses the map of anomaly density created through kriging to define the density contour above which densities are considered indicative of a target area. This second approach is applied to only the Set 1 transect data and results are evaluated against the Set 2 transect data later in the report.

3.4.1 Distribution Analysis

Figure 26 shows the densities across the site for three different window sizes. Windows represent circular areas with centers spaced 1/6 of the window diameter apart along each transect. For each window, VSP computes the density by dividing the number of detected anomalies within the window by the total area traversed by transects.

Analysis was conducted using an assumed background density of 60 anomalies per acre using Set 1. Figure 26 shows the densities across the site for three different window sizes.



45.72m (150) ft. diameter 91.44m (300 ft.) diameter 137.16m (450 ft.) diameter

Figure 26. Set 1 – every other transect. Frequencies of site densities for different window sizes.

Ideally, the histograms in Figure 26 would show high frequencies generally towards the left side of the histogram that represent mostly background areas. As densities increase, we look for abrupt decreases in frequencies that suggest these higher densities are less common. Less common, higher densities are more likely to be potential target areas. Figure 26 shows that a 45.72m (150 ft.) diameter does not distinguish background areas from potential target areas very well. However, the 91.44m (300 ft.) and 137.16m (450 ft.) diameter windows both show many areas with densities under 100 anomalies per acre followed by a drop-off in frequencies as densities increase.

In Figure 27, VSP "flags" 91.44m (300 ft.) windows where the number of anomalies is significantly greater than a background density of 60 anomalies per acre. This is based on an upper confidence limit for the number of anomalies using a Poisson distribution as was done on the full dataset.

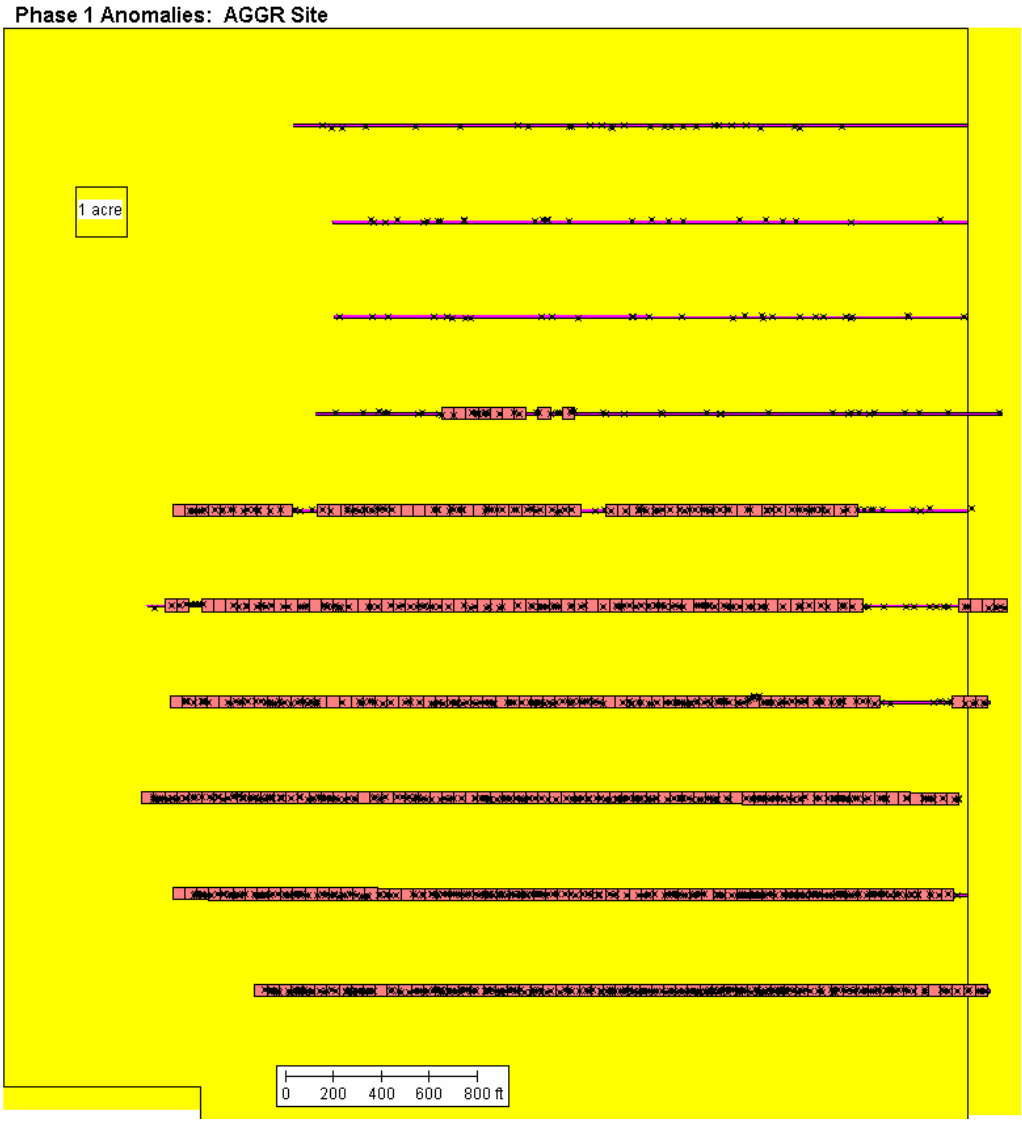
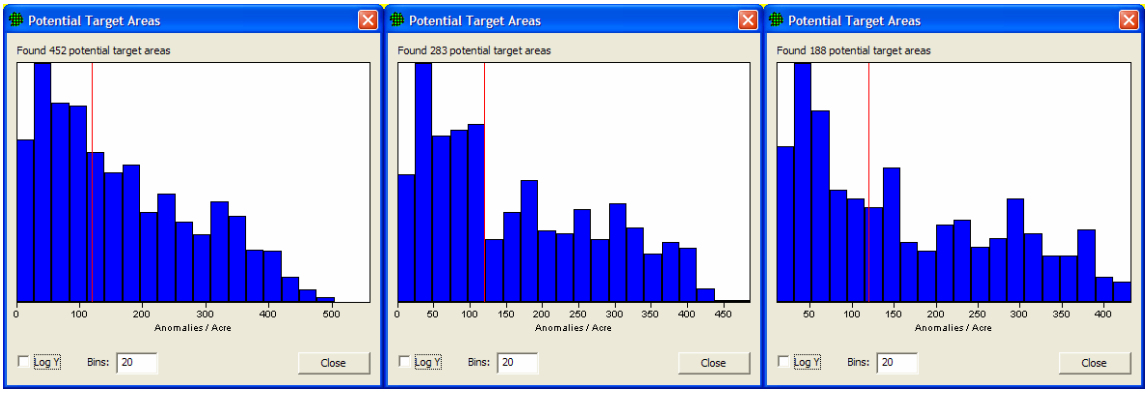


Figure 27. Set 1 – every other transect. Flagged Areas for 91.44m (300 ft.) diameter window, 60 anomalies/acre background density, and 95% confidence an area is greater than background density. "Flagged" coordinates along transects are identified with a small box. Results are for CM1.

For Set 2, analysis was conducted using an assumed background density of 65 anomalies per acre. Figure 28 shows the densities across the site for three different window sizes.



45.72m (150) ft. diameter 91.44m (300 ft.) diameter 137.16m (450 ft.) diameter

Figure 28. Set 2 – every other transect. Frequencies of site densities for different window sizes.

Figure 28 shows that a 45.72m (150 ft.) diameter does not distinguish background areas from potential target areas very well. However, the 91.44m (300 ft.) and 137.16m (450 ft.) diameter windows both show many areas with densities under 100 anomalies per acre followed by a drop-off in frequencies as densities increase.

In **Error! Reference source not found.**, VSP "flags" 91.44m (300 ft.) windows where the number of anomalies is significantly greater than a background density of 65 anomalies per acre. This is based on an upper confidence limit for the number of anomalies using a Poisson distribution as was done on the full dataset.

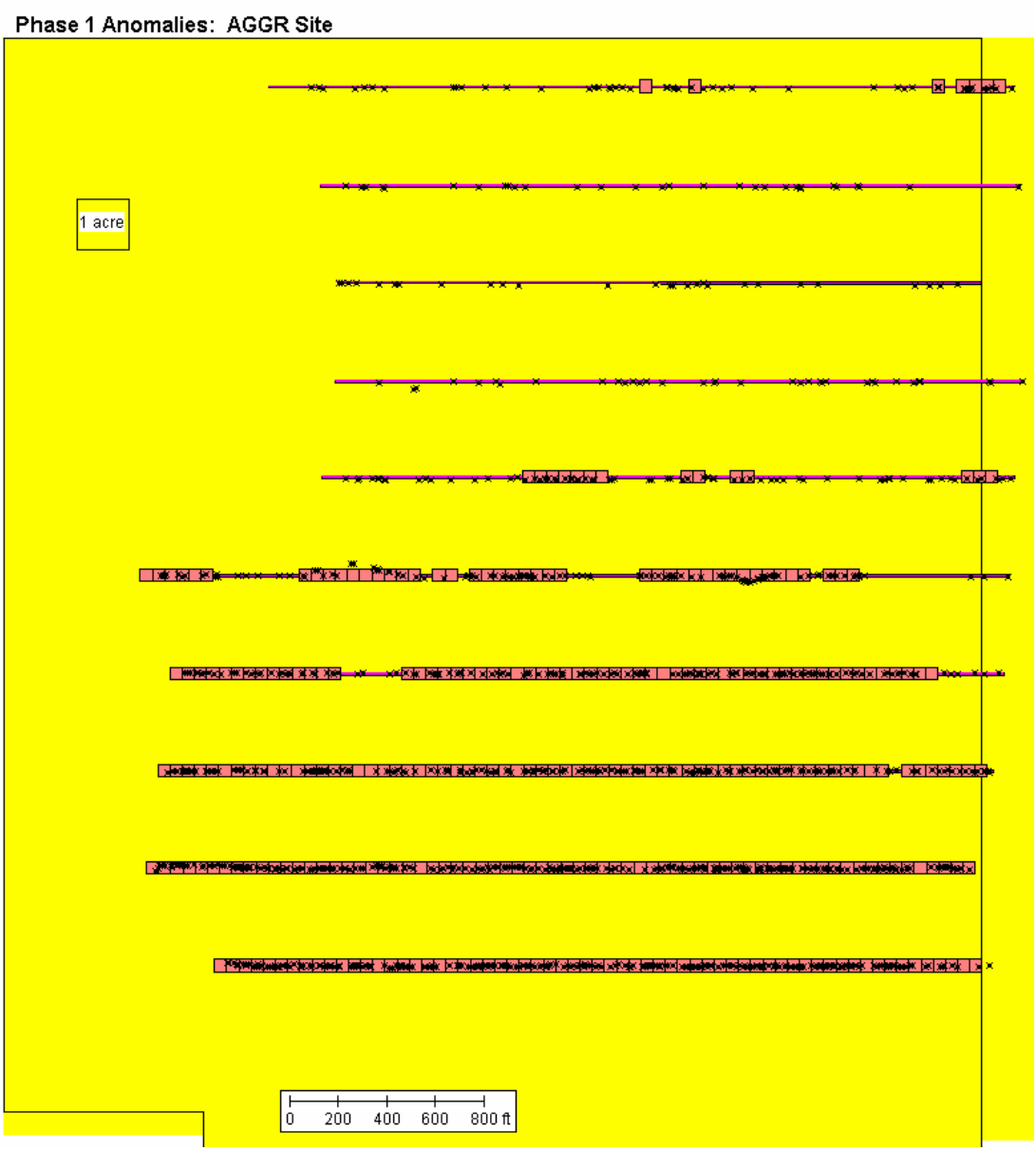


Figure 29. Set 2 – every other transect. Flagged Areas for 91.44m (300 ft.) diameter window, 65 anomalies/acre background density, and 95% confidence an area is greater than background density. "Flagged" coordinates along transects are identified with a small box. Results are for CM1

Figure 30 shows the site with suggested boundaries for potential target areas in the Phase 1 area based on the analysis of Set 1. The solid green line shows a well-defined upper border of an area with high numbers of anomalies compared to an assumed uniform background density. On the left, bottom, and most of the right-hand side of the map, the boundary is unclear since high numbers of anomalies were detected right up to the edge of the transects. This raises the possibility that the potential target area(s) may extend beyond the creek beds where transects end.

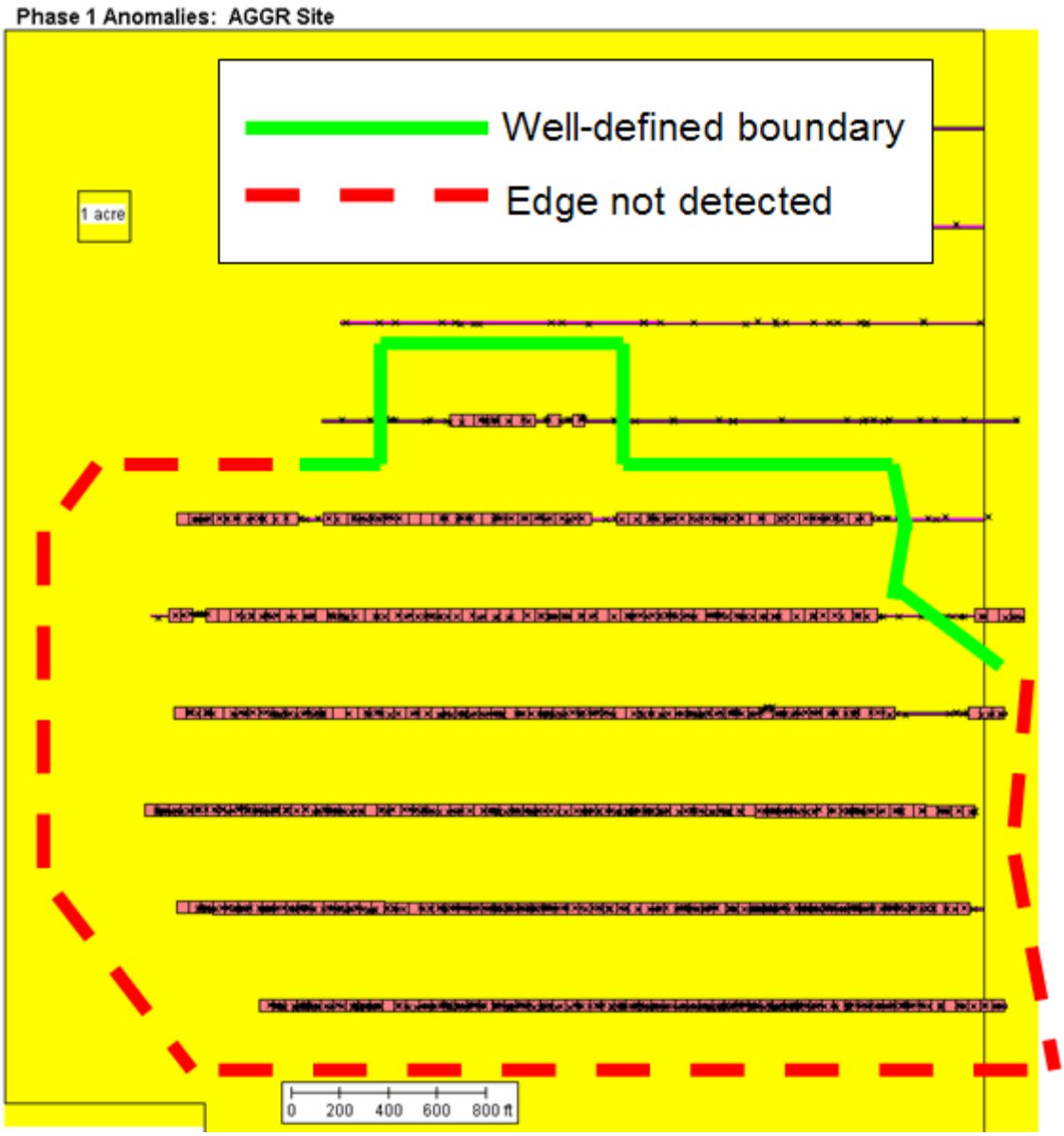


Figure 30. Set 1 – every other transect. Possible target area boundaries under CM1.

Figure 31 shows the site with suggested boundaries for potential target areas in the Phase 1 area based on the analysis of Set 2. The solid green line shows a well-defined upper border of an area with high numbers of anomalies compared to an assumed uniform background density. On the left, bottom, and most of the right-hand side of the map, the boundary is unclear since high numbers of anomalies were detected right up to the edge of the transects. This raises the possibility that the potential target area(s) may extend beyond the creek beds where transects end.

Figure 31. Set 2 – every other transect. Possible target area boundaries under CM1.

Target area boundaries are similar for the analysis of the full data set, Set 1, and Set 2. Much of this is due to the fact that edges were not detected for much of the detected targets areas. Well-defined boundaries were similar for the full data set, Set 1, and Set 2. Set 1 and Set 2 had more conservative boundaries due to wider transect spacings.

3.4.2 Density Threshold Contour

Each of the estimated anomaly density values in Figure 24 is the mean of a Poisson distribution defining the expected number of anomalies for that estimation cell. These estimates are compared to the 60 anomaly per acre threshold used in the distribution analysis by determining the 95th percentile cumulative probability of a Poisson

distribution with a mean of 60 anomalies per acre. If the mean of the estimated Poisson distribution is greater than the 95th percentile value of the Poisson distribution with a mean of 60 anomalies per acre, then that cell is defined as being within the target area. The 95th percentile value of the Poisson distribution with a mean of 60, is 73 anomalies per acre. Thus any estimated anomaly density greater than or equal to 73 is considered part of the target area. This decision process turns the map of estimated anomaly densities into two classes: 1) regions where the estimated anomaly has a 95 percent confidence of being greater than the background density of 60 anomalies per acre; and 2) regions where the confidence of exceeding the background density is less than 95 percent. These two classes are shown in Figure 32 as the red and grey regions respectively. This classification can be compared qualitatively to the portions of the transects that are flagged in Figure 27 and Figure 29 above.

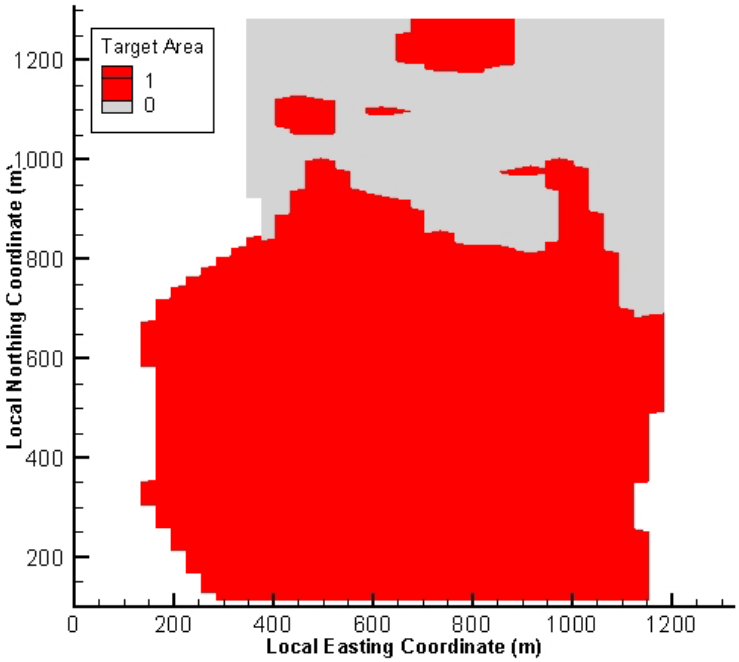


Figure 32. Definition of target areas (red) and background areas (grey) based on the estimated anomalies per acre under CM1 as shown in Figure 20. Estimated number of anomalies per 150m^2 cell (left) and per acre (right) using one-half the transects under CM1.

3.5 Target Area Boundary Delineation (CM2)

For target area boundary delineation under the assumption of CM2, the goal is to identify deviations from the spatially varying background anomaly density that would be indicative of a target area. For the CM2 boundary delineation, it is necessary to define a threshold within residual space to divide background and target areas. The distribution of the residuals from transect data Set1 is shown in Figure 33, and the proximity of the distribution to a normal distribution, especially at the lower 90 percent of the distribution, is evident. We use the same definitions of target area here as was used for the full data set: the lowest 5 and 10 percent of the residual distribution. The lower end of the Set 1 residual distribution is very similar to that of the full data set residual distribution and the thresholds corresponding to the lowest 5 and 10 percent of the distribution are again -3.5 and -2.5, respectively (Figure 33). Recall that the residual values are defined as the spatially varying background density minus the measured density such that negative residuals indicate areas where the model of spatially varying background density underestimates the measured anomaly density. The thresholds of -3.5 and -2.5 define the locations with the largest underestimation of the measured data by the background density model.

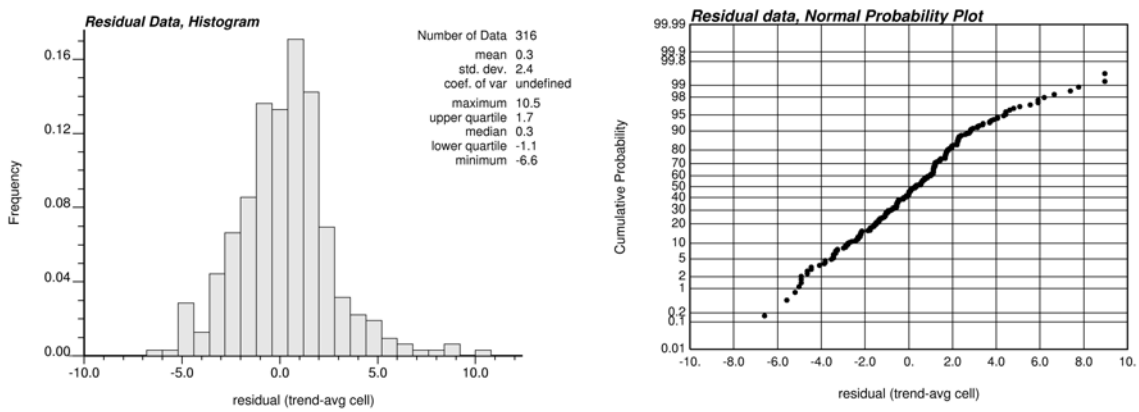


Figure 33. Histogram and normal probability plot of the residual data as determined from the Set 1 half of the transect data

The two different thresholds are applied to the estimated residual map shown in Figure 23 to produce two different target area maps as shown in

Figure 34. While the defined target areas show a few random cells, there are several larger, spatially contiguous target areas in each map of

Figure 34. Compared to the analysis using the full data set, there are more contiguous regions identified as potential target areas and, in general, these potential target areas are smoother than those identified with the full transect data set. The potential target areas are 100-200 meters wide and 100-400 meters long. Even with only half of all transects, the use of a model of spatially varying background density for detrending the data and the spatial estimation of residuals is again able to pull out the subtle expression of what may be target areas from the much larger variation in the background. Additional site

knowledge must be applied to determine whether or not *CM1* or *CM2* is the more plausible conceptual model for this site.

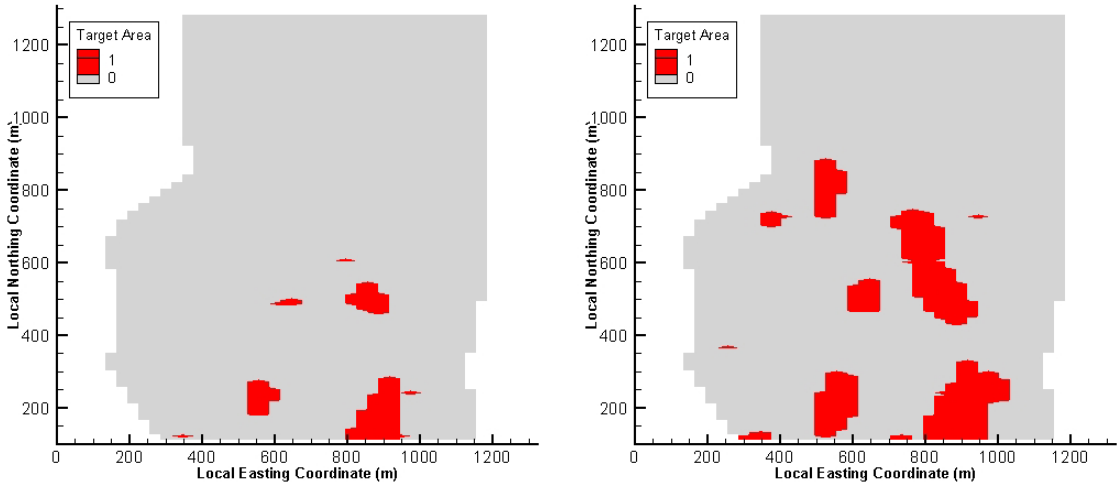


Figure 34. Delineation of target areas using the residuals about the background density trend (CM2). The image on the left is for a threshold of -3.5 and the image on the right is for a threshold of -2.5.

4 Conclusions

Results of these analyses show that a large high-density area is present in the middle to lower half of the Phase 1 area of the AGGR, while the upper section of the site has a generally lower anomaly density. Anomaly densities tend to follow a smoothly increasing trend from north to south and the peak density occurs near the southern edge of the Phase 1 area. This behavior shows possible evidence that the background density may not be uniform across the site. Two different conceptual models were considered to define the background density: *CM1* considers the background density to be uniform across the site. If we assume a uniform background density across the site, background density is approximately 60-70 anomalies per acre or lower across the site. *CM2* considers the background anomaly density to be a function of the northing coordinate. This non-uniform background anomaly density trend could be due to a large target area or several close knit target areas in the southern half of the Phase 1 area, or due to natural variation in the conductivity of the soils. For both conceptual models, the anomaly density is estimated across the site area. A jackknife analysis using half of the transect data showed that the anomaly density estimates were accurate under both the *CM1* and *CM2* assumptions. For *CM1*, two approaches to delineating the target boundary were used: distribution analysis and defining a density threshold contour. Target delineation with the density contour approach was also used for the *CM2* density estimates.

The portion of the site area determined to be part of the target area is compared across the different approaches used to estimate the anomaly density in Table 1. Results in Table 1 reiterate that the largest differences in the boundary delineations are due to the choice of the conceptual model. If the large north to south increase in anomaly density is considered to be due to the transects covering the northern flank of a very large target (*CM1*), then the majority of the site (65-70 percent) is identified as a potential target area.

If the large increase in anomaly density is considered to be part of the background variation (*CM2*), then the identified potential target areas are a much smaller fraction of the total site (1 to 7 percent depending on the data set and chosen threshold). Additional site information should be able to quickly determine which conceptual model is most appropriate.

Analytical results are consistent regardless of whether all transects are used, or every other transect. The same high-density areas were identified using the full data set and data sets where only every other transect was used. For either conceptual model, use of only half of the measured transects in the analysis resulted in only slight increases (a few percent) in the amount of the site identified as a potential target area relative to using all of the transect data.

Table 1. Results of target area boundary delineation.

	All Transect Data		
	Number of Cells in Target Area	Area within Target Area (m2)	Percent of Site within Target Area
CM1, 70 anomalies/acre	4866	7.30E+05	65.1
CM2, lowest 5% residuals	95	1.43E+04	1.3
CM2, lowest 10% residuals	401	6.02E+04	5.4
	Set 1 Transect Data		
	Number of Cells in Target Area	Area within Target Area (m2)	Percent of Site within Target Area
CM1, 60 anomalies/acre	5154	7.73E+05	69.0
CM2, lowest 5% residuals	151	2.27E+04	2.0
CM2, lowest 10% residuals	541	8.11E+04	7.2

The current nominal transect spacing of 60.96m (200 ft.) between transect centers is adequate if high-density areas are expected to be as large and dense as those found in Phase 1. A 121.92m (400 ft.) spacing also appears to be adequate for detecting these kinds of large, dense areas. The mapping of the edge of the target area was similar for transect spacings of 60.96m and 121.92m (200 and 400 ft.) other than variation along the edges due to the wider transect spacings when using every other transect.

In conclusion, this report has examined anomaly densities of Phase 1 sampling at the AGGR site. The analysis identified a large high density area in the middle to lower sections of the Phase 1 area regardless of whether a 60.96m or 121.92m (200 ft. or 400 ft.) transect spacing was used. Similar boundaries can be drawn around this high-density area for both the 60.96m and 121.92m (200 ft. and 400 ft.) transect spacing.