

**Final Report**

**Intelligent Data Fusion for Wide-Area Assessment  
of UXO Contamination**

SERDP Project MM-1510

Naval Research Laboratory

31 January 2008

Susan Rose-Pehrsson and Kevin Johnson  
Naval Research Laboratory  
Chemistry Division, Code 6181  
4555 Overlook Ave. SW  
Washington, DC 20375

Christian Minor  
Nova Research, Inc.  
1900 Elkin St, Suite 230  
Alexandria, VA 22308

Verner Guthrie  
U.S. Army Engineering Research and Development Center  
7701 Telegraph Road  
Alexandria, VA 22315-3864

**Approved for public release; distribution is unlimited**

This report was prepared under contract to the Department of Defense Strategic Environmental Research and Development Program (SERDP). The publication of this report does not indicate endorsement by the Department of Defense, nor should the contents be construed as reflecting the official policy or position of the Department of Defense. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the Department of Defense.

08-1226-0376

NRLINST 5600.2

PUBLICATION OR PRESENTATION RELEASE REQUEST

1. REFERENCES AND ENCLOSURES	2. TYPE OF PUBLICATION OR PRESENTATION	3. ADMINISTRATIVE INFORMATION
Ref: (a) NRL Instruction 5600.2 (b) NRL Instruction 5510.40D  Encl: (1) Two copies of subject paper <del>XXXXXXXXXX</del>	<input type="checkbox"/> Abstract only, published <input type="checkbox"/> Book <input type="checkbox"/> Conference Proceedings (refereed) <input type="checkbox"/> Invited speaker <input type="checkbox"/> Journal article (refereed) <input type="checkbox"/> Oral Presentation, published <input checked="" type="checkbox"/> Other explain: <u>SERDP Final Report</u> (Websites - see comments below)	<input type="checkbox"/> Abstract only, not published <input type="checkbox"/> Book Chapter <input type="checkbox"/> Conference Proceedings (not refereed) <input checked="" type="checkbox"/> Multimedia report <input type="checkbox"/> Journal article (not refereed) <input type="checkbox"/> Oral Presentation, not published
		STRN <u>NRL/O/6180/08/0012</u> Route Sheet No. <u>61/0012</u> Job Order No. <u>61-9010-0-7-5</u> Classification <u>X</u> U <u>    </u> C <u>    </u> Sponsor <u>SERDP-see attached approval</u> * approval obtained <u>X</u> yes <u>    </u> no

4. AUTHOR

Title of Paper or Presentation  
Intelligent Data Fusion for Wide-Area Assessment of UXO Contamination - SERDP Project MM1510 - Final Report

Author(s) Name(s) (First, Mi, Last), Code, Affiliation if not NRL  
Susan Rose-Pehrsson, Kevin Johnson, Christian Minor (Nova Research, Inc.) and Verner Guthrie (U.S. Army Engineering Research and Development Center)

It is intended to offer this paper to the \_\_\_\_\_  
(Name of Conference)

and/or for publication in \_\_\_\_\_  
(Name and Classification of Publication)

After presentation or publication, pertinent publication/presentation data will be entered in the publications data base, in accordance with reference (a).

It is the opinion of the author that the subject paper (is \_\_\_\_\_) (is not X) classified, in accordance with reference (b). This paper does not violate any disclosure of trade secrets or suggestions of outside individuals or concerns which have been communicated to the Laboratory in confidence. This paper (does \_\_\_\_\_) (does not X) contain any militarily critical technology. This subject paper (has \_\_\_\_\_) (has never X) been incorporated in an official NRL Report.

SUSAN ROSE-PEHRSSON Name and Code (Principal Author)  
Susan L. Rose Pehrsson (Signature)

5. ROUTING/APPROVAL

CODE	SIGNATURE	DATE	COMMENTS
Author(s)	<u>Susan L. Rose Pehrsson</u>	<u>1-24-08</u>	* Websites paper will be posted on: SERDP on-line library - <a href="http://docs.serdp-estcp.org">http://docs.serdp-estcp.org</a> Defense Technical Information Center (DTIC) <a href="http://www.dtic.mil">http://www.dtic.mil</a>
Section Head	<u>Susan L. Rose Pehrsson</u>	<u>1-24-08</u>	
Branch Head Code 6180	<u>Rose Johnson</u>	<u>1-29-08</u>	
Division Head Code 6100	<u>M. Koderski</u>	<u>1/30/08</u>	1. Release of this paper is approved. 2. To the best knowledge of this Division, the subject matter of this paper (has _____) (has never <u>X</u> ) been classified.
Security, Code 1226 1226	<u>[Signature]</u>	<u>2/18/08</u>	1. Paper or abstract was released. 2. A copy is filed in this office.
Office of Counsel, Code 1008.2	<u>[Signature]</u>	<u>2/12/08</u>	Sponsor's approval attach
ADOR/Director NCST Code 6000	<u>[Signature]</u>	<u>14 FEB 2008</u>	This is a Final Security Review.
Public Affairs (Unclassified/ Unlimited Only), Code 1030	<u>RLJ</u>	<u>2/15/08</u>	Any changes made in the document after approved by Code 1226 nullify the Security Review.
Division, Code 6180			
Author, Code 6180A	cy to: SRP, KJ, CM	on: <u>2/20/08</u>	

FEB 5 PM 10:53

# Table of Contents

List of Acronyms .....	iv
List of Figures .....	v
Acknowledgement .....	viii
Executive Summary .....	1
Objective .....	6
Background .....	6
Problem Statement .....	6
Current Technology .....	7
Data Fusion Approach .....	8
Methods .....	10
Automated Feature Extraction .....	11
Feature Layer Development .....	12
Data Fusion Architecture Development .....	13
Results and Discussion .....	21
Data Registration and Importation .....	21
<i>Ground truth data.</i> .....	27
Feature Extraction and Feature Layer Generation .....	28
<i>Manually delineated features.</i> .....	28
<i>Aerial orthophotography features.</i> .....	33
<i>LiDAR data features.</i> .....	34
<i>Helicopter magnetometry feature extraction.</i> .....	42
<i>Enhancement of magnetic anomalies.</i> .....	47
<i>Summary of feature layers.</i> .....	53
Data Fusion Results .....	54
<i>Heuristic approach.</i> .....	54
<i>Bayesian approach.</i> .....	55
<i>Dempster-Shafer approach.</i> .....	60
<i>Hybrid heuristic/Dempster-Shafer approach.</i> .....	72
Summary of Results .....	77
Conclusions .....	79
References .....	81
Appendix A: Supporting Data .....	83
Appendix B: List of Technical Publications .....	83

## **List of Acronyms**

AMTADS	Airborne Multi-Sensor Towed Array Detection System
EMI	Electromagnetic Induction
ESRI	Environmental Systems Research Institute
ESTCP	Environmental Security Technology Certification Program
GIS	Geographic Information System
GPR	Ground Penetrating Radar
GPS	Global Positioning System
KDE	Kernel Density Estimation
LiDAR	Light Detection And Ranging
NRL	Naval Research Laboratory
nT	nanoTesla
PBR	Precision Bombing Range
ROC	Receiver Operating Characteristic Curve
SAR	Synthetic Aperture Radar
SERDP	Strategic Environmental Research and Development Program
TFM	Total Field Magnetometry
UTM	Universal Transverse Mercator Grid Coordinates
UXO	Unexploded Ordnance
WAA	Wide-Area Assessment
WAAPP	Wide-Area Assessment Pilot Program

## List of Figures

Figure 1. Hypothetical probability mass assignments in Dempster-Shafer data fusion. Example (A) depicts the assignment of complex functions to each of the three focal elements over the observed range of feature intensity values. Example (B) depicts an assignment where probability masses are assigned to the feature intensity extremes for each of the three focal elements and intermediate values are estimated through linear interpolation.....	19
Figure 2. Dempster’s rule of combination implemented for a UXO assessment application.....	19
Figure 3. Helicopter magnetometry survey from the Pueblo ESTCP WAAPP survey site .....	22
Figure 4. Helicopter magnetometry survey from the Kirtland ESTCP WAAPP survey site.....	23
Figure 5. LiDAR survey from the Pueblo ESTCP WAAPP survey site. ....	24
Figure 6. LiDAR survey from the Kirtland ESTCP WAAPP survey site. ....	25
Figure 8. Aerial orthophotographic survey from the Kirtland ESTCP WAAPP survey site.....	27
Figure 9. Pueblo PBR #2 topographical map with ground truth data. Red dots indicate locations of ordnance-related, blue dots indicate non-UXO scrap, green dots indicate geologic feature (i.e., an empty dig site), and magenta dots indicate locations of intact ordnance.....	28
Figure 10. Feature layers resulting from (A), manually identified ship targets dilated to a larger effective area, and (B), manually identified munitions areas.....	30
Figure 11. Feature layers resulting from manually identified roads, fences, pipelines, and man-made structures within the Kirtland site.....	31
Figure 12. Enlarged region of the Kirtland feature layer containing bombing targets (A) before, and (B), after dilation.....	32
Figure 13. Feature layers resulting from (A), manually identified ship targets, and (B), manually identified munitions areas, and (C), manually identified man-made structures in the Pueblo site. ....	33
Figure 14. Orthophoto of a sub region of the Pueblo PBR #2.....	34
Figure 15. Subsection of LiDAR generated surface estimate for the Pueblo site showing ordnance-related cratering. Typical crater depths are smaller than variations in surface elevation. ....	35
Figure 16. Estimated LiDAR surface gradient (A) and direction (B) before and after (C) gradient magnitude threshold application. (D) A direction of 0 radians indicates a positive surface gradient estimate orientation of due east. ....	36
Figure 17. Surface gradient direction template for a crater of four meter radius in a one meter resolution image.....	37
Figure 18. Residual between data and ideal crater template for (A) a non-crater containing region, and (B) a crater containing region. ....	37
Figure 19. Crater detection at the Pueblo site. (A) depicts automatically identified craters via an NRL crater-detection algorithm. (B) depicts a crater density estimate determined via Kernel Density Estimation. Colors are scaled from blue (=0) to red (=1).....	38

Figure 20. Craters detection at the Kirtland site. (A) depicts manually identified craters (Versar, Inc.) (B) depicts automatically identified craters via an NRL crater-detection algorithm. .... 40

Figure 21. Crater density estimate determined via Kernel Denisty Estimation. (A) depicts manually identified craters (Versar, Inc.). (B) depicts automatically identified craters via an NRL crater-detection algorithm. Colors are scaled from blue (=0) to red (=1). .... 41

Figure 22. Shown in (A) are pixel islands in the magnetometry data. (B) is a map of estimated feature density generated from (A)..... 43

Figure 23. Filtering of pixel islands: (A) all pixel islands, (B) islands with 2 – 75, (C) 75 – 200, and (D) 200 – 1000 member pixels..... 44

Figure 24. Average pixel island intensities, capped at 55 nT, for island pixels surrounding the target in the southern portion of the Pueblo site. Surface clutter objects such as fence lines have significantly greater intensity per pixel..... 45

Figure 25. Shown in (A) are magnetometer signal pixel islands with average intensity less than 50 nT. Shown in (B) is a map of estimated feature density generated from the pixel islands in (A)..... 46

Figure 26. Shown in (A) are magnetometer signal pixel islands with average intensity more than 50 nT. Shown in (B) is a map of estimated feature density generated from the pixel islands in (A)..... 47

Figure 27. Magnetic anomalies at the Kirtland site. (A) depicts algorithmically identified anomaly locations (Sky Research). (B) depicts anomaly locations after filtering with an NRL clustering algorithm. .... 49

Figure 28. Magnetic anomalies at the Pueblo site. (A) depicts algorithmically identified anomaly locations (Sky Research). (B) depicts anomaly locations after filtering with an NRL clustering algorithm. .... 50

Figure 29. Histogram of nearest neighbor distances between magnetic anomalies at the Kirtland site (yellow bars), overlay of Poisson distribution of mean 6.0 (= variance) (blue line)..... 51

Figure 30. Magnetic anomalies at the Kirtland site. (A) depicts estimated anomaly density for Fig. 29(A). (B) depicts estimated anomaly density for Fig. 29(B)..... 52

Figure 31. Magnetic anomalies at the Pueblo site. (A) depicts estimated anomaly density for Fig. 30(A). (B) depicts estimated anomaly density for Fig. 30(B). .... 53

Figure 32. Heuristic-based data fusion output utilizing magnetic anomaly density, manually detected crater density, bombing target, and munitions area feature layers. (Map is color scaled from blue=0 to red=1). .... 54

Figure 33. Initial conditional probability assignments for Bayesian data fusion combining helimag anomaly density, crater density, known bombing targets, and manually delineated munitions areas. .... 55

Figure 34. Bayesian data fusion output for the Pueblo site utilizing magnetic anomaly density, automatically detected crater density, bombing target, and munitions area feature layers. .... 56

Figure 35. Bayesian data fusion output for the Kirtland site utilizing magnetic anomaly density, manually detected crater density, bombing target, and munitions area feature layers. .... 57

Figure 36. Relationship between ground truth survey data from the Pueblo site and magnetic anomaly density feature layer. (A) depicts the distribution of density values for locations where UXO or UXO-related objects were found while (B) depicts the distribution of density values for locations determined to be free of UXO. .... 59

Figure 37. Probability mass assignments for Dempster-Shafer data fusion for extreme feature intensity values (0 and 1) of layers: helimag anomalies, manually detected craters, known bombing targets, and manually delineated munitions areas. .... 61

Figure 38. Dempster-Shafer data fusion output combining helimag anomaly density, manually detected crater density, known bombing targets, and manually delineated munitions areas at the Kirtland site. .... 63

Figure 39. Probability mass assignments for Dempster-Shafer data fusion for extreme feature intensity values (0 and 1) of layers: thresholded helimag signal with morphological filtering, automatically detected craters, known bombing targets, and manually delineated munitions areas. .... 64

Figure 40. Dempster-Shafer data fusion output combining thresholded helimag signal with morphological filtering, automatically detected craters, known bombing targets, and manually delineated munitions areas at the Kirtland site. .... 66

Figure 41. Dempster-Shafer data fusion output combining helimag anomaly density, automatically detected crater density, known bombing targets, and manually delineated munitions areas at the Pueblo site. .... 67

Figure 42. A preliminary delineation of areas of likely UXO contamination utilizing Dempster-Shafer data fusion output for the Pueblo site (A), threshold set at 0.3. Overlaid on this plot are limited truth data showing agreement with this assessment (B). .... 68

Figure 43. Dempster-Shafer data fusion output combining manually detected craters, known bombing targets, and manually delineated munitions areas. .... 70

Figure 44. Dempster-Shafer data fusion output combining helimag anomalies, known bombing targets, and manually delineated munitions areas. .... 72

Figure 45. UXO assessment (A) before and (B) after incorporation of a heuristic rule blocking magnetic signal in regions containing known interferences. .... 73

Figure 46. Site-wide values for the helicopter magnetometry data density metric. .... 74

Figure 47. UXO assessment (A) before and (B) after incorporation of a heuristic rule incorporating a data density metric for helicopter magnetometry feature layers. .... 75

Figure 48. Dempster-Shafer data fusion output combining thresholded helimag signal with morphological filtering, automatically detected craters, known bombing targets, and manually delineated munitions areas at the Pueblo site. Heuristic rules involving feature blocking and data point density metrics are implemented. .... 76

## **Acknowledgement**

This research was supported wholly by the U.S. Department of Defense, through the Strategic Environmental Research and Development Program (SERDP). The authors thank the SERDP staff and team members for their assistance, particularly Dr. Herb Nelson and Dr. Dan Steinhurst.

## Executive Summary

**Background.** The remediation of sites contaminated with unexploded ordnance (UXO) remains an area of intense focus for the Department of Defense. Current estimates place the total area of possibly UXO-contaminated sites at 10 million acres, with an overall cost of remediation with current methods and sensing technologies in the tens of billions of dollars. Fortunately, studies have estimated that up to 80% of typical sites of potential contamination are actually UXO-free. What is needed to take advantage of this ratio is a means to quickly and reliably scan large sites (on the order of 10,000 acres) in order to rapidly identify regions that are free of UXO and regions that must be subjected to more detailed and time-intensive examination and remediation with established UXO detection tools. Recent investigations have focused on wide-area assessments (WAA) aimed at rapidly determining the approximate density and spatial distribution of UXO objects over regions of wide area, rather than identification of individual UXO objects. Several wide-area assessment projects have been completed under the auspices of Strategic Environmental Research and Development Program (SERDP) and Environmental Security Technology Certification Program (ESTCP). [1] These projects utilized various detection techniques, each with different strengths and weaknesses. However, no single sensing technology has been proven superior in wide-area assessment of UXO. It is therefore logical to examine data fusion approaches, which take advantage of all the available evidence, combining the strengths of each sensing technology while minimizing the weaknesses.

**Objective.** The objective of this work is to develop a data fusion framework that will form the basis of a cohesive data management and decision making utility for processing information acquired in the course of performing wide-area surveys of potential UXO remediation sites. This framework will be capable of capturing UXO-related information from all available data and effectively combining this information to provide site-wide assessments of the likelihood of UXO contamination that are more accurate than any single information source on its own. The final data fusion framework is intended to allow site managers to more efficiently direct the expenditure of time, labor and resources in remediation efforts.

The first year of this project examined potential wide-area assessment information streams, and determined the feasibility of feature selection methods for data fusion. The second year of the project focused on the development of an architecture and data fusion algorithms for the framework. Future work will focus on optimization and refinement of the methodology and algorithms with the development of a prototype implementation of the data fusion framework and its evaluation at several sites.

**Methods.** An overview of the data fusion approach developed in Project MM-1510 is as follows: First, sites of potential UXO remediation are identified and surveyed with multiple wide-area sensing technologies. Additional information about the site, including suspected and known locations of UXO contamination and relevant historical, geological, and topological data, are also acquired. All data sets are geo-referenced, processed,

and/or analyzed to extract UXO-related features. As part of this process, each data set is registered to a common survey area map grid and stored in a data structure for easy reference and subsequent calculations. Each feature set is converted into a site-wide feature map, referred to as a feature layer, which serves as an input into the data fusion framework. Feature layers are registered to the common map grid as well, and stored in a similar data structure. Each feature layer is associated with UXO contamination in a quantitative fashion, and the input feature maps are transformed into assessments of the presence or absence of UXO conditioned on evidence provided by that feature. These quantitative assessments are combined on a point-by-point basis throughout the survey area map grid via a data fusion algorithm to provide a site-wide, overall assessment of UXO contamination. The process is inductive by design. New information and more refined data sets can be incorporated as they become available.

Project MM-1510 leveraged data and feature sets acquired in the ESTCP Wide-Area Assessment Pilot Program (WAAPP). Various feature sets derived from both ESTCP performers and from customized feature extraction algorithms from Naval Research Laboratory (NRL) were developed as inputs for data fusion of wide-area surveys of the Pueblo Precision Bombing Range #2 in Colorado and the former Kirtland Bombing Targets N1 and N3 in New Mexico, along with associated expert information regarding each site. Sensor data included magnetometry data acquired via helicopter survey (helimag), LiDAR data in two resolution scales from airborne surveys, and orthophotography data from airborne surveys. Expert information was comprised of topological maps, geological survey data, and UXO-related features that were manually identified and delineated by expert reviewers (such as bombing targets), or otherwise algorithmically identified (such as magnetic anomalies).

As the data and features from wide-area surveys arrived in various formats, a fundamental aspect of data fusion approach was the development of a standardized input method for sensor data, features, and expert information. A common storage format for subsequent analysis and algorithm development was devised to address this issue. To assess the suitability of feature extraction for data fusion, several automated feature extraction algorithms were developed and evaluated against data supplied from SERDP in the first year of the project. Automated feature extraction methods are desirable as the large size of the data sets acquired in wide-area surveys can make manual identification of features impractical. Methods to transform both manually and automatically extracted feature sets into site-wide feature intensity maps were identified and implemented for each of the three distinct classes of feature type observed in this work: point located, delineated areas, and continuously-valued surfaces.

Finally, data fusion algorithms were developed to effectively combine the information contained within the extracted feature sets and expert knowledge. The two most well-known theoretical data fusion approaches are those formulated by Bayesian inference theory and by Dempster-Shafer theory. Bayesian theory, the older and more established of the two, is founded on a rigorous statistical framework and generally requires significant statistical assumptions regarding the input data. Dempster-Shafer can be thought of as a relaxation of Bayesian theory in which subject belief assessments are

utilized and can be assigned to supersets of hypotheses as well as to single hypotheses. In this work, UXO WAA data-specific approaches based on both theories were implemented and evaluated, along with a simple heuristic approach representing current GIS techniques relying on data overlay and visualization.

**Results & Conclusions.** The major accomplishments for Project MM-1510 are listed below:

- A common geo-referenced database for registering data, features, and associated meta-data resulting from wide-area surveys.
- The evaluation of standard feature extraction strategies on wide-area survey data for their potential in assessing UXO contamination.
- Automatic feature extraction algorithms for identification of craters in LiDAR data and anomalies in helicopter magnetometry data.
- A wavelet filtering algorithm was developed to improve the automatic crater identification algorithm. Filtering reduced false positive crater identifications by removing surface texture artifacts.
- A statistical nearest-neighborhood clustering filter algorithm was developed to improve the magnetic anomaly data by locating anomalies that were part of non-uniform groupings. The clustering algorithm reduced background noise and false positive assessments of UXO in the final data fusion output.
- A prototype Data Fusion Framework based on a hybrid Dempster-Shafer theory.
- A methodology for the input and registration of disparate data and feature streams from wide-area sensing technologies and other site-specific knowledge.
- A methodology allowing the input of meta-information regarding the relationship between these lines of evidence and the presence (or absence) of UXO or UXO-related objects.
- A method by which heuristic rules can be incorporated into the Data Fusion Framework in order to take advantage of specific known interdependencies between feature layers.
- Demonstration of the Data Fusion Framework prototype's ability to successfully provide useful output assessment of UXO likelihood from an assortment of WAA assessment data and features.

The principal accomplishment of Project MM-1510 was the development of a prototype Data Fusion Framework suitable for wide-area assessment of UXO contamination. A key enabling technology was the development of a generalized method for processing input data feature streams from UXO WAA survey efforts. The development of this method is significant, as it requires only a limited number of specifications to be imposed on input features, allowing a wide range of feature sets and relationships to be formatted and input for data fusion. Such flexibility is crucial, as the disparate nature of the data

and features available from potential WAA survey techniques presents a significant impediment towards adoption of more basic data fusion approaches.

Features relevant for wide-area assessment of UXO contamination were successfully obtained from the various data sources. Automatic crater extraction from the LiDAR data was successful. An algorithm based on the circular Hough transform was able to extract a value of four meters as the characteristic diameter of craters at the Pueblo site. Further, a pattern recognition algorithm based on the morphology of the craters was developed to locate them in the LiDAR data. This information was then converted into a feature map describing the density aspect of craters. Feature maps describing the intensity and quality aspects to craters provide data fusion algorithms with additional discriminatory information.

Feature extraction algorithms were also developed for magnetometry data. Due to the minimal geomagnetic features at the Pueblo site, a simple threshold proved effective in eliminating geologic background, which is expected to be more significant at other ESTCP WAA Pilot Program sites. A pattern recognition algorithm was developed to separate ordnance-related signal from the ferromagnetic background of man-made structures. This information was then converted to a feature map describing the density aspect of ordnance-related material. Separate feature maps describing the intensity and morphology aspects of the ordnance-related and man-made components of data are also expected to provide additional discriminatory information for data fusion.

Methods for generating feature layers from extracted features for input to data fusion were developed and implemented. For each input feature set, a corresponding feature intensity map and specification of a functional relationship between a feature's intensity and the likelihoods for or against the presence of UXO that are supported by the feature's intensity values. Heuristic, Bayesian, and Dempster-Shafer theoretic algorithms for combining evidence presented in feature layers were investigated as possible engines for a UXO WAA data fusion framework prototype. These were implemented as MATLAB code and evaluated with feature layers generated from both the Pueblo and Kirtland site data acquired by performers in the ESTCP WAAPP.

The Dempster-Shafer approach, with its ability to quantify uncertainty about evidence, was shown to be the most appropriate data fusion strategy for the UXO problem and proved to be the most successful of the three. The ability to incorporate heuristic rules regarding specific dependencies between input feature layers into the Dempster-Shafer based data fusion framework prototype was described and demonstrated utilizing two specific examples. The first demonstrated a reduction of false positive indications of UXO by utilizing a feature layer comprised of manually identified man-made structures to selectively block magnetometry-derived features. The second demonstrated an adjustment of the impact of magnetometry-derived features on the output assessment of UXO to accurately reflect the uncertainty associated with increased magnetometry data sparseness in some areas of the helimag survey. The prototype data fusion framework developed was able to delineate areas of likely contamination while providing reasonable estimates of the likelihood of that contamination given supporting observational evidence

and a priori knowledge. Preliminary results were compared with limited ground truth data available at the Pueblo site and agreed well.

**Benefits.** The key theoretical advantage of a data fusion approach to wide-area assessment is the ability to reduce false positives while retaining high detection rates. The framework described is flexible, tolerating missing data and allowing multiple configurations of potential input data streams, as well as scalable, allowing new data streams to easily be included in the assessment. Further, the impact of available and new data streams on the output can be readily quantified. One challenge is that the structured input methodology requires the specification of each feature layer's relationship to the presence or absence of UXO. However, the input methodology allows specification to be accomplished in a highly flexible manner. The user has the ability to input specifications that vary from simple, intuitive estimations based on expert knowledge to detailed functional relationships based on empirical evidence of sensor performance. Thus, the data fusion framework is capable of utilizing all the information and observation evidence available, without necessarily requiring that the exact same inputs be present for assessment. This flexibility is an important feature of the data fusion approach as it is expected that, for a number of reasons, it will rarely be the case that exactly the same types or quality of data will be available for analysis each time a wide-area UXO assessment is performed.

**Future Applications.** MATLAB was chosen as the computational platform for data exploration and initial feature and algorithm development. [2] An extremely useful research tool, MATLAB provides an extensive code base of algorithmic resources together with the ability to rapidly port prototype implementations to other computational platforms. Future work is needed to further the development, evaluation, and optimization of a prototype Dempster-Shafer based data fusion framework with the intended result being a prototype software implementation of the data fusion framework suitable for demonstration and evaluation at various sites of potential UXO contamination. A final, production-grade data fusion framework that is well-suited for independent use by site administrators would most likely require close coordination with SERDP and potential end-users and partnership with an existing GIS software vendor.

## Objective

The objective of project MM-1510 was to enhance assessment of unexploded buried ordnance (UXO) in large geographic areas through data fusion of outputs of multiple sensing technologies with any available expert knowledge of the sites in order to reduce overall false alarm rates. Project MM-1510 centered on the development of a cohesive data management and decision making strategy for analyzing UXO survey data and site-specific information. The development of such a data fusion framework required the development of a customized algorithmic data fusion approach that enabled the effective combination of all available knowledge about a UXO survey site (including both data feature analysis and manually-acquired site information) into output wide-area maps of potentially contaminated sites delineating areas of likely UXO contamination from those that were likely to be free of UXO. The intention was that these assessments would serve to direct the acquisition of data using more accurate, sensitive, and costly local-area UXO surveys. Data from local-area surveys can then be fed back into the data fusion framework and used to further refine maps of these regions and to provide more detailed assessments and allowing more efficient direction of time, labor, and resources.

This report serves as the final report for project MM-1510 and describes two years of accomplishments in the development of a viable prototype data fusion framework for wide-area assessment. The first year objective was to determine the feasibility of feature extraction methods to provide complementary data from the various detection sources that was also suitable for intelligent data fusion. For this effort, a geo-referenced database container was constructed to hold the data and several successful automated feature extraction techniques were developed. The overall goal for the second year of Project MM-1510 was the development of appropriate architecture and algorithm components of the data fusion framework for enhancing wide-area assessment of UXO. Towards this end, data fusion strategies and algorithms were explored and developed, feature layer development efforts begun in year one of the project were continued with an eye towards ensuring compatibility with and suitability for potential data fusion frameworks, and a prototype implementation of the developed data fusion framework was created to test and evaluate output assessments.

## Background

### Problem Statement

Current estimates place the total area of possibly UXO-contaminated sites at 10 million acres, with an overall cost of remediation with current methods and sensing technologies in the tens of billions of dollars. Studies have estimated that up to 80% of typical sites of potential contamination are actually UXO-free, indicating a need to quickly and reliably scan large sites (on the order of 10,000 acres) in order to rapidly identify regions that are

free of UXO and regions that must be subjected to more detailed and time-intensive examination and remediation with established UXO detection tools.

Effective wide-area UXO assessment is centered on the capability to rapidly scan large tracts of land and obtain relevant, useful information in the process. Two possible modes of enhancement to wide-area UXO assessment are, first, decreasing the false alarm rate of current UXO sensing technologies, and second, utilizing alternate sensing technologies and survey methods that scan larger areas or more rapidly cover large areas than current sensing platforms.

## **Current Technology**

Standard, ground-based UXO sensing technologies include methods such as vehicular-mounted time and frequency domain electromagnetic induction (EMI), total field magnetometry (TFM), and ground penetrating radar (GPR).<sup>\*</sup> These UXO sensors have been deployed on ground-based platforms such as portable devices, push carts, and towed sensor arrays. Standard analysis methods of these types of data are well described in the literature and have been implemented with success. These methods typically rely on generating theoretical sensor response models, or measuring pure responses of various UXO items and then comparing survey data to these models in order to make a detection or classification. Typically, the ability of such techniques to cover wide survey areas is limited, although current SERDP projects are assessing the utility of ground survey transects to increase this capability for wide-area applications.

As the sensitivity/ ground penetration depth of direct UXO sensing technologies, such as magnetometry, drops steeply with distance, alternate techniques for wide-area assessment will generally not have the benefit of sensing deeply buried UXO objects directly. Instead, they must rely instead on sensing UXO-related phenomena like spectral chemical signatures, variations in heat capacity, and measurements of surface clutter and micro-topological features. The techniques of synthetic aperture radar (SAR), light detection and ranging (LiDAR), and high-resolution aerial photography all yield information about micro-topological features and surface clutter of wide areas under assessment. These sensors have the benefit of functioning over much greater distances than electromagnetic sensors, allowing them to be deployed on fixed wing aircraft. Towards this end, six wide-area assessment projects have been completed under the auspices of SERDP and ESTCP. [1] These projects utilized techniques that varied from airborne infrared laser imaging combined with thermal imaging (UX-9523), synthetic aperture radar (UX-0126, UX-1070, and UX-1173), and airborne magnetometer arrays (UX-0031 and UX-3002). Additionally, ESTCP has sponsored a “Wide-Area Assessment Pilot Program” to fund development of wide-area UXO sensing technologies. [3]

Wide-area sensing platforms are typically designed and optimized to detect UXO or UXO-related features without correlation to other sources of data or expert information.

---

<sup>\*</sup> See “*Proceedings of SPIE: Detection and Remediation Technologies for Mines and Minelike Targets V through VIII*” and “*Transactions on Geoscience and Remote Sensing, Vol 2, No. 3, (2001)*” for examples.

Data are acquired and processed via algorithmic feature extraction, manual inspection and feature identification, or some combination of both. Final site assessment is provided by evaluating the extracted UXO-related features, taking into account a priori knowledge as to how accurate the features are at indicating the presence of UXO and/or UXO-related objects. Thus, two types of information are obtained from sensing: UXO-related features extracted from survey data, and meta-data regarding the relationship between UXO and these features. Further information can be obtained from manual assessment of the site. Areas associated with UXO-related activities can be delineated through visual inspection of sensing data or through knowledge of historical usage patterns at the site. Examples of this type of information include delineations of visible bombing targets and descriptions of known munitions ranges.

Thus, at each survey site, site managers potentially have multiple, disparate lines of evidence for or against the presence of UXO or UXO-related items. The goal of data fusion is to produce the best assessment of UXO contamination at any location in the site, given all the available evidence, and in turn, allowing site managers to more accurately delineate areas of likely UXO contamination.

### **Data Fusion Approach**

To date, effective wide-area assessment has been hindered by a lack of accurate target and range information. In addition, no single sensing technology has been both accurate and cost-effective in surveying entire sites. One potential avenue of enhancing wide-area assessments of UXO is to combine multiple data streams, taking advantage of the strengths of individual sensing platforms, while minimizing the weaknesses. [4]

Limited data fusion approaches have been explored, although these tend to be data-stream specific rather than generic data aggregators and decision utilities. [5-8] The approach undertaken in this effort is to enhance wide-area assessment of UXO by developing an algorithmic framework for data management and data fusion that provides a structured, intuitive means for incorporating and inputting UXO WAA survey sensor data, algorithmically-generated data features, manually located features, site-specific information, and any relevant associated meta-data, such as sensor performance characteristics and specific heuristic rules regarding data stream interdependencies.

The central task in this effort is to find a structured way of combining these multiple lines of evidence, and information. The nature of the wide-area UXO assessment problem suggests three main criteria for any proposed data fusion framework. The first criterion is that a potential data fusion framework must be capable of accepting a variable number of disparate data streams as input. Different types of information are conveyed by each stream, and not every line of data will be available at every geographic grid point in the survey area. Thus, the data fusion framework must be capable of producing output assessment that utilize the data available, rather than a fixed set of data inputs.

The second criterion is that a potential data fusion framework has a structured, intuitive means of capturing the relationship between a given data feature or other form of

information and the likelihood of the presence of UXO. The reality of the UXO WAA problem is that extensive ground truth data will likely never be available, due to the considerable expense required to acquire it. This complicates efforts at statistically modeling the relationship between extracted features and UXO contamination and makes precise probabilistic assessments of UXO contamination highly difficult to achieve. Thus, it makes sense to focus on approaches that allow the user to incorporate and benefit from knowledge that is known regarding the specific performance characteristics of the individual sensing technologies, as well as that which can be gleaned from the expertise of seasoned site managers and other program personnel.

Further, if expert information is known regarding the performance characteristics of one of the wide-area assessment technologies in detecting UXO, then the data fusion framework should be capable of directly accepting and utilizing this information. Likewise, if the relationship between an information stream and UXO is not well-known, then there should be a structured, intuitive means of inputting reasonable estimates of this relationship into the data fusion framework. These relationships should not be “hard wired” into the framework itself, as this would not allow for updating in the face of improved knowledge of sensor characteristics, or in site-specific conditions that may alter such relationships from site to site. Finally, the data fusion framework should be able to incorporate these relationships independently of their functional form to reflect the fact that, in some cases, complex functional dependencies may be known from empirical testing while in others, only simple estimates arising from the site manager’s personal experience may be available.

Finally, the third criterion is that a data fusion framework must be reasonably amenable to the incorporation of specific heuristic rules that would otherwise be difficult to capture in a strictly probabilistic fashion. These rules may reflect specific, known interdependencies between two or more different input data streams and would operate on the input information to adjust the output assessments accordingly.

## Methods

In the broadest terms, data fusion for wide-area assessment of UXO proceeds as follows: multiple data sets utilizing different wide-area sensing technologies are acquired at a site potentially in need of UXO remediation. Additional information about the site, including locations of potential UXO contamination and relevant historical, geological, and topological data, are also acquired. All data sets are geo-referenced, processed, and/or analyzed in some fashion to extract UXO-related features. Each feature set is converted into a site-wide feature map which is input into the data fusion framework. Each type of feature is associated with UXO contamination in a quantitative fashion, and the input feature maps are transformed into assessments of the presence or absence of UXO, conditioned on evidence provided by that feature. These quantitative assessments are combined on a point-by-point basis throughout the survey area via a data fusion algorithm to provide an overall assessment of UXO contamination. The process is inductive by design. New information and more refined data sets can be incorporated as they become available.

Fundamental to the data fusion approach is the development of a standardized input method for sensor data, features, and expert information. To assess the suitability of feature extraction for data fusion, several automated feature extraction algorithms were developed and evaluated against data supplied from SERDP in the first year of the project. The specification of a common format for features from varying sites was also a focus of year one and has continued to figure importantly in work done in year two. The approaches pursued are described in the sections on automated feature extraction and feature layer development that follow.

With the inputs established, the development of a data fusion architecture began with the formulation of an appropriate frame of discernment for the problem of wide-area assessment of UXO. Several theoretical data fusion approaches were evaluated for their suitability. Methods for performing data fusion with heuristic, Bayesian theoretic and Dempster-Shafer theoretic approaches are presented below. Finally, a prototype implementation of a data fusion framework embodying the identified data fusion strategies was constructed in software in order to evaluate the strengths and weaknesses of the approaches and to examine possible avenues for optimization.

Project MM-1510 leveraged data acquired by SERDP/ESTCP programs, and in particular, data acquired for two or the three former Department of Defense sites during the ESTCP Wide-Area Assessment Pilot Program. The two sites surveyed in 2004 and 2005 were: the Pueblo Precision Bombing Range #2 in Colorado (CO), the former Kirtland Bombing and Targets N1 and N3 in New Mexico (NM). The principle sensing modalities utilized in each survey were low altitude (helicopter) airborne magnetometry, often referred to as "helimag," high-altitude airborne LiDAR, and orthophotography. Along with sensor data, a number of extracted features were made available. These included manually identified regions of interest from LiDAR and photographic data, and magnetic anomalies from helimag data. Additional feature sets were generated

algorithmically in the first year of Project MM-1510, including automatically identified craters and magnetic signatures. [9] Although the work described in this report was applied to these data sets, it should be noted that the underlying approach developed for assimilating data is intended to apply to any UXO-related feature data that is available for inclusion in the data fusion framework.

### **Automated Feature Extraction**

The first steps were the acquisition, conversion, and registration of data sources to a common grid and format. The principle sensing modalities utilized in each survey were low altitude (helicopter) airborne magnetometry often referred to as helimag, and high-altitude airborne LiDAR and orthophotography. MATLAB [2] was chosen as the computational platform for initial algorithm prototyping, and this choice dictated the data format. MATLAB provided an extensive code base for rapid prototype development, and the ability to port MATLAB implementations to other computational platforms. For a production platform, it was expected that a partnership of NRL, SERDP, and an existing GIS software vendor would be the preferable data analysis platform and the subject of future work.

Two grid spacings were selected for feature development, half meter and one meter. Note that it is possible to develop features in the native space of the data and then register the features to a common grid. However, this approach was not taken as it does not allow for an examination of the impact of data resolution on feature extraction, which may be relevant to data acquisition costs.

Available ground truth data and auxiliary information regarding potential UXO contamination at the selected sites were acquired to assess their potential to enhance data fusion. Archive search reports for each site that consisted of items such as historical records of site usage, data from previous surveys, first hand experience of local inhabitants, and other sources of expert knowledge were also obtained for review. A literature survey was performed in order to research current UXO data analysis strategies and to assess their utility for extracting features for data fusion. Software packages commonly used in the evaluation of UXO data were obtained and evaluated for their ability to efficiently process large data streams, to locate features relevant for wide-area assessment, and to export such information to the development platform. Data provided in multiple, algorithmically-altered forms were evaluated to determine the impact of the alterations to subsequent feature analyses and data fusion.

The development of automated feature extraction algorithms began with investigations focused on locating features relevant to UXO assessment within data gathered, and also on determining what types of signal processing algorithms would be useful in extracting those features in an automatic fashion. Utilizing expert information and ground truth regarding the survey site, regions of interest were located. Data features exhibiting a causal or correlative link with UXO-related items were flagged and examined as well as features associated with the absence of UXO-related items. The utility of various data analysis techniques was evaluated for potential in automatic extraction of these features,

or otherwise marking regions associated with known UXO contamination. The acquired raw and processed sensor data were also examined to evaluate the ability of external signal processing techniques to remove redundancies and to identify previously unseen correlations and features in and between the data sets.

The approach to feature analysis was based upon examining three aspects of identified features, their density, intensity, and morphology. The density aspect of identified features quantified the number of features identified per unit survey area, irrespective of feature strength or shape. The intensity aspect of features quantified their strength, and the morphology aspect quantified their shape. For a variety of reasons, it was not expected that, on a feature by feature basis, all three of these aspects would necessarily prove to be simultaneously useful to achieving this goal.

### **Feature Layer Development**

In general, UXO-related features are extracted from acquired data either manually or through application of a feature extraction algorithm. These features fall into one of three categories of feature type: features with continuously-variable intensity, binary features that are point located, and binary features resulting from the delineation of regions of interest. For example, potential crater locations can be extracted from LiDAR and aerial photography through either visual inspection or automated crater detection algorithms. These features are binary in nature and point-located. Other extracted features could be continuously valued parameters, such as background-corrected total field magnetometry signals extracted from magnetometry data. In this case, the extracted features have both location and intensity. Finally, manual assessment of the site, either through visual inspection of the data or through knowledge of historical usage patterns at the site, can result in delineated areas that are associated with UXO-related activities. Examples of this type of data include delineations of visible bombing targets and known munitions ranges.

Extracted features constitute multiple, disparate lines of evidence for or against the presence of UXO or UXO-related items at each survey site. While each feature can be visually displayed, and even overlaid on a site map, the disparity between the different feature sets makes it difficult to algorithmically combine the evidence provided by them into an overall assessment. In order to overcome this difficulty, a data fusion architecture was developed in which individual feature types were used to produce corresponding feature “layers.” In this approach, the information encapsulated by each available extracted feature was converted into a map indicating the strength of that feature across the survey site. The map formulation allowed for a pixel-by-pixel determination of feature specific evidence related to UXO contamination throughout the survey site. The salient property of each feature layer was that it conveyed a site-wide map filled with feature intensity values ranging between a known minimum and maximum value. Survey grid points where no feature information was available were allowed and were treated as missing data or estimated via interpolation from nearby data. Along with a structured representation of each layer’s relationship to UXO contamination, these feature layers were utilized as inputs into subsequent data fusion algorithms responsible for

synthesizing the evidence presented by each feature layer into a site-wide, point-by-point assessment of the likelihood of UXO contamination.

Generation of individual feature layers was accomplished within the MATLAB computational platform. Feature layers were organized in the same file structure as that used to store imported survey data. [9] The method by which each feature layer was generated depended entirely on the nature of the extracted feature. Continuously-valued features, such as magnetometry survey data, were used directly as feature layers. Manually delineated feature areas, such as visually identified bombing targets or historically known munitions areas, were converted into a binary feature layer in which grid points within the features were given a value of one and points outside a value of zero. Delineated feature areas for bombing targets were further dilated to encompass an “effective area” that more closely resembled the likely area of UXO contamination from repeated bombing runs. Collections of point-location binary features, such as the locations of identified craters or magnetic anomalies, were transformed into site-wide feature density estimates via a kernel density estimation (KDE) algorithm. [10] For simplicity, the generated feature density maps were scaled according to their maximum value, and thus reflected the relative density of the feature across the site.

By utilizing a feature density estimate, the likelihood of UXO contamination was correlated to the intensity of feature activity in a nearby region, rather than to the presence or absence of the feature itself at any specific point. The size of the region of influence was chosen to roughly correspond to the effective area of a bombing target by doubling the approximate length of the major axis of each target and replacing the target feature with a circle of that diameter, centered upon the bombing target’s location. In kernel density estimation, points most distant from the center point being estimated are scaled to make the least contribution, nearby points the most contribution. Gaussian scaling was used for all KDE-generated feature density maps.

### **Data Fusion Architecture Development**

The next task was to develop a data fusion methodology to accurately assess UXO contamination given the evidence provided by an arbitrary set of feature layers. Three algorithmic approaches were considered for implementation: a purely heuristic-based linear combination, a Bayesian theoretic approach, and a Dempster-Shafer theoretic approach. In addition to developing an algorithmic approach for combining the evidence presented by multiple feature layers, a standardized, structured means for defining the relationship between each feature layer and the presence of UXO was required.

The simplest approach involved the use of heuristic rules to weight each feature layer according to its contribution to the evidence that UXO was present and then summing the contributions to form an output assessment. Intuitively, a heuristic approach is equivalent to overlaying each available feature layer in a visual display, and then adjusting the opacity of each layer according to its importance in signifying UXO contamination. Specific heuristic rules can come from expert knowledge regarding the features and their relationship to UXO, or through analysis of empirical data to discern relationships. These

rules express UXO likelihood as a function of feature layer value (i.e., intensity) that varies between zero and one, depending on the degree to which that feature layer expresses support for the presence of UXO. While a heuristic approach is simple computationally and allows for detailed input of specific UXO-feature relationships, a number of problems present themselves in implementation. Chief among these is the lack of a measure of uncertainty or probability in the output of such an algorithm, and no means for incorporating such information as an input. This lack puts reliable information streams on par with less reliable lines of evidence in terms of its impact on the final output assessment. Additionally, experience with fusion of multiple sensor outputs at NRL has shown that a simple “AND” operation such as this often provide little benefit as they typically increase false alarms along with detections. For this reason, the purely heuristic approach was eliminated from further consideration.

Next, a Bayesian theoretic approach was considered. Bayesian reasoning and Bayesian belief network architectures are powerful tools for inductive reasoning within a probabilistic framework. [11,12] The core of the approach was formulated from Bayes’ rule, eq. 1, which states that for a hypothesis,  $H$ , and evidence,  $E$ , the probability of the  $H$ , conditioned on  $E$  is:

$$p(H/E) = p(H) \times \frac{p(E/H)}{p(E)} \quad (1)$$

The first term,  $p(H)$ , is known as the prior probability of  $H$  and represents knowledge of  $H$  before observing evidence  $E$ . The second term is known as the normalized likelihood and represents knowledge of how likely it is to observe  $E$ , given  $H$ . Multiplication of these two terms provides the posterior probability of  $H$ , given observed evidence  $E$ . If multiple lines of evidence are available, the joint distribution must be considered as in , eq. 2.

$$p(H/e1, e2, e3, \dots) = \frac{p(H, e1, e2, e3, \dots)}{p(H, e1, e2, e3, \dots) + p(\neg H, e1, e2, e3, \dots)} \quad (2)$$

Fortunately, in many cases calculating the joint distribution of several variables is made simpler by examining the causal relationships among the variables and identifying those which are independent. A specification of these relationships and the conditional probability values associated with them forms what is known as a Bayesian belief network.

A diverse range of applications have been described for belief networks, such as automated medical diagnosis and intelligent software help utilities. Of particular utility is the ability of such networks to optimally estimate unknown network parameters from known parameters, allowing, for example, diagnosis of a particular disease with a limited set of observations and test results. This ability, however, stems from the belief network’s complete encapsulation of all pertinent probabilistic relationships between

network parameters. As such, belief networks are only useful when an appropriate network topology is defined with appropriate causal links such that the resulting conditional probability distributions assigned to each link are fully specified. This is problematic for the UXO WAA task, as estimating either of these quantities can be difficult, and possibly site-specific. In other words, it's possible that not only will the types of evidence available change from site to site and, indeed, from location to location within a given site, but also that the conditional probability distributions and network topology itself will also change.

A simpler, so-called “naïve” Bayesian theoretic approach is to assume independence amongst each line of evidence, as in eq. 3. Beginning with an initial prior probability for UXO, the prior can be updated to a posterior probability of UXO by multiplying it with the normalized likelihood of each available line of evidence.

$$p(UXO/E) = p(UXO) \times \prod_i^n \frac{p(e_i | UXO)}{p(UXO)p(e_i | UXO) + p(\neg UXO)p(e_i | \neg UXO)} \quad (3)$$

where  $E$  is the set of available lines of evidence:

$$E = \{e_1, e_2, e_3, \dots, e_n\}$$

Utilizing this formulation, a naïve Bayes data fusion algorithm was implemented within MATLAB. In this implementation, conditional probability functions for  $p(e/UXO)$  and  $p(e/\neg UXO)$  must be defined across the range of observed feature intensity values for each feature data layer, representing the probability of observing a particular feature value, given that UXO is present and not present, respectively. These functions were defined a priori and were estimated from knowledge regarding the feature's relationship to UXO. To begin the data fusion process, an uninformed prior probability of UXO in the absence of any evidence was assumed, and was represented by setting  $p(UXO)$  to an initial value of 0.5 at each point in the survey grid. For each input feature layer,  $p(UXO)$  was updated at each point in the survey grid to a posterior probability of UXO,  $p(UXO/E)$ , given the evidence observed in that feature layer. The specified conditional probability functions were used to convert feature intensity values into conditional probability assignments on a point-by-point basis across the survey grid. These assignments were then used to calculate the normalized likelihood ratio at every point in the survey grid and then multiplied by the corresponding prior probabilities and generating a posterior probability for every grid point. These posterior probabilities were then updated with subsequent feature layers in the same fashion, until all available feature layers were included. The resulting output was a single site-wide map containing values for the posterior probability that UXO was present, given the observed feature data layers. Thus, values close to one indicated a high likelihood that UXO was present, those close to zero indicated little likelihood that UXO was present, and those close to a value of 0.5 indicated relative uncertainty as to whether UXO was or was not present.

The main difficulty in implementing a Bayesian theoretic approach is that it requires specification of the conditional probabilities  $p(e/UXO)$  and  $p(e/\neg UXO)$ , which can be

difficult to do in an intuitive, rational manner depending on the nature of the feature layer and its relationship to UXO. A related problem is that within this formalism, a lack of evidence for a supposition is generally interpreted as evidence for that supposition's negation (i.e.,  $p(x) + p(\neg x) = 1$ ). For example, if a feature layer represents a direct detection of UXO-related objects, then a specific feature layer value can be considered as a detection threshold and  $p(e/UXO)$  can be related to the feature's UXO detection rate at that threshold, and  $p(e/\neg UXO)$  can be related to the feature's false positive rate at that same threshold. However, other feature layers may encode a relationship that is difficult or even nonsensical to parse in this manner. For example, crater density may be indirectly associated with UXO-related objects, but cannot be viewed as a detection mechanism in the classic signal detection framework, as it would involve assigning a probability to particular crater density conditioned on the fact that UXO is or isn't present. Unfortunately, these probability assignments are tenuous at best due to the lack of a causal mechanism between the two features. For the UXO problem, the fundamental difference between the two lines of evidence,  $p(e/UXO)$  and  $p(e/\neg UXO)$ , is the type of uncertainty associated with them. The former (i.e.,  $p(e/UXO)$  that links crater density with UXO-related objects) is associated with objective uncertainty and is well-described by traditional probability theory. The latter is associated with subjective uncertainty in so much that a lack of craters reflects ignorance towards either the presence or absence of UXO contamination.

In order to avoid the complications encountered in probability assignments for the Bayesian method, an approach based on Dempster-Shafer theory was also developed. Briefly, Dempster-Shafer theory can be described as a generalization of Bayes theory in which observed evidence can support not only specific hypotheses, but also sets of hypotheses. [13,14] The approach is attractive for problems involving evidence pooling from multiple sources because it does not require complete specification of the underlying conditional probabilities and it allows for assignment of a degree of belief to a specific hypothesis without necessarily assigning any belief to the negation of that hypothesis. These properties enable Dempster-Shafer frameworks to address evidence associated with subjective uncertainty in a more satisfying manner while simultaneously retaining the ability to incorporate evidence associated with objective uncertainty. [15] Dempster-Shafer data fusion has been utilized in applications such as land cover classification, machine vision, and medical diagnoses. [16-18]

An example of Dempster-Shafer theory applied to sensors follows: consider a hypothetical sensor system that attempts to identify a sensed object as belonging to one of four object types: A, B, C, or D. The exhaustive and mutually exclusive hypothesis space, or frame of discernment, is represented by the set  $\{A,B,C,D\}$ . The evidence observed from a given sensor response provides support for one or more of these elements including, potentially, supersets combining two or more elements. Assignment is accomplished through the apportionment of unit probability mass across the focal elements for which the sensor's response provides support. The amount of mass assigned to any given focal element is representative of the relative amount of certainty with which the sensor response can make that declaration. For example, say the output of sensor  $I$  indicates that the sensed object is of type A or B with 80% certainty. Thus, an

assignment of 0.8 is made to the hypothesis that the object is type A or B,  $m_1(A,B) = 0.8$ . The remaining mass of 0.2 can be assigned to this hypothesis's negation, namely that the object is type C or D,  $m_1(C,D) = 0.2$ . Alternatively, it can be assigned to superset of all elements in the frame of discernment, reflecting total uncertainty as to the object's identification,  $m_1(A,B,C,D) = 0.2$ . The former assignment is useful in situations where the evidence is well-described by traditional probability assessments, while the latter assignment is useful in describing evidence associated with subjective uncertainty or incomplete probability specifications.

Probability mass assignments from two sensors are pooled according to Dempster's rule of combination. Pooling is accomplished by calculating the cross products of the probability mass assignments for each sensor. Each cross product is assigned to a hypothesis that represents the intersection of those of the two component probability masses used to generate it. Cross products assigned to the same hypothesis are summed, resulting in a series of unique output hypotheses with corresponding probability mass assignments. If any cross products result from conflicting hypotheses, they are removed from the set and the remaining masses are normalized to sum to unit mass. Such renormalization has the effect of redistributing the mass assigned to conflict proportionally across the remaining hypotheses. For example, eq. 4 combines the evidence provided by the sensor in the previous example with that of a second sensor with probability mass assignments,  $m_2(A,C) = 0.7$ ,  $m_2(A,B,C,D) = 0.3$ .

$$\begin{array}{c}
 \text{Sensor 1} \left\{ \begin{array}{l} m_1(A,B)=0.8 \\ m_1(A,B,C,D)=0.2 \end{array} \right. \quad \boxed{\begin{array}{cc} m_{1,2}(A)=0.56 & m_{1,2}(A,B)=0.24 \\ m_{1,2}(A,C)=0.14 & m_{1,2}(A,B,C,D)=0.06 \end{array}} \quad (4) \\
 \underbrace{\begin{array}{cc} m_2(A,C)=0.7 & m_2(A,B,C,D)=0.3 \end{array}}_{\text{Sensor 2}}
 \end{array}$$

Here, the net effect of data fusion is the assignment of most of the probability mass to  $m(A)$  ( $=0.56$ ), a reduction of those assigned to  $m(A,B)$  ( $=0.24$ ) and  $m(A,C)$  ( $=0.14$ ), as well as a reduction of the mass assigned to uncertainty  $m(A,B,C,D)$  ( $=0.06$ ).

Finally, the quantities of support and plausibility for each focal element are central to Dempster-Shafer theory. Support is defined as the sum of probability masses that can be directly attributed to that focal element. Plausibility is defined as the sum of probability masses not assigned to the focal element's negation. Together, the two quantities form an uncertainty interval bounded on the lower side by the support and on the upper side by the plausibility. The uncertainty interval conveys information regarding the proportion of evidence that directly supports a focal element versus that which merely fails to negate it. For example, support for A in the example above is 0.56 following fusion, while the plausibility is 1. This indicates that A receives partial support from the evidence, while no evidence directly refutes it. Thus, an uncertainty interval of (0,1) indicates complete uncertainty, intervals of (0,0) and (1,1) indicate focal elements known to be false and

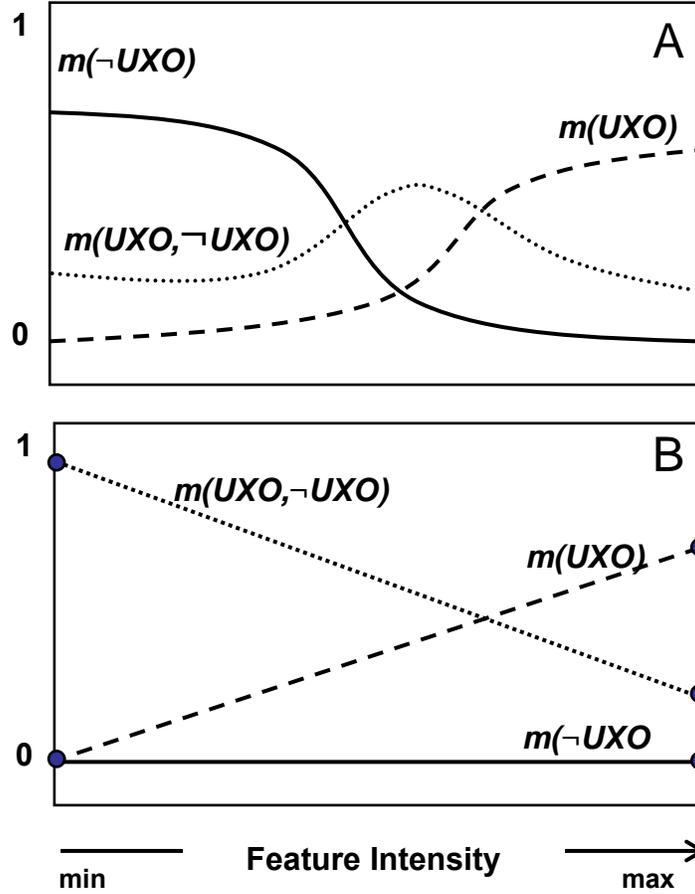
true, respectively, and intervals in which the support and plausibility are equal indicate that all of the evidence available directly supports that focal element.

A customized Dempster-Shafer data fusion algorithm was implemented within MATLAB. For this implementation, a frame of discernment of  $\{UXO, \neg UXO\}$  was used. Thus, there were three separate focal elements to which probability assignments can be made by feature-based evidence:  $\{UXO\}$ , reflecting the belief that UXO was present,  $\{\neg UXO\}$ , reflecting belief that UXO was not present, and  $\{UXO, \neg UXO\}$ , reflecting ignorance (i.e., uncertainty) regarding the presence of UXO.

As with the Bayesian approach, the heart of the implementation involved the input of a series of feature layers in the form of site-wide maps of feature intensity values. Each feature layer was accompanied by a corresponding probability mass assignment function that related the probability mass assignments for each focal element to feature intensity over the range of feature intensities observed.

These functions were assigned a priori and were based on expert knowledge of the feature and its relation to the presence of UXO. If only limited knowledge is available, simpler estimations can be made. Two hypothetical examples are shown in Figure 1, the first involving a specified functional dependence and the other based simply on probability mass assignments made at the extreme values of feature intensity and estimated through linear interpolation elsewhere. The purpose of these probability mass functions was to convert feature intensity values on a point-by-point basis to probability assignments for each of the focal elements of the frame of discernment for every point in the survey grid at which the feature layer was defined. At points on the grid where feature intensity values were missing, the algorithm was programmed to assign all probability mass to uncertainty, reflecting that this situation represents a lack of knowledge. This process results in a site-wide map of probability mass assignments for each focal element for each feature layer.

These resulting probability mass assignments were then combined on a point-by-point basis throughout the survey grid according to Dempster's rule of combination, as illustrated in Figure 2.



**Figure 1.** Hypothetical probability mass assignments in Dempster-Shafer data fusion. Example (A) depicts the assignment of complex functions to each of the three focal elements over the observed range of feature intensity values. Example (B) depicts an assignment where probability masses are assigned to the feature intensity extremes for each of the three focal elements and intermediate values are estimated through linear interpolation.

$$\begin{array}{l}
 \text{Feature Layer 1} \left\{ \begin{array}{l} m_1(UXO) \\ m_1(\neg UXO) \\ m_1(UXO, \neg UXO) \end{array} \right. \begin{array}{|c|c|c|} \hline m_{1,2}(UXO) & m_{1,2}() & m_{1,2}(UXO) \\ \hline m_{1,2}() & m_{1,2}(\neg UXO) & m_{1,2}(\neg UXO) \\ \hline m_{1,2}(UXO) & m_{1,2}(\neg UXO) & m_{1,2}(UXO, \neg UXO) \\ \hline \end{array} \\
 \underbrace{\begin{array}{ccc} m_1(UXO) & m_1(\neg UXO) & m_1(UXO, \neg UXO) \end{array}}_{\text{Feature Layer 2}}
 \end{array}$$

**Figure 2.** Dempster's rule of combination implemented for a UXO assessment application.

At each point in the survey map, the cross products of the probability mass assignments from the three focal elements of the UXO frame of discernment form a set of nine terms. Three of these support the focal element  $\{UXO\}$ , three support  $\{\neg UXO\}$ , one supports  $\{UXO, \neg UXO\}$ , and two represent evidence assigned to null set,  $\{\}$ , which indicates the amount of conflict between the different lines of evidence. The output probability masses for  $\{UXO\}$ ,  $\{\neg UXO\}$ , and  $\{\}$  were calculated by summing the appropriate terms. Evidence was combined in a serial fashion, with the probability mass assignments of additional layers combined with the output assignments generated by Dempster-Shafer combination of the previous layers. In order to preserve the transitive nature of this operation (i.e., to ensure that evidence combined in any order provided the same output) assignments made to the null set were propagated separately until the final feature layer was combined. Final output was generated by unit normalization of the non-empty frame element assignments:  $\{UXO\}$ ,  $\{\neg UXO\}$ , and  $\{UXO, \neg UXO\}$ .

With the final output, decisions regarding area delineation can be made that take into account both the weight of evidence attributed to UXO being present and the weight assigned to uncertainty. The values of the Dempster-Shafer quantities of support and plausibility for focal element  $\{UXO\}$  at each point in the survey map can be calculated as, respectively, the output assignment to  $\{UXO\}$  itself, and one minus the output assignment to  $\{\neg UXO\}$ , or, equivalently, the sum of the assignments output to  $\{UXO\}$  and  $\{UXO, \neg UXO\}$ .

## **Results and Discussion**

Sensor data and UXO-relevant feature sets were obtained from ESTCP performers in that agency's wide-area assessment pilot program. To augment the ESTCP-generated feature sets, automatic feature extraction strategies and algorithms were evaluated and developed for each data set. As previously discussed, automatic feature extraction algorithms are desirable in wide-area assessment applications due to the potential enormity of the acquired data sets. For each feature set, an appropriate strategy for conversion to a site-map feature layer for input to data fusion was developed and implemented. Heuristic, Bayesian, and Dempster-Shafer algorithmic approaches for fusing feature layers into UXO assessments were developed and implemented in software prototypes. Evaluation of the prototypes demonstrated that the Dempster-Shafer formalism was the best data fusion approach for the wide-area assessment. Finally, approaches for implementing specific heuristic rules to incorporate particular feature interdependencies were developed and implemented in the prototype data fusion framework.

### **Data Registration and Importation**

Data acquired during the ESTCP WAAPP survey of both the Pueblo Precision Bombing Range #2 and former Kirtland Bombing Targets N1 and N3 were obtained through SERDP. The survey data consisted of image files containing aerial LiDAR and orthophotographic surveys, and a series of text files containing data from helicopter magnetometry survey. The data were imported into a common format for subsequent viewing and analyses within the MATLAB computational environment. A suite of MATLAB routines was assembled to provide a turn-key approach to importing WAA data from multiple different formats into this common format, allowing for convenient and rapid access to all survey data from within the MATLAB workspace, and making possible the rapid importation of new survey site data in the future.

A common survey map grid was specified for both the Pueblo and Kirtland site data sets. The grid parameters were based on an examination of the furthest extents of each of the different survey data sets. For the Pueblo site, this grid spanned the range between UTM Northing 4,169,390 to 4,178,634 and UTM Easting 614,590 to 618,627 (Zone 13), and represented a rectangular survey area of approximately 37 square kilometers. For the Kirtland site, the grid spanned the range between UTM Northing 3,888,460 to 3,894,568 and UTM Easting 330,026 to 337,537 (Zone 13), and represented a rectangular survey area of approximately 46 square kilometers. In each case, the grid was divided into regularly sized blocks that were 1024 meters to a side, starting at the SW corner and moving south to north and west to east. For each block, the appropriate data were loaded from the source file and interpolated to fit a 1m or 0.5m resolution grid. The data were saved on a block by block basis in a MATLAB database where each data layer occupied a different field in the structure. Thus, correctly registered, survey-wide data were made available for subsequent feature extraction and processing.

Figures 3 and 4 depict the surface generated via interpolation of helicopter magnetometry data acquired from the Pueblo and Kirtland ESTCP WAAPP sites, respectively. As can be seen, data at the Kirtland site were acquired in two separate continuous regions, a northern section and a smaller southern section.

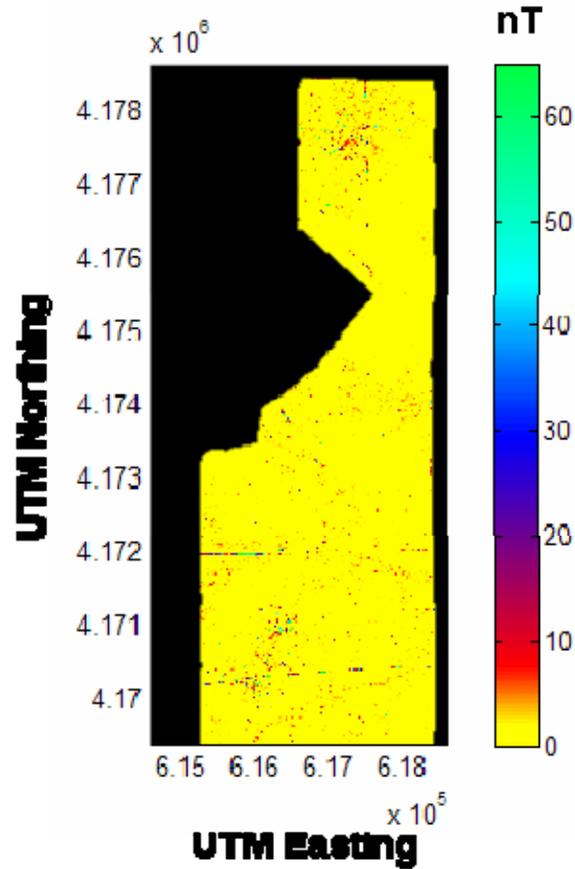
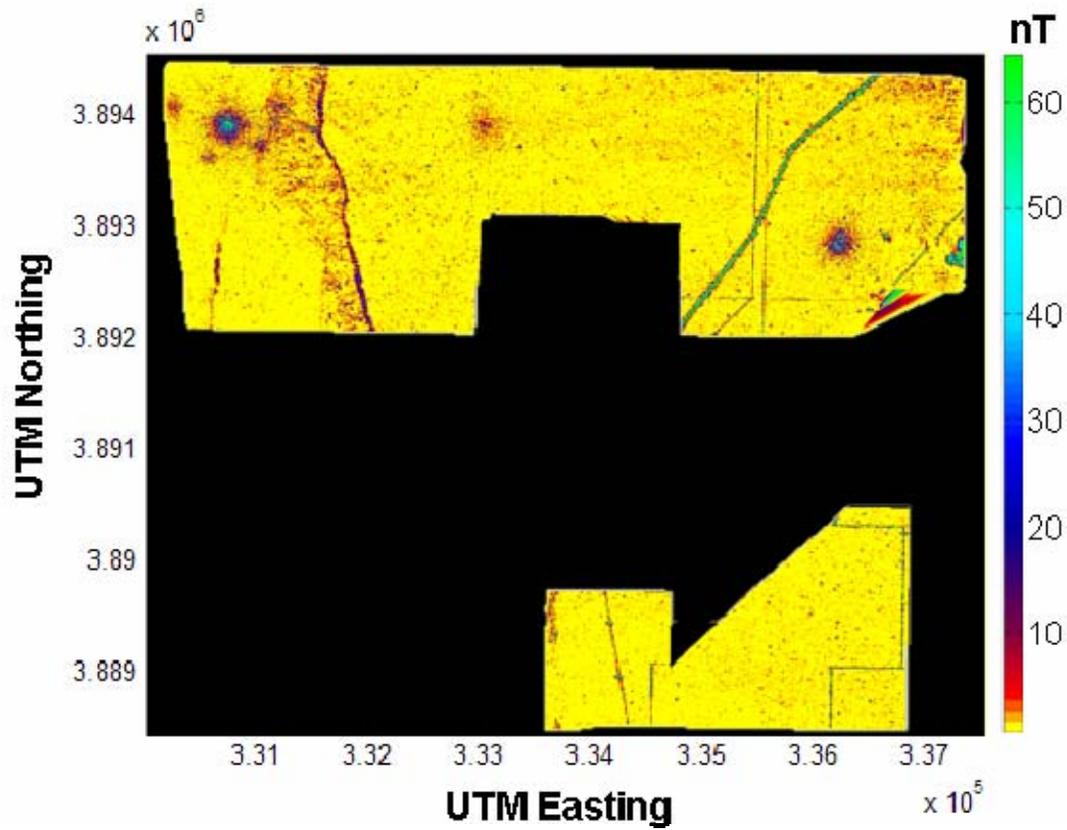
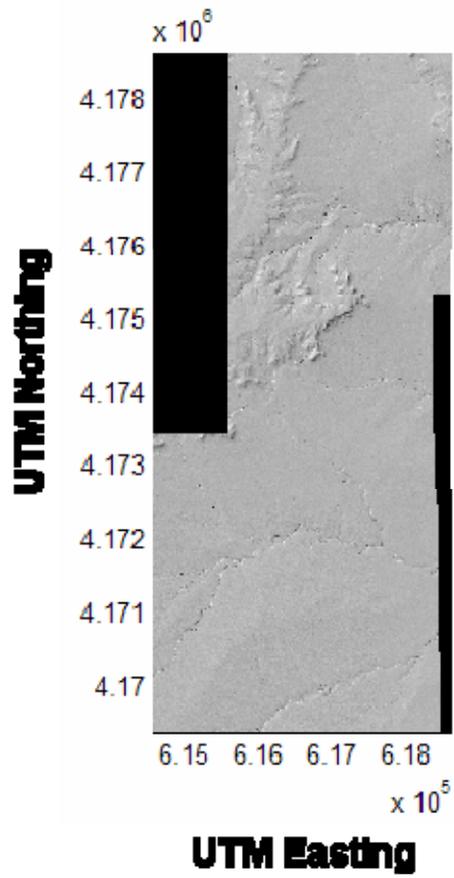


Figure 3. Helicopter magnetometry survey from the Pueblo ESTCP WAAPP survey site.

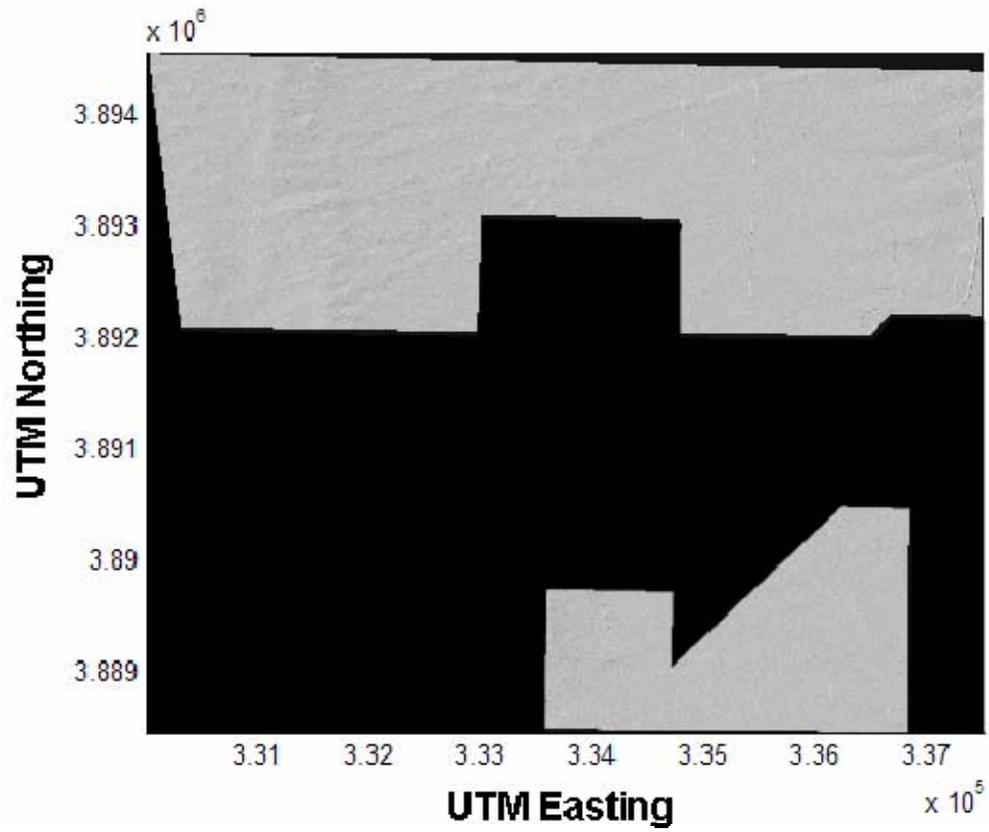


**Figure 4.** Helicopter magnetometry survey from the Kirtland ESTCP WAAPP survey site.

Figures 5 and 6 depict the surface generated from the corresponding aerial LiDAR surveys while Figures 7 and 8 depict the imagery from the aerial orthophotographic surveys, which were acquired at the same time as the LiDAR.



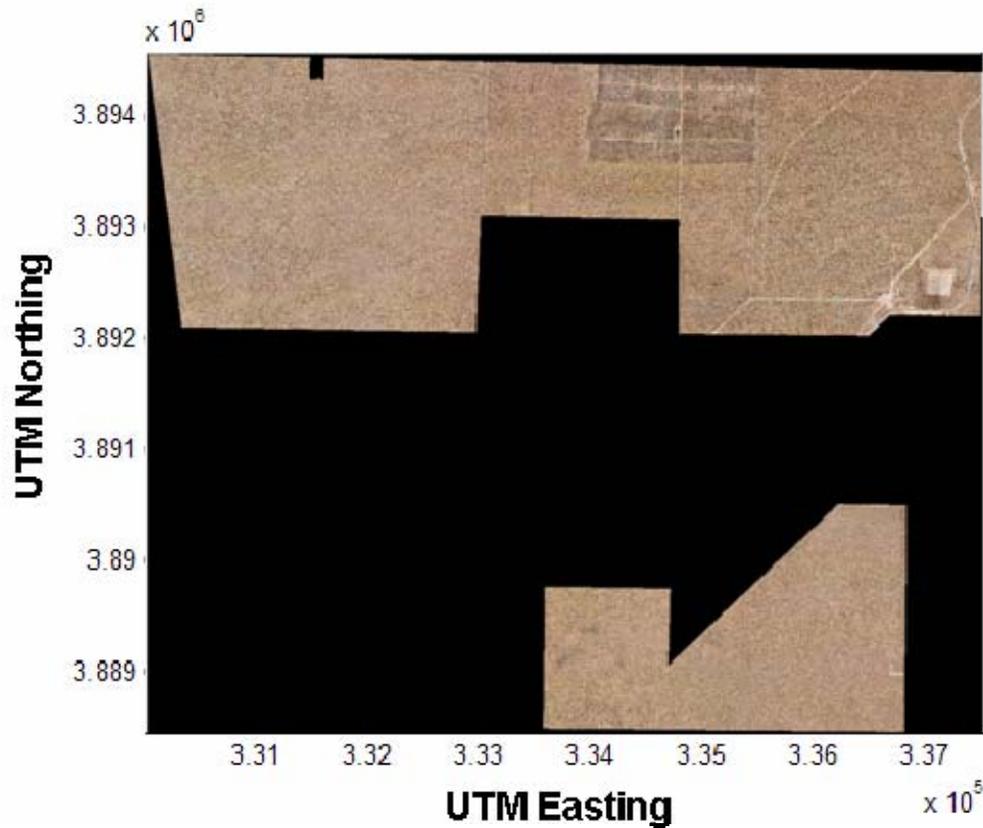
**Figure 5.** LiDAR survey from the Pueblo ESTCP WAAPP survey site.



**Figure 6.** LiDAR survey from the Kirtland ESTCP WAAPP survey site.

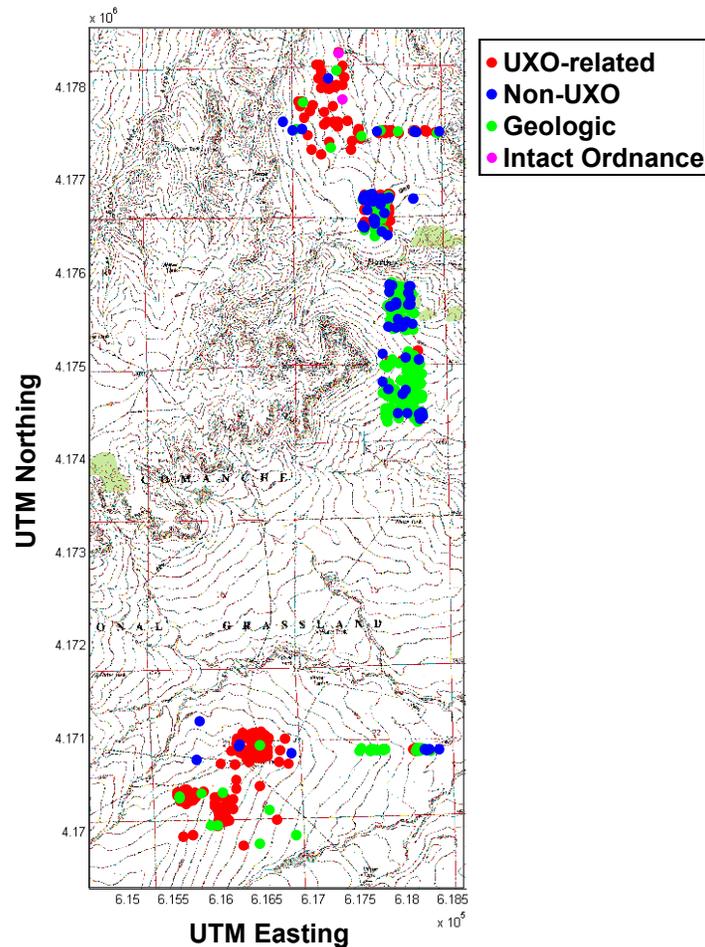


**Figure 7.** Aerial orthophotographic survey from the Pueblo ESTCP WAAPP survey site.



**Figure 8.** Aerial orthophotographic survey from the Kirtland ESTCP WAAPP survey site.

*Ground truth data.* In the summer of 2006, SERDP and ESTCP personnel conducted a ground-truth survey of the Pueblo site in order to provide data with which to validate WAAPP results. These data were obtained by Project MM-1510 researchers and are depicted in Figure 9.



**Figure 9.** Pueblo PBR #2 topographical map with ground truth data. Red dots indicate locations of ordnance-related, blue dots indicate non-UXO scrap, green dots indicate geologic feature (i.e., an empty dig site), and magenta dots indicate locations of intact ordnance.

## Feature Extraction and Feature Layer Generation

*Manually delineated features.* Various manually delineated regions of interest were provided for both the Pueblo and Kirtland sites. Derived from LiDAR and orthophotographic data, these were generated by ESTCP WAAPP performers, or aggregated by those performers from other third-party sources. The individual regions were provided as ESRI “shape files” containing UTM geocoordinates for the boundaries described by the region and, typically, a manually entered text label describing the nature of the region. These labels included both highly specific terms and rather vague characterizations. In order to generate appropriate feature layers, regions were grouped according to their relationship to UXO. Thus, all regions describing visible bombing targets were combined into one feature layer, all regions characterized as a “potential MRA” (munitions remediation area) were grouped into a second layer, and all those describing features that could potentially lead to significant non-UXO related

magnetometry signal (e.g., “structure”, “fence line”, “road”) were grouped into a third layer. Other regions with more vague labels (e.g. “linear feature”), more uncertain relationships to either the presence or absence of UXO (e.g., “bumps”, “berms”, and “depressions”) were grouped into other feature layers, but were not used for subsequent data fusion in this work. Figures 10 and 11 depict feature layers generated for the Kirtland site.

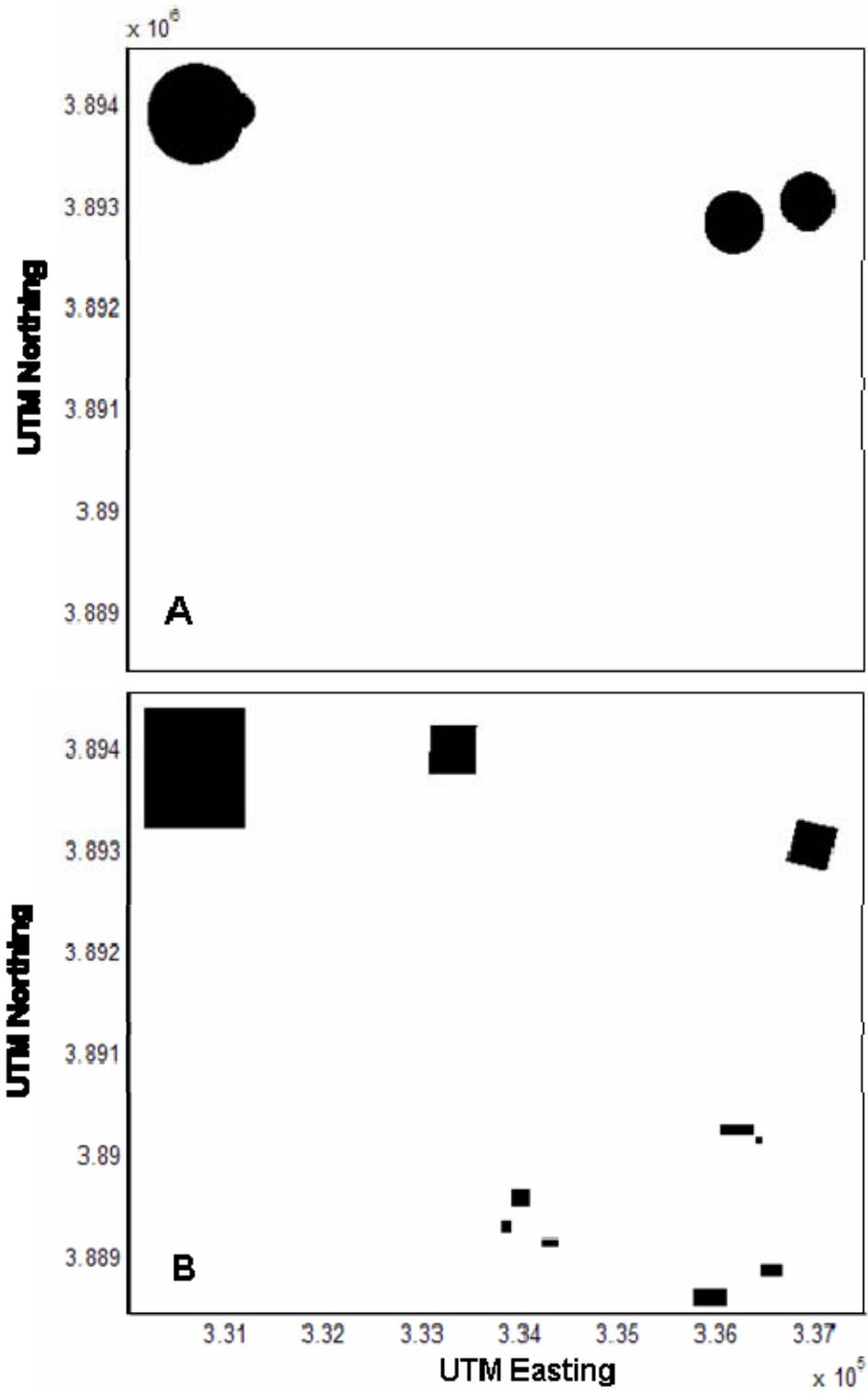
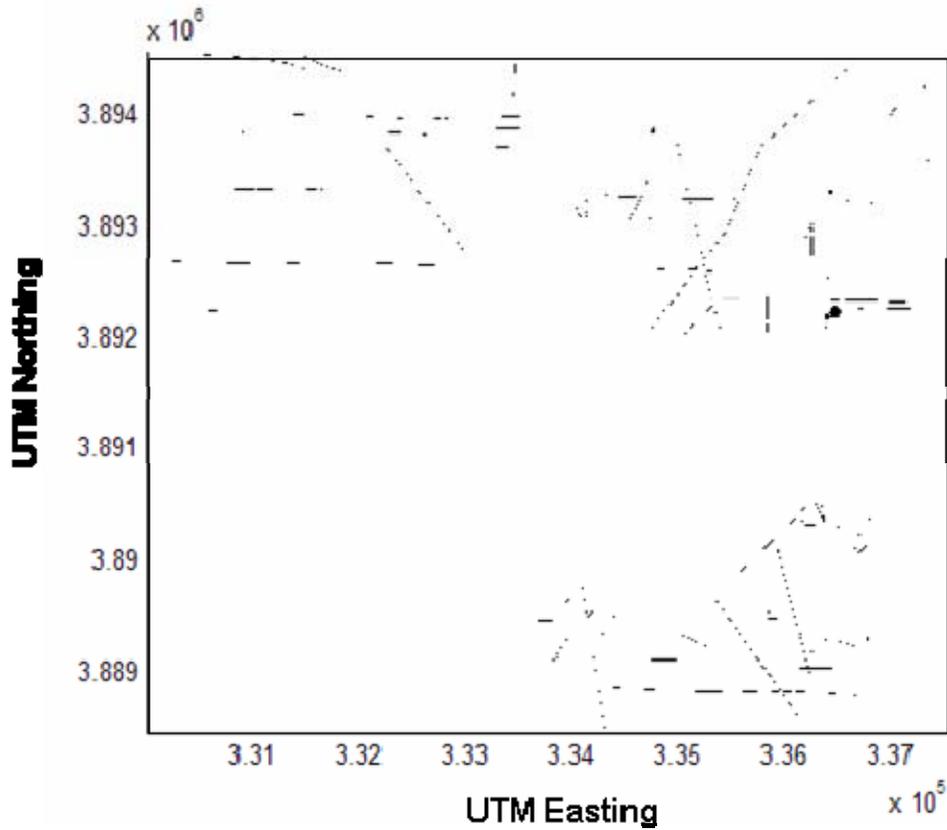
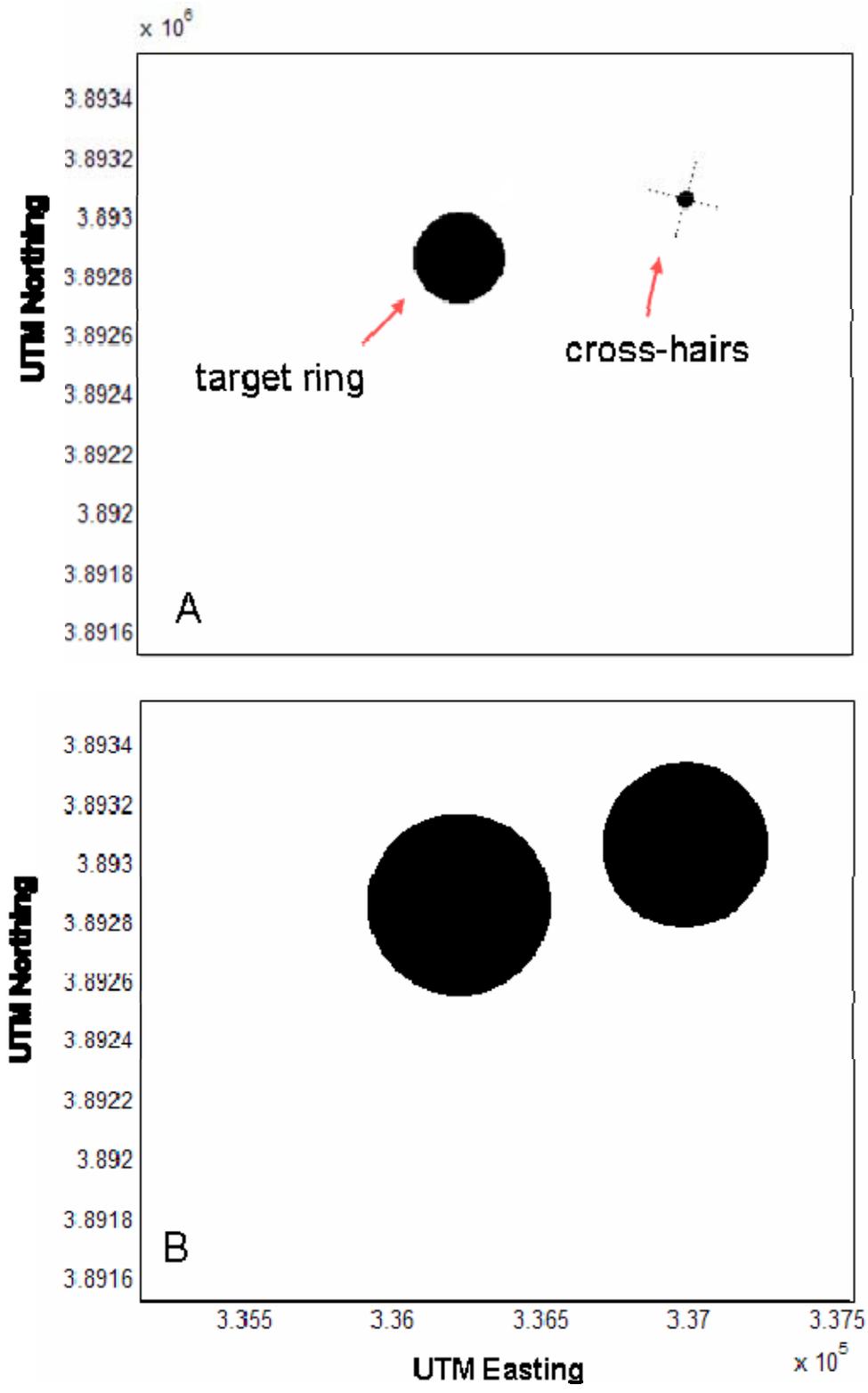


Figure 10. Feature layers resulting from (A), manually identified ship targets dilated to a larger effective area, and (B), manually identified munitions areas.



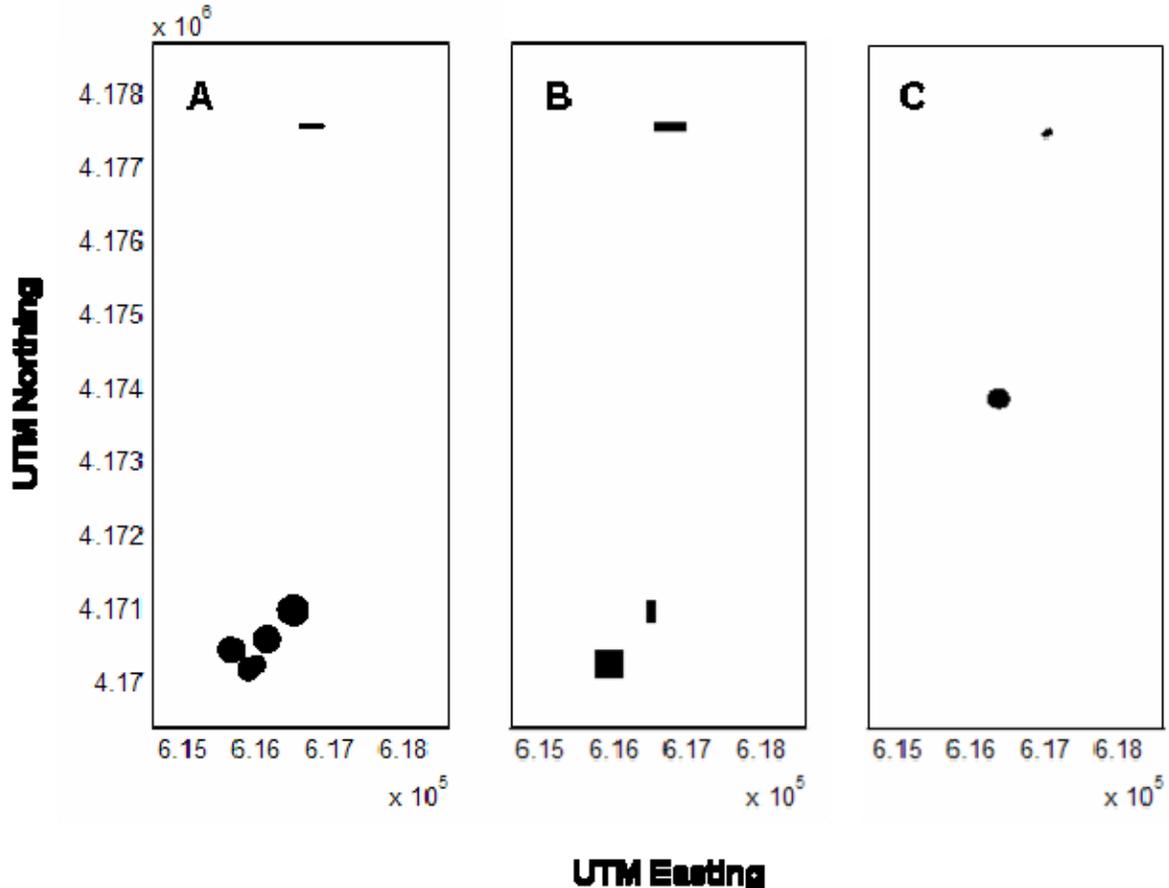
**Figure 11.** Feature layers resulting from manually identified roads, fences, pipelines, and man-made structures within the Kirtland site.

An enlarged region in Figure 12 demonstrates the process of region dilation to account for a given feature's area of influence. Visible bombing targets were expanded to include a circle of a diameter twice the length of the major axis of, and centered on, the original target. Dilation is necessary for any manually identified region of interest in which the area of influence of that feature exceeds the boundaries of the region itself.



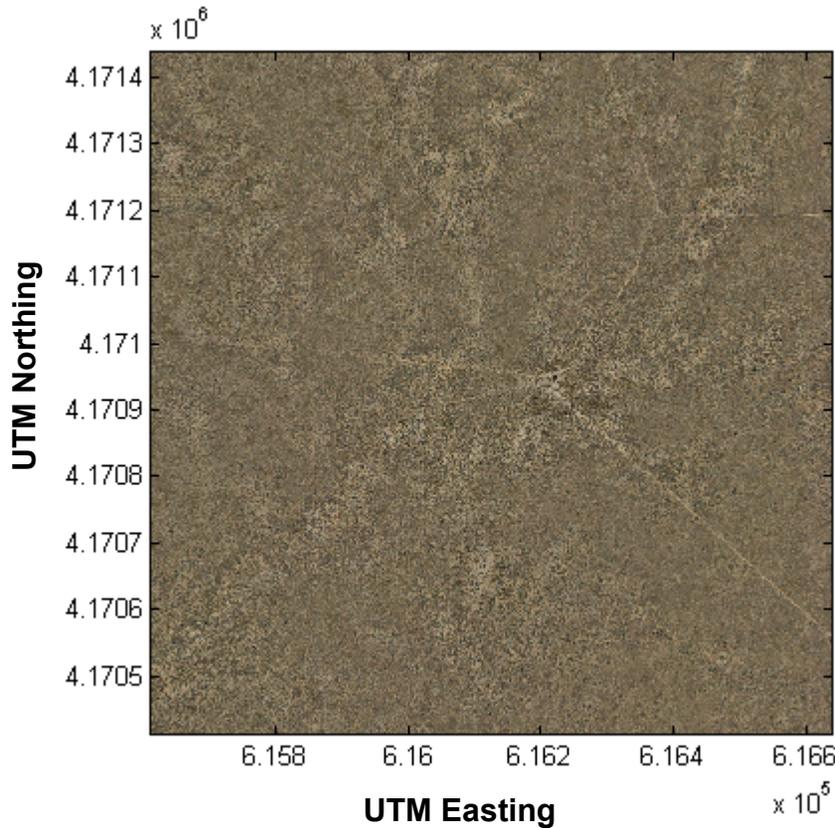
**Figure 12.** Enlarged region of the Kirtland feature layer containing bombing targets (A) before, and (B), after dilation.

A similar procedure utilizing manually delineated regions of interest at the Pueblo site resulted in three feature layers: manually identified ship targets, manually identified munitions areas, and manually identified man-made structures. The layers are depicted in Figure 13.



**Figure 13.** Feature layers resulting from (A), manually identified ship targets, and (B), manually identified munitions areas, and (C), manually identified man-made structures in the Pueblo site.

*Aerial orthophotography features.* Of the three survey data sets, the orthographic data were the least informative source for the purposes of automated feature extraction. An orthophotographic image at one-meter resolution of a sub region of the Pueblo survey site is depicted in Figure 14. The main features depicted by the orthophotographic survey of the site were different types of ground cover, evidenced by color and shape morphology. Subjects such as dirt roads, trees, and brush were plainly visible.

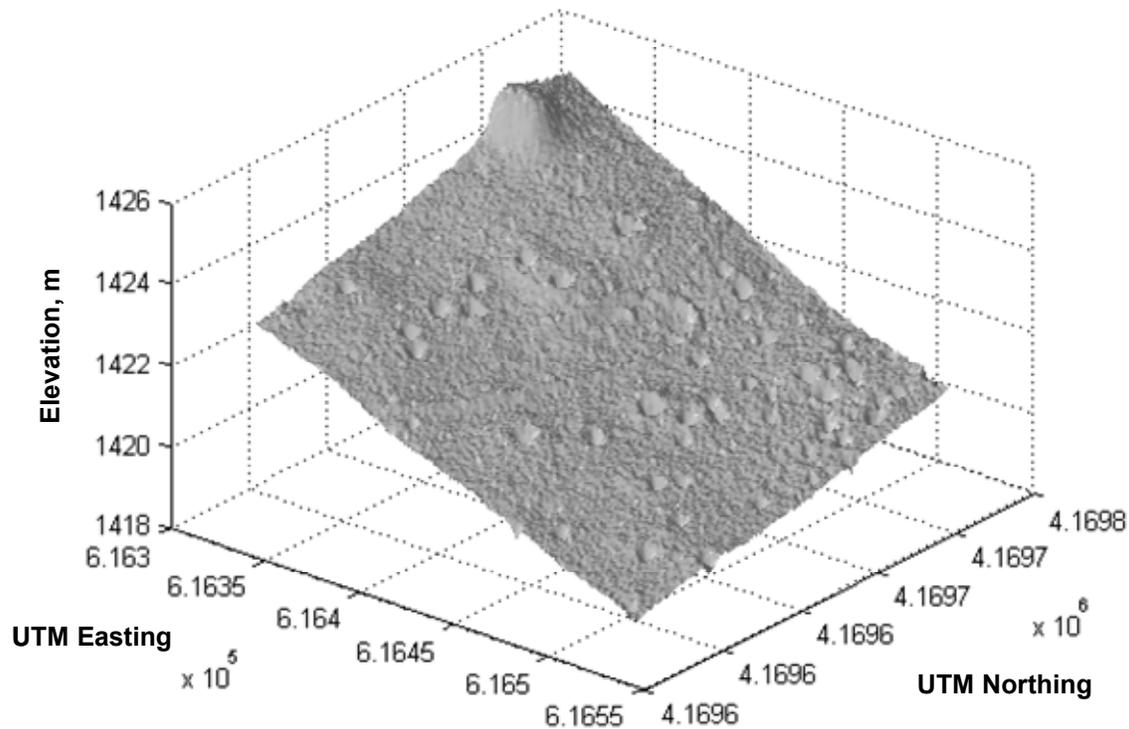


**Figure 14.** Orthophoto of a sub region of the Pueblo PBR #2.

However, none of these features were strongly associated with the presence or absence of UXO objects. In areas with heavy cratering, it was observed that vegetation was present within the craters, but vegetation was also observed in areas without cratering, negating this feature's usefulness for UXO indication. While the location of man-made structures such as fences and buildings was useful in assisting in the interpretation of magnetometry data, an automatic means of doing so was frustrated by the highly variable background present in the orthophotography images. Searching for regular geometric shapes (circular targets, fenced areas, buildings and structures) was difficult, as edges were generally very poorly defined, due primarily to the relatively low contrast of the image and the variability of the background. While various algorithmic approaches have been described for automatic identification of features from aerial imagery, none were found to be relevant to the UXO wide-area assessment problem. The utility of the orthophotographic data stream thus remains in manual assessment of the site by experts, and in verification of features detected within other data sources.

*LiDAR data features.* The principal feature in LiDAR data associated with UXO is cratering of the ground surface. While unexploded ordnance itself does not cause crater formation, craters indicate a land usage pattern of active ordnance bombardment, and through the assumption of collocation, a commensurate increase in the likelihood that UXO is present. One challenge to automatic crater is that the observed depth of ordnance-related craters is small relative to the overall changes in elevation experienced

within the survey area, as shown in the shaded relief map in Figure 15. Thus, the detection of craters must occur through the location of crater-specific arrangements of pixel values within the LiDAR image rather than directly through pixel values themselves.

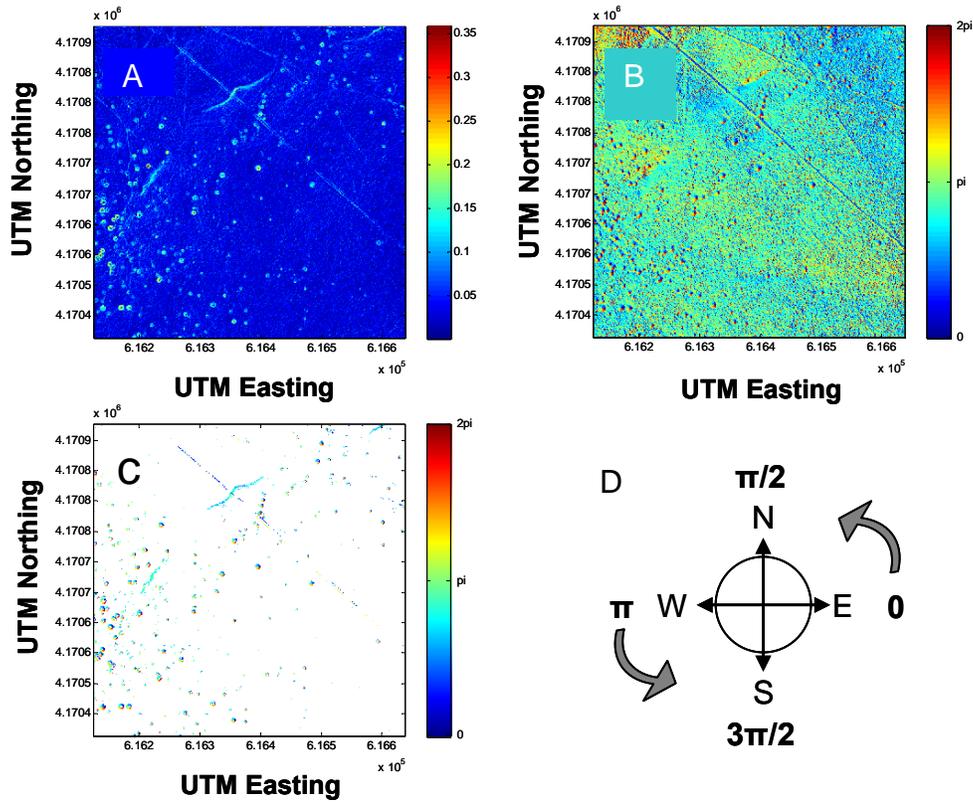


**Figure 15.** Subsection of LiDAR generated surface estimate for the Pueblo site showing ordnance-related cratering. Typical crater depths are smaller than variations in surface elevation.

A popular algorithm used in machine vision for locating objects of radial symmetry is the circular Hough transform. [19] The output of the transform is a map of the same dimensions as the input image, with each pixel value in the map indicating the degree to which it is surrounded by a constrained circular geometry. Map pixels with a high value indicate areas at the center of a circular feature, and an appropriate threshold can be applied to classify these features. However, implementation of standard circular Hough transform on the LiDAR survey data proved to be relatively nonselective for craters, hampering the utility of the circular algorithm output as an input into a data fusion.

To improve the selectivity, a more restrictive algorithm was implemented within MATLAB. First, the gradient and gradient direction for each pixel in the raw LiDAR surface were estimated, shown in Figure 16(A) and (B). This step was carried out in a manner similar to that utilized by Bai, Shen and Wang in their radial symmetry feature

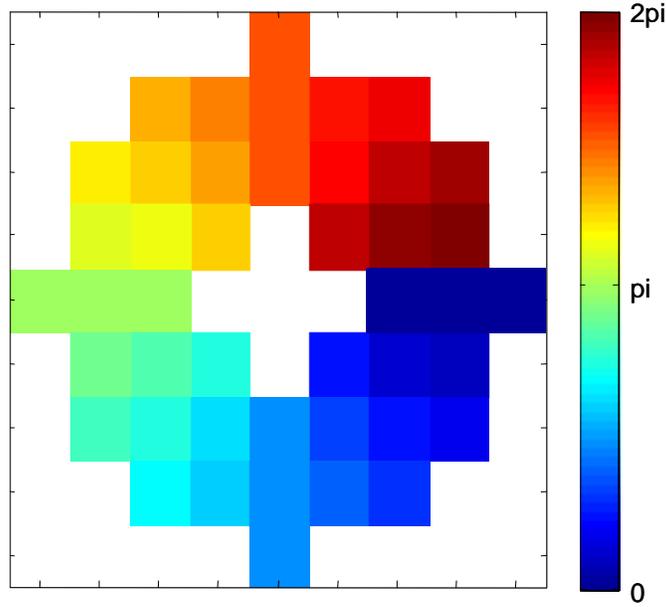
finding algorithm. [20] Pixels with gradient magnitudes less than 0.1 units were removed from consideration, resulting in the map shown in Figure 16(C). One consequence of this algorithmic approach is that shallower, less steeply sloped craters are more difficult to detect. The ability to detect shallower craters depends on the surface texture of the survey site and the uncertainty in the LiDAR estimated surface. By specifying a particular crater radius range and surface gradient threshold, the crater detection algorithm can be utilized to selectively search for specific crater sizes. Expert information regarding soil density and depth as well as historical information regarding which types of ordnance were used at the site could conceivably be used to fine tune these parameters.



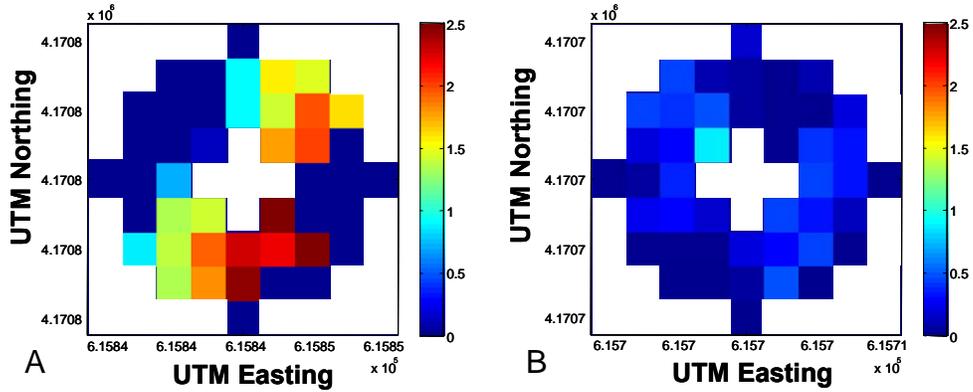
**Figure 16.** Estimated LiDAR surface gradient (A) and direction (B) before and after (C) gradient magnitude threshold application. (D) A direction of 0 radians indicates a positive surface gradient estimate orientation of due east.

Next, a windowed filter was applied across the gradient direction image that calculated the difference between the region in the local area of the center pixel of the filter and an “ideal crater” gradient direction template. The crater template, shown in Figure 17, was generated by assuming a circular crater of radius four meters and calculating the direction of the surface gradient at each pixel in the hypothetical crater. Pixels that were in the center of the crater and beyond the rim were disregarded in this difference calculation. Pixels with gradients that were within 10 degrees of their counterpart in the crater template were identified, with each one counting as one vote. The filter generated the number of votes each potential crater location garnered, normalized by the total number

of potential votes. Figure 18 depicts the difference in direction between the crater template and (A), a non-crater region, and (B), a crater-containing region.

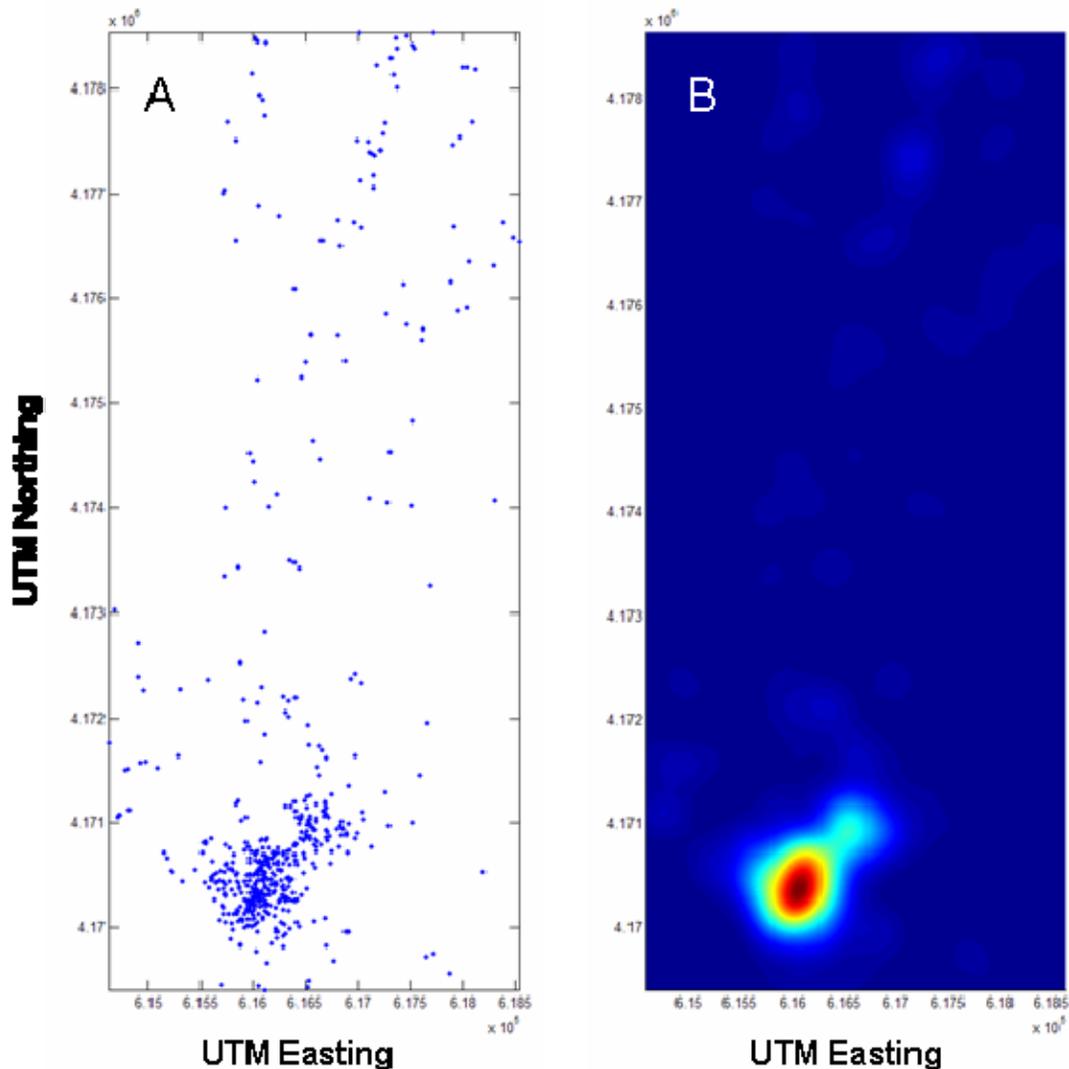


**Figure 17.** Surface gradient direction template for a crater of four meter radius in a one meter resolution image.



**Figure 18.** Residual between data and ideal crater template for (A) a non-crater containing region, and (B) a crater containing region.

Pixels with a value of less than 0.30 (i.e., only thirty percent of surrounding pixels pointing inward to center) were discarded and the remaining pixels were flagged as crater center locations. This method greatly enhanced rejection of geographic features within the survey area when compared to the circular Hough algorithm. Automatic crater detection algorithm output for the Pueblo site is shown in Figure 19(A).



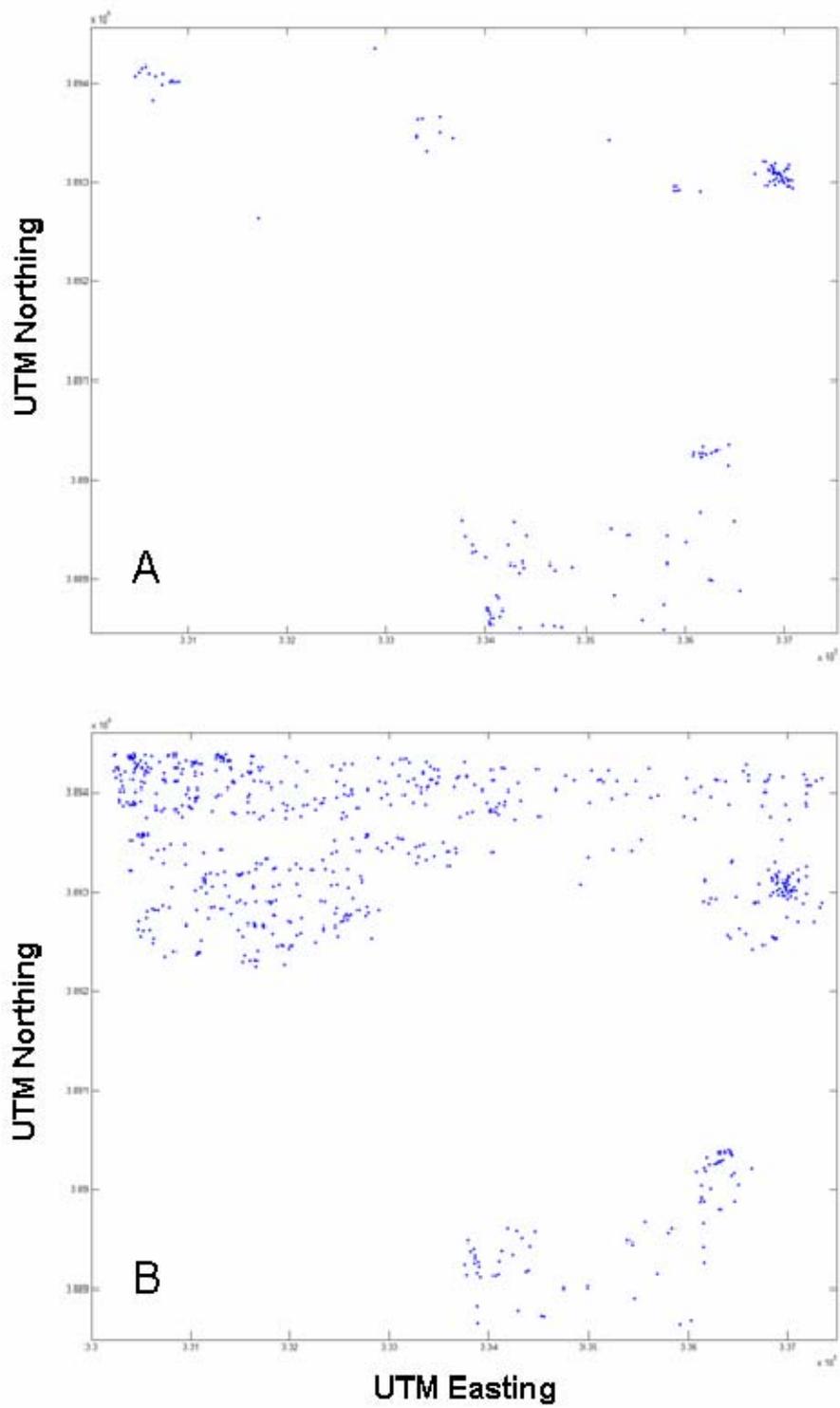
**Figure 19.** Crater detection at the Pueblo site. (A) depicts automatically identified craters via an NRL crater-detection algorithm. (B) depicts a crater density estimate determined via Kernel Density Estimation. Colors are scaled from blue (=0) to red (=1).

Additional data quality metrics such as crater depth and edge slope statistics can also be easily calculated for data fusion purposes. These metrics may become important in situations where distinguishing between different types of ordnance usage and/or ordnance and non-ordnance sources of cratering are desired. The final step in the crater detection algorithm was to generate a feature layer from the output set of crater locations. A standard KDE algorithm was employed to estimate the feature density of identified craters feature found throughout the Pueblo survey area, shown in Figure 19(B). The feature layer was then suitable for input to data fusion algorithms.

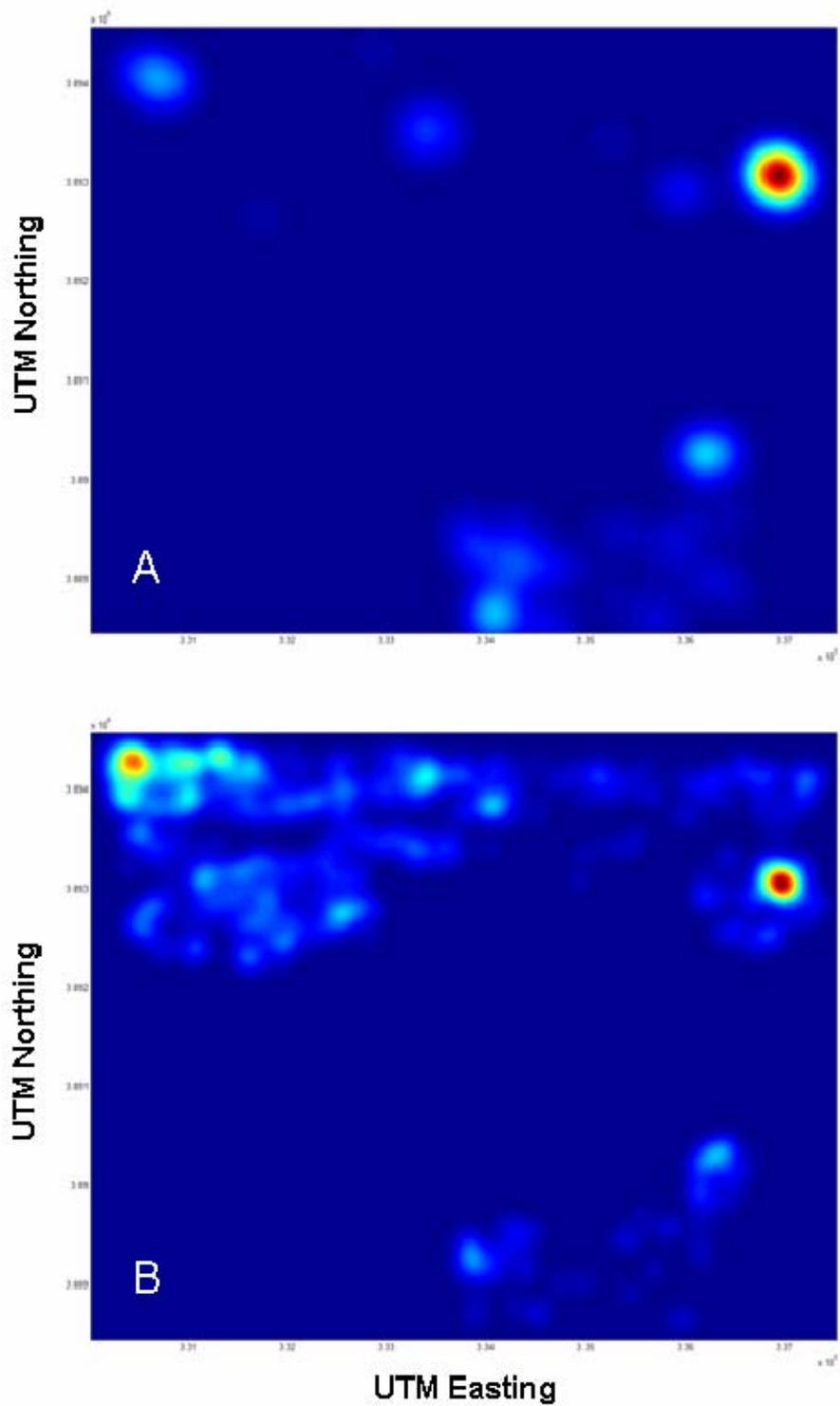
The automatic crater detection algorithm performance was compared against manual detection of craters, as shown in Figure 20. Manually identified crater locations were provided for the Kirtland site as part of the data acquired through SERDP. Figure 21

depicts feature intensity maps generated from these two feature sets. Examination of Figures 19-21 demonstrates the relative performance between the Pueblo and Kirtland sites.

In particular, the comparison of site to site performance of the automatic crater detection algorithm indicated that thresholds for maximum detection rates were similar at each site, although the surface topography presented by the Kirtland site made accurate crater detection more challenging as it presented greater background interference. Thus, the chief difference between the results attained at the two sites was that the Kirtland site demonstrated higher apparent false positive rates, and, unlike Pueblo, these false positive rates did not correlate well with algorithm threshold setting. However, the concept of “false positive” used here is a misnomer, as ground truth data for crater locations were not available for either site, preventing accurate assessment and comparisons of algorithmic performance. At both sites, the addition of a wavelet-based filter enhanced automatic crater detection. Although in the absence of truth data, the improvement and false positive rates were difficult to quantify.



**Figure 20.** Craters detection at the Kirtland site. (A) depicts manually identified craters (Versar, Inc.) (B) depicts automatically identified craters via an NRL crater-detection algorithm.

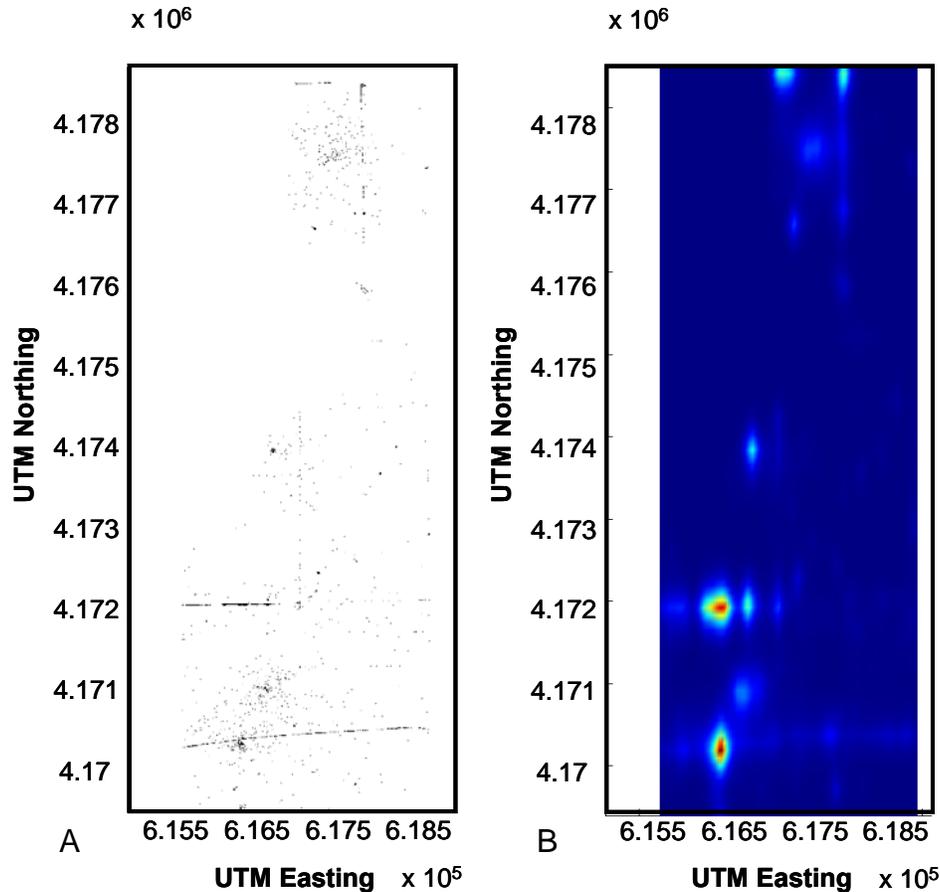


**Figure 21.** Crater density estimate determined via Kernel Density Estimation. (A) depicts manually identified craters (Versar, Inc.). (B) depicts automatically identified craters via an NRL crater-detection algorithm. Colors are scaled from blue (=0) to red (=1).

*Helicopter magnetometry feature extraction.* The phenomena measured via airborne magnetometry sensing are deviations in the background magnetic field of the earth due to the presence of ferromagnetic materials on or below the surface of the ground. Ferrous materials that may be present at a survey site include not only UXO and ordnance-related scrap but also a background of non-UXO related signal comprised of man-made features such as buildings, fences, pipelines, general metallic detritus, and magnetically active geological features. Thus, the main criterion for a successful detection algorithm for wide-area assessment of ordnance-related material is the accurate separation and extraction of UXO-related signal from background with computationally-reasonable methods.

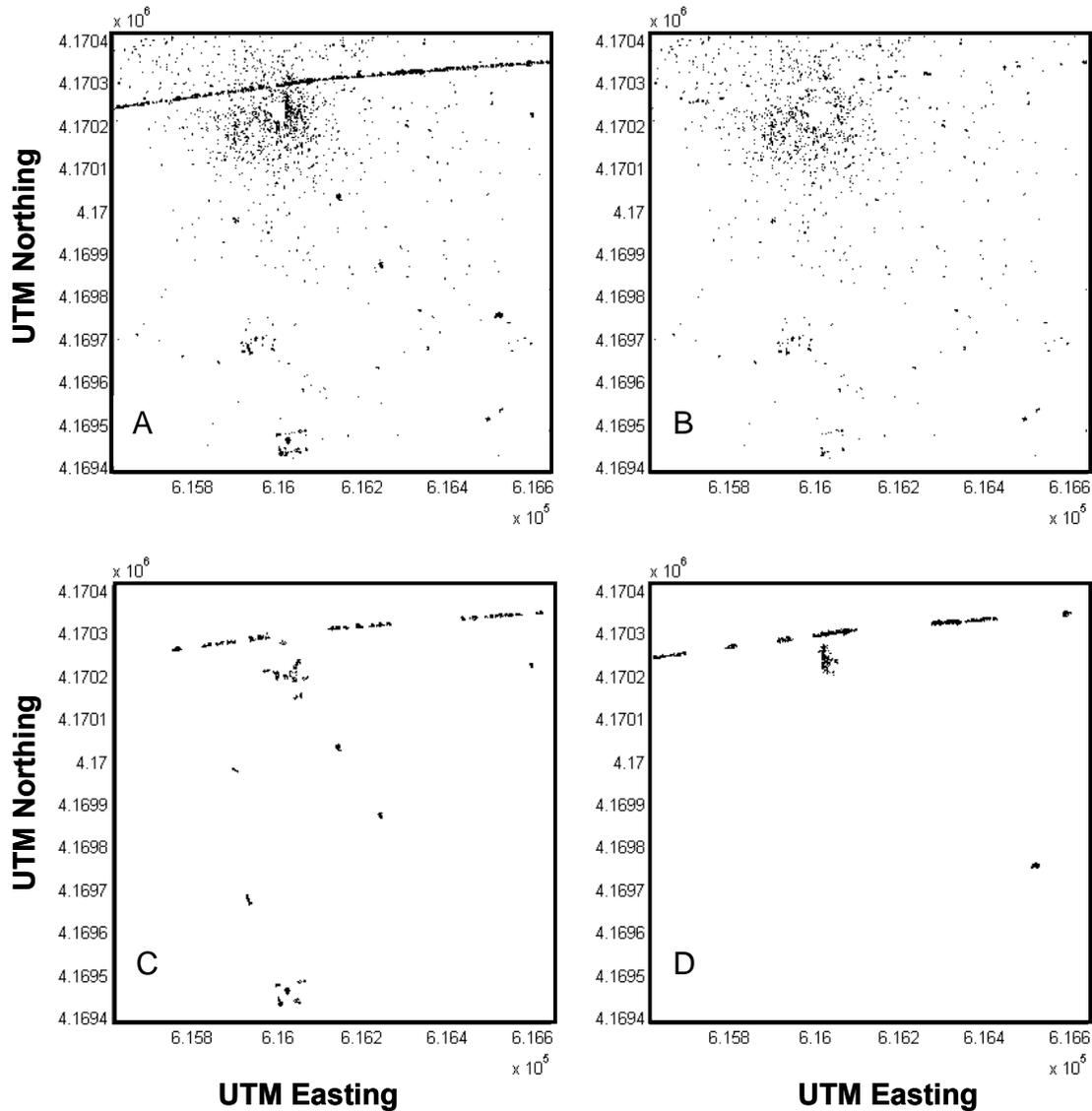
Typically, analysis of magnetometer readings involves fitting a physics-based model to predict quantities such as object depth, size, and orientation. This approach can be complicated by highly overlapped signals from multiple objects and from excessive distance between the object and the sensor. For the purposes of wide-area assessment, it was decided that, rather than approaching the UXO recognition from a physics-based modeling approach, it would be advantageous to consider the problem from an image analysis standpoint, examining the morphological characteristics of total magnetic signal distributions throughout a given survey site. A morphological approach has the potential to loosen requirements on magnetometer data gathering and to speed analysis times while continuing to provide useful assessments of potential UXO contamination for input to a data fusion framework.

The principle algorithmic challenge to feature detection within the magnetometry data was the similarity of the ordnance-related signal to the non-ordnance background, both in terms of feature intensity and spatial location. In order to overcome this, a pixel clustering technique was developed that focused on extracting features relevant for ordnance-related signal in the context of density, intensity, and morphology aspects to the data. Magnetometry data were thresholded to ten units to remove spurious and small-valued background. An algorithm was developed to identify clusters of adjacent, above-threshold pixels and group them into “pixel islands.” Adjacent pixels were defined as a pixel’s eight nearest neighbors. An island was defined as any cluster of two or more adjacent and above-threshold pixels. Pixels meeting these criteria are shown in Figure 22(A) for the Pueblo site. As with the crater features, in order to generate a feature layer for input to data fusion, the pixel clusters determined to be UXO-related magnetic anomalies were used to estimate a site-wide anomaly density map. Figure 22(B) displays a preliminary density estimate obtained from the density of pixel islands in pane (A). Pixel islands are able to capture ordnance-related features in the magnetometry data, but significant ferromagnetic background from man-made structures (e.g., the clearly visible horizontal fence lines in the lower half of the survey region) is equally well-captured.



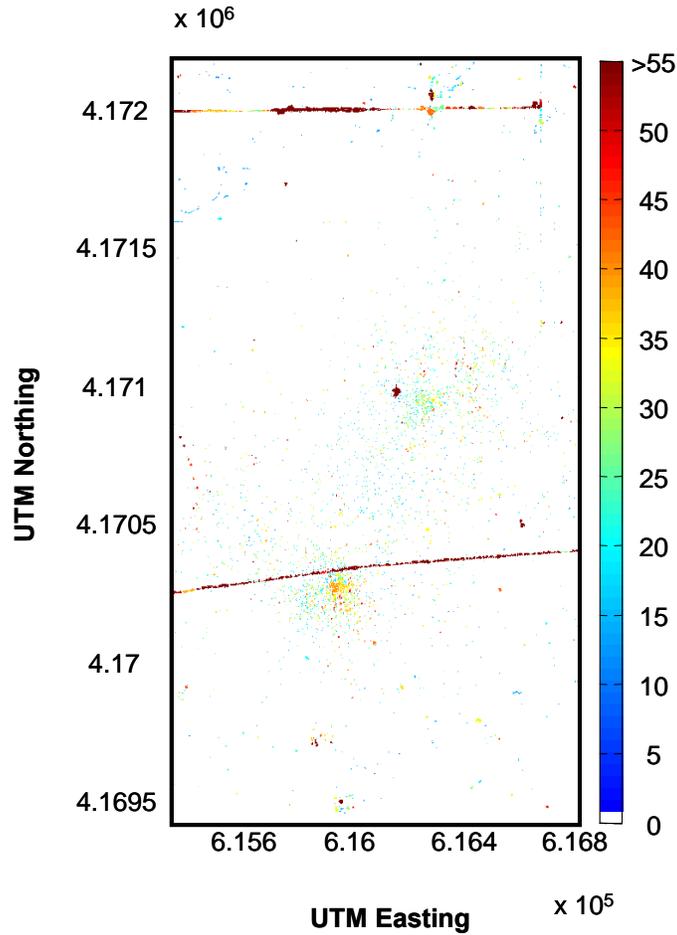
**Figure 22.** Shown in (A) are pixel islands in the magnetometry data. (B) is a map of estimated feature density generated from (A).

In order to distinguish ordnance-related features from structured ferromagnetic background, the morphology of the pixel islands was explored. Filtering pixel islands based on the number of pixels each island contained was effective in separating signal from background features. For example, Figure 23(A) displays all pixel islands in the southern portion of a sub region of the Pueblo containing both a grouping of UXO-related scrap and a fence line. Figure 23(B) displays only those pixel islands with two to 75 member pixels. Note that only a few islands from the man-made structures remain. The scatter clouds of ordnance-related signal associated with targets in the northern and southern sections are clearly the dominant features. Panes (C) and (D) display pixel islands with 75 – 200 and 200 – 1000 member pixels, respectively. In these ranges, nearly all pixel islands belong to man-made structures. Only a few islands in the densest regions of the ordnance-related scatter cloud are present in these ranges with the fence line crossing the southern section clearly visible.



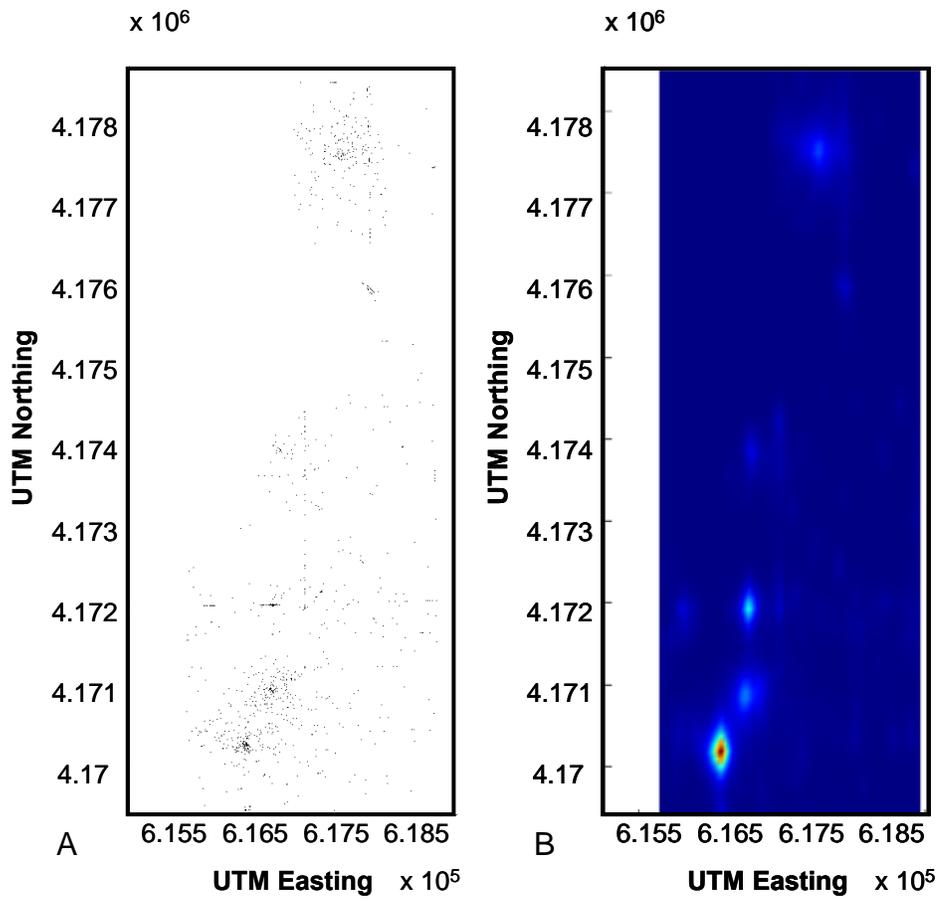
**Figure 23.** Filtering of pixel islands: (A) all pixel islands, (B) islands with 2 – 75, (C) 75 – 200, and (D) 200 – 1000 member pixels.

The pixel island features contain additional information about the magnetometry data and make possible further separation of ordnance-related signal from structured ferromagnetic background. First, the intensities of member pixels seemed to be larger for man-made structures than for ordnance. This could have been due to the location of structures on the ground surface, appreciably closer to helicopter than buried ordnance, as well as to more concentrated ferromagnetic content, typical of fence posts and pipelines. Figure 24 shows average pixel intensities for each pixel island in a sub region near the target in the southern region of the Pueblo site. Intensities in the fence feature were much higher than those in the dense cloud of UXO-related scrap around the bombing target. With proper site-specific calibration, an intensity-based approach may provide a probabilistic estimate to determine if features in the magnetometry data are on or below the surface of the ground.

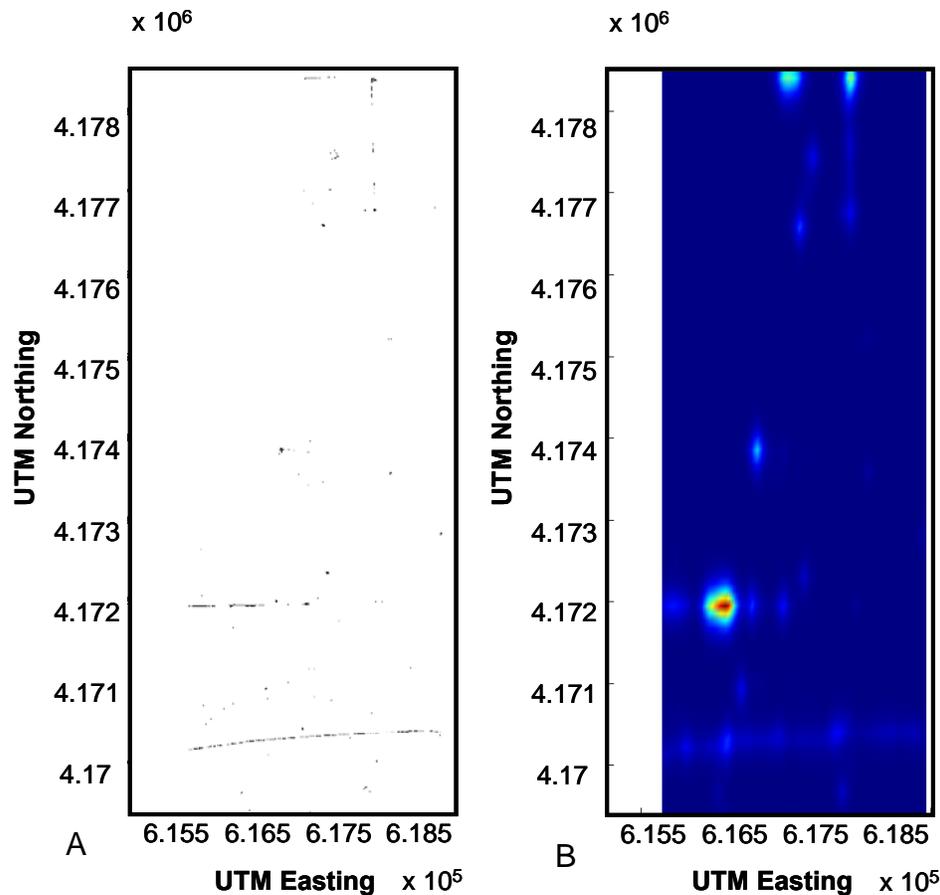


**Figure 24.** Average pixel island intensities, capped at 55 nT, for island pixels surrounding the target in the southern portion of the Pueblo site. Surface clutter objects such as fence lines have significantly greater intensity per pixel.

Figures 25 and 26 depict the distribution of pixel islands in the airborne magnetometer data with average signal intensities below and above a threshold of 50 nT, respectively. As with filtering based on pixel island size, good discrimination was observed between signal associated with buried UXO-related scrap and that associated with metallic objects on the surface.



**Figure 25.** Shown in (A) are magnetometer signal pixel islands with average intensity less than 50 nT. Shown in (B) is a map of estimated feature density generated from the pixel islands in (A).



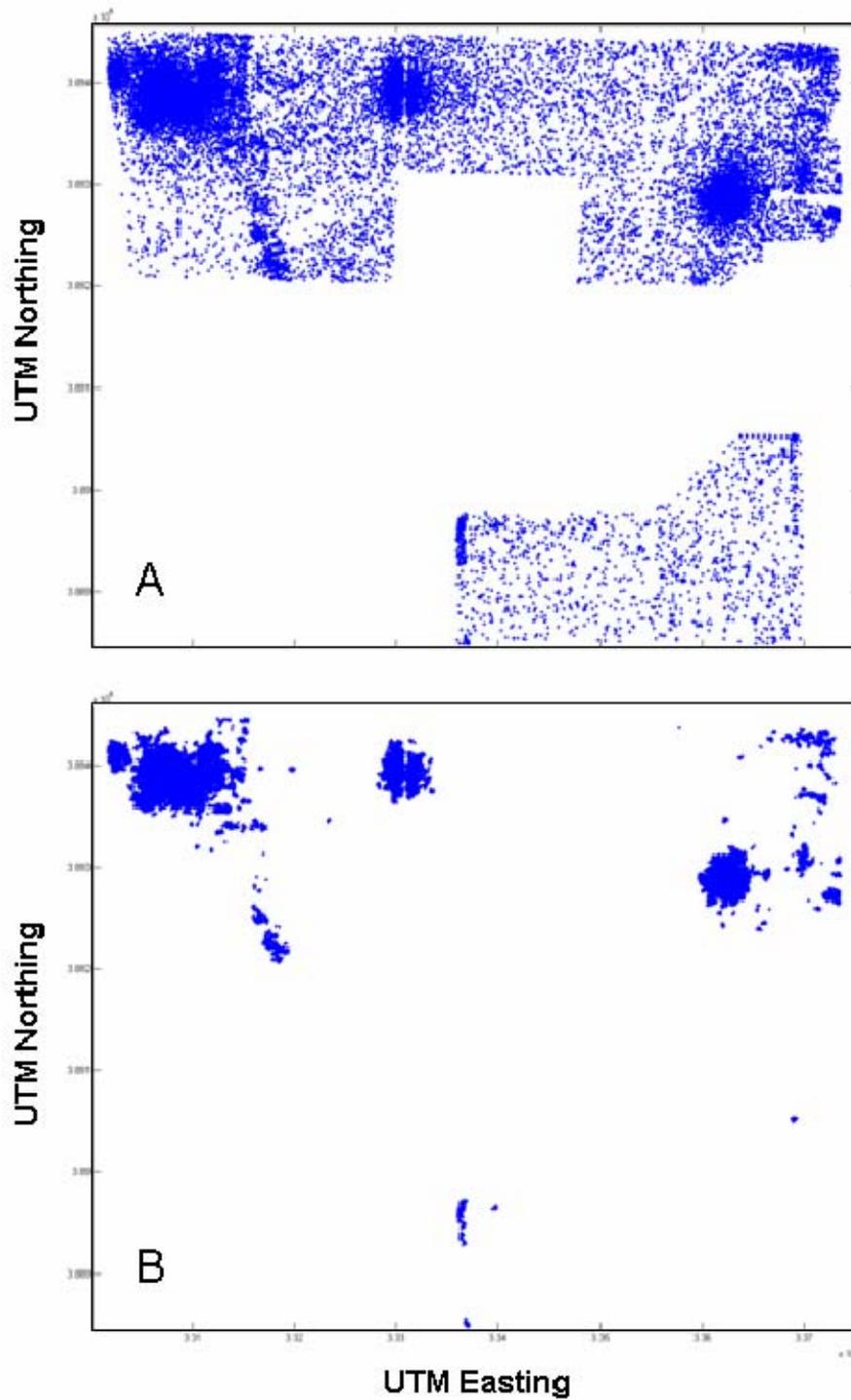
**Figure 26.** Shown in (A) are magnetometer signal pixel islands with average intensity more than 50 nT. Shown in (B) is a map of estimated feature density generated from the pixel islands in (A).

*Enhancement of magnetic anomalies.* In addition to the NRL-developed algorithms, magnetic anomalies were also algorithmically identified from helicopter magnetometry data by Sky Research, Inc., the ESTCP WAAPP performer that acquired helicopter magnetometry data at both the Pueblo and Kirtland sites. An identified anomaly indicated a survey grid location at which the acquired magnetometry signal was consistent with the presence of a UXO-related object. These anomaly locations were obtained from SERDP and were utilized to create corresponding feature layers consisting of maps of estimated anomaly densities for both the Pueblo and Kirtland sites. These maps were generated using the same techniques employed in generating crater density maps from detected crater locations described earlier.

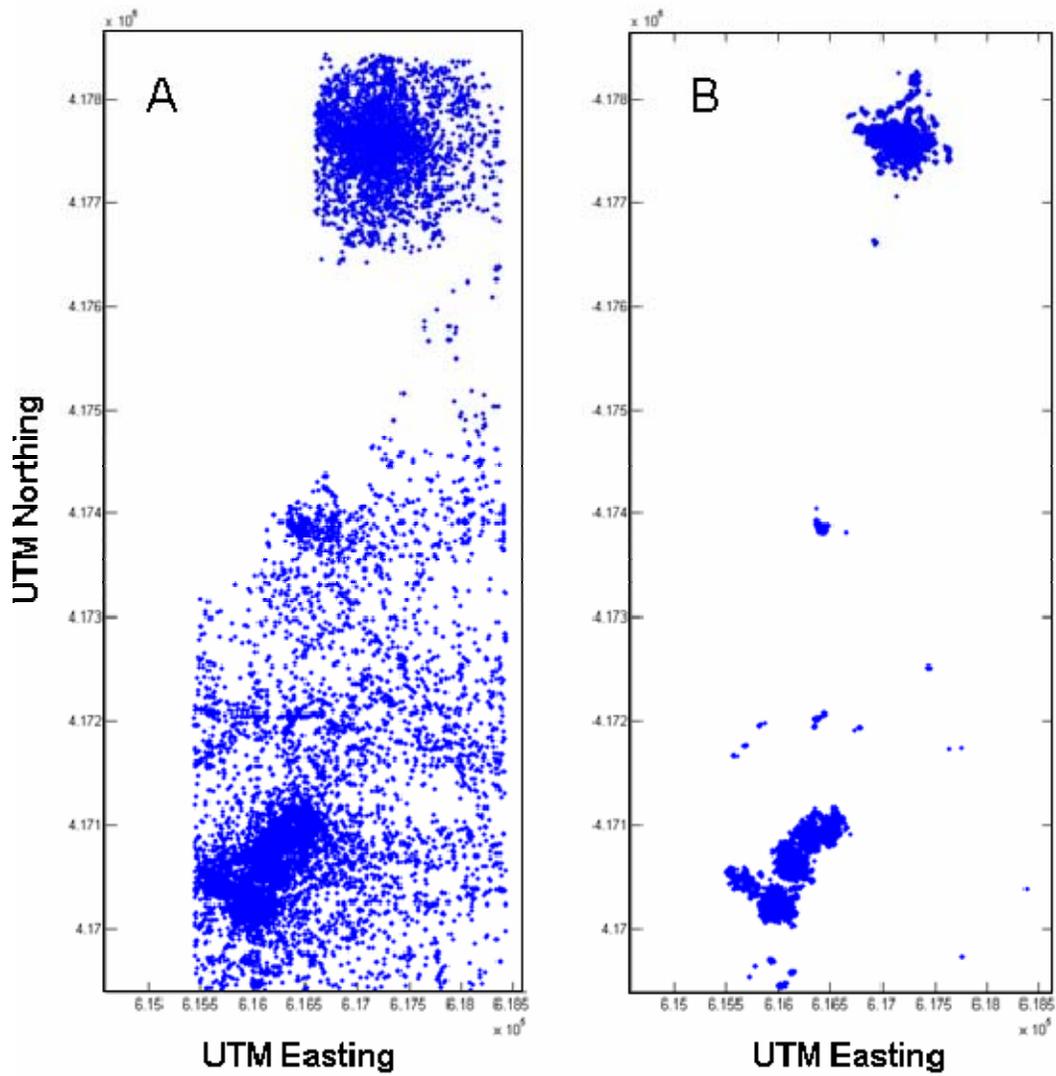
In identifying the anomalies, a Receiver Operating Characteristic (ROC) curve analysis was employed to select a threshold that resulted in a “false positive rate” of approximately 25% when compared to manually identified anomalies at the Pueblo and Kirtland sites. [21] The automatically identified anomalies from SERDP are shown in Figure 27(A) for Kirtland and Figure 28(A) for Pueblo. Both figures exhibit a nearly uniform spatial distribution of detected anomalies punctuated by clusters of anomalies

near known targets and gaps at known locations of roads and pipelines. A histogram of the minimum (nearest neighbor) distances between magnetic anomalies at the Kirtland site is shown in Figure 29. The data were distributed similar to a Poisson distribution with a mean and variance of 6.0 with additional data points present in the tail region (i.e., larger distances between anomalies.)

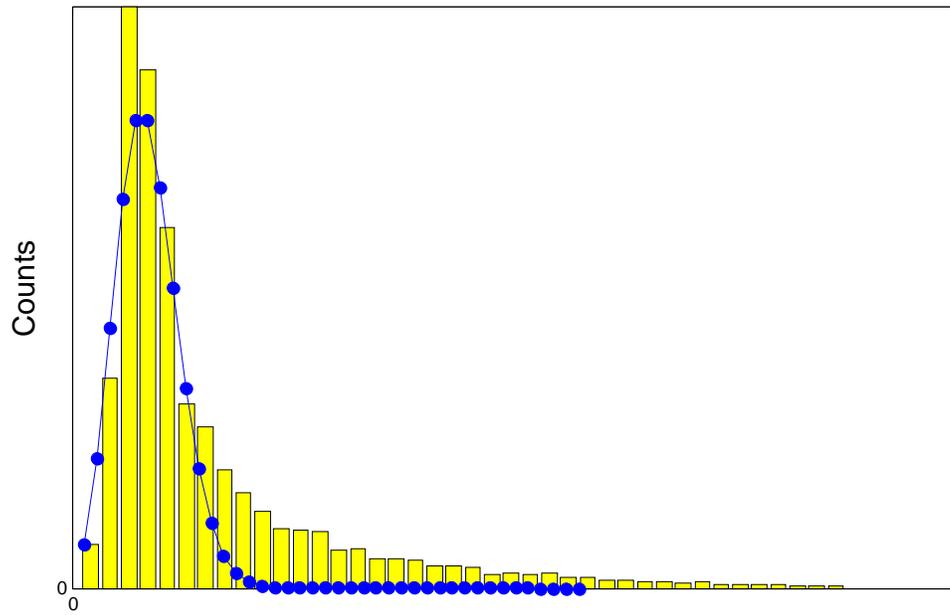
A nearest neighbor clustering algorithm was developed to distinguish anomalies associated with UXO from those due to background. The algorithm was built on the assumption that anomalies due to non-UXO phenomena were likely to be more uniformly distributed throughout the surveyed areas, due to their random origin, than UXO-related anomalies, which were more likely to be clustered around bombing targets or areas replete with magnetic rocks. Application of the automatic clustering algorithm to filter helimag anomaly feature sets demonstrated significant reduction in background signal, as shown in Figures 27 and 28, and in the associated feature density maps, shown in Figures 30 and 31. The algorithm removed anomalies with 10 or fewer neighboring anomalies within a circle of radius 64 meters. Anomalies passing these criteria were highly clustered around known target areas, though some clusters were known to be due to the presence of magnetic geology.



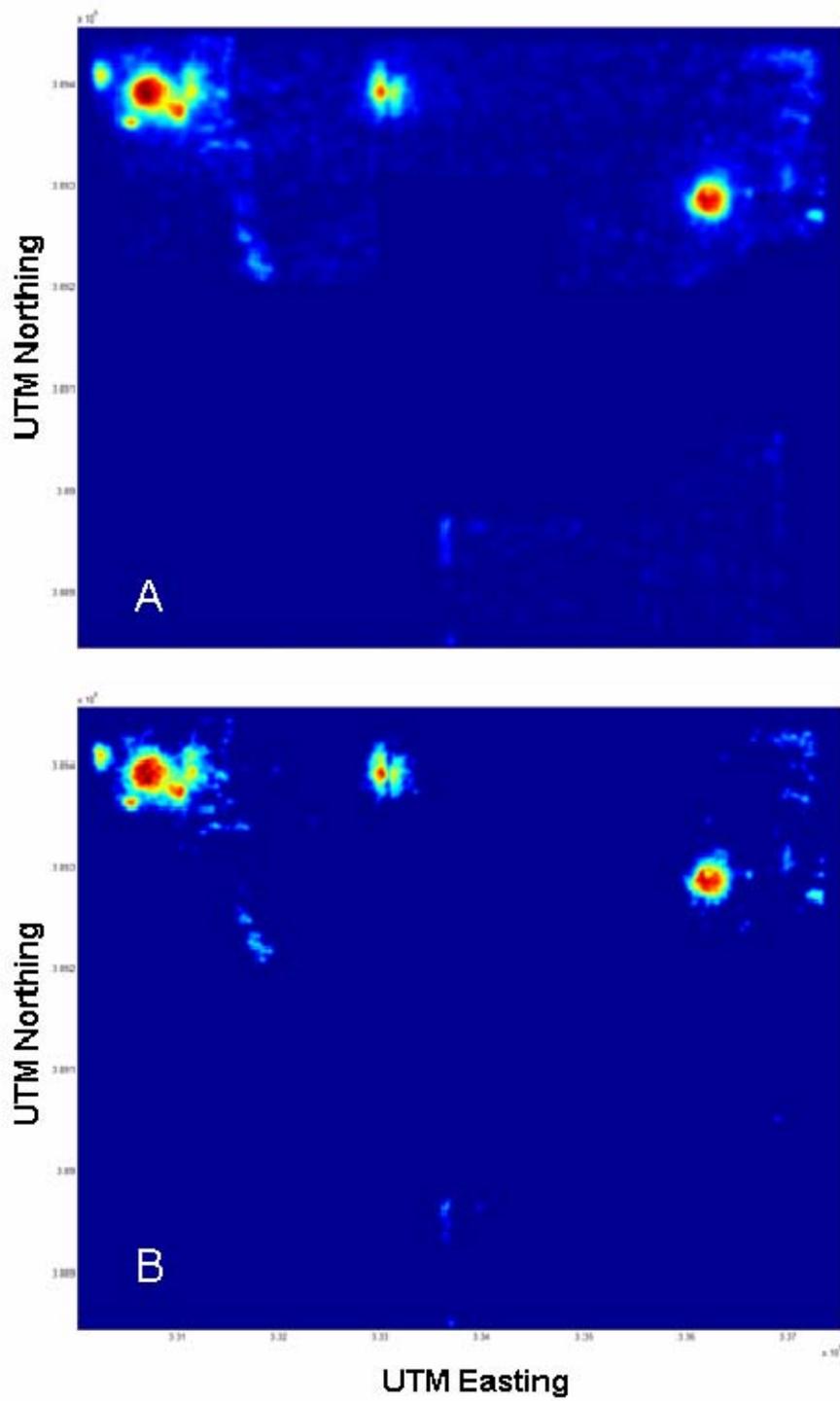
**Figure 27.** Magnetic anomalies at the Kirtland site. (A) depicts algorithmically identified anomaly locations (Sky Research). (B) depicts anomaly locations after filtering with an NRL clustering algorithm.



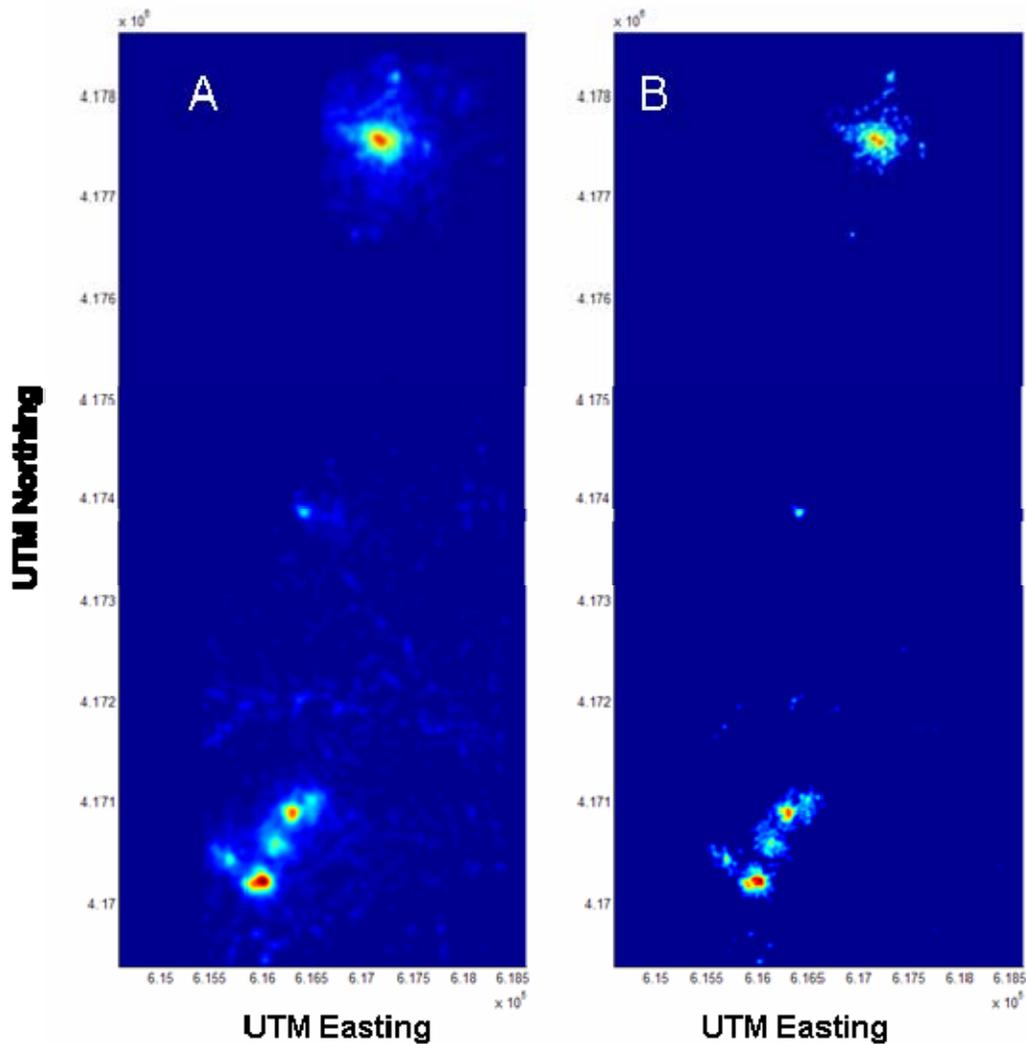
**Figure 28.** Magnetic anomalies at the Pueblo site. (A) depicts algorithmically identified anomaly locations (Sky Research). (B) depicts anomaly locations after filtering with an NRL clustering algorithm.



**Figure 29.** Histogram of nearest neighbor distances between magnetic anomalies at the Kirtland site (yellow bars), overlay of Poisson distribution of mean 6.0 (= variance) (blue line).



**Figure 30.** Magnetic anomalies at the Kirtland site. (A) depicts estimated anomaly density for Fig. 29(A). (B) depicts estimated anomaly density for Fig. 29(B).



**Figure 31.** Magnetic anomalies at the Pueblo site. (A) depicts estimated anomaly density for Fig. 30(A). (B) depicts estimated anomaly density for Fig. 30(B).

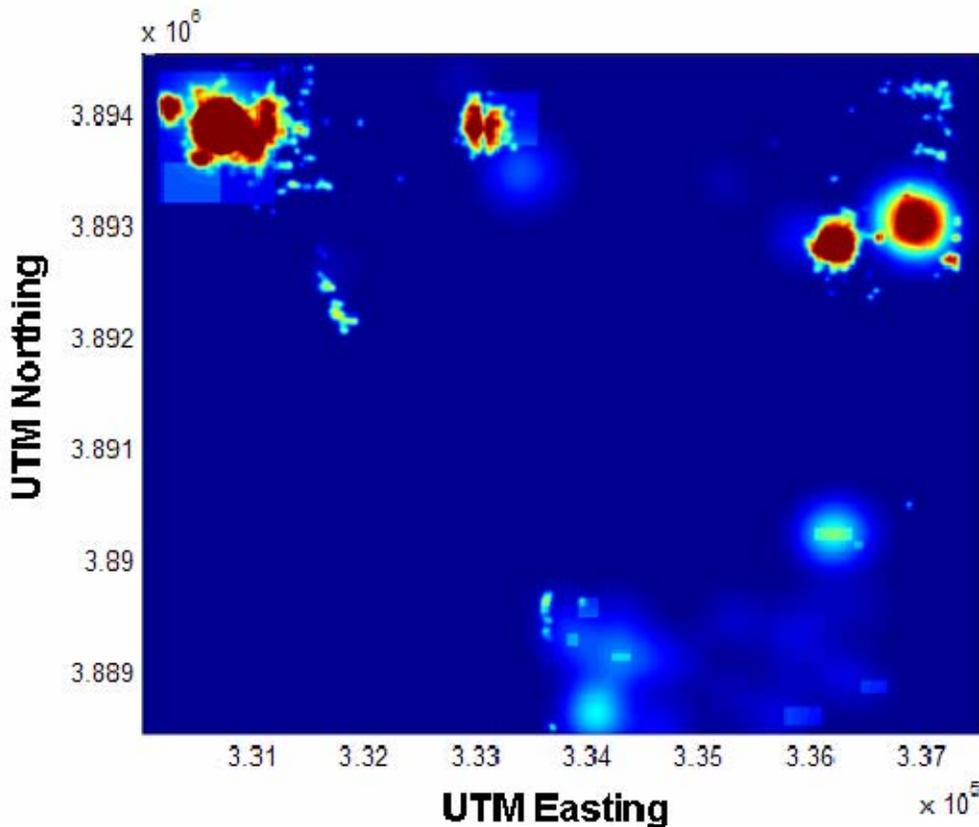
*Summary of feature layers.* Feature layer development and generation resulted in the following feature data layers for data fusion:

- Density of helicopter magnetometry anomalies acquired from Sky Research, both with and without NRL nearest-neighbor clustering filter.
- Density of threshold-applied, morphologically filtered magnetometry anomalies from the NRL pixel clustering algorithm
- Density of automatically detected craters
- Density of manually detected craters from Versar (Kirtland only)
- Manually delineated bombing targets from Versar
- Manually delineated munitions areas from Versar
- Manually delineated man-made structures from Versar

In general, the generated feature layers were of two types: First, layers that were direct indications of the presence of UXO, such as the magnetometry data and anomalies for which  $p(\text{UXO})$  and  $p(\neg\text{UXO})$  were well-defined. And second, those layers that were indirect indications of the presence of UXO, for which only  $p(\text{UXO})$  was well-defined. These latter feature layers included crater anomalies, manually identified target areas, munitions areas, and man-made structures.

### Data Fusion Results

*Heuristic approach.* The output of heuristic-based fusion of a subset of available data layers is shown in Figure 32. In this example, the output was the result of a linear combination of intensity values from the magnetic anomaly density, manual crater density, bombing target, and munitions area feature layers available for the Kirtland site. In essence, the linear combination of layers is functionally equivalent to the overlay of multiple feature layers and represents roughly the same level of knowledge that simple visualization tools provide. The output denotes areas of likely UXO contamination, but says nothing about the likelihood of UXO contamination in those areas.



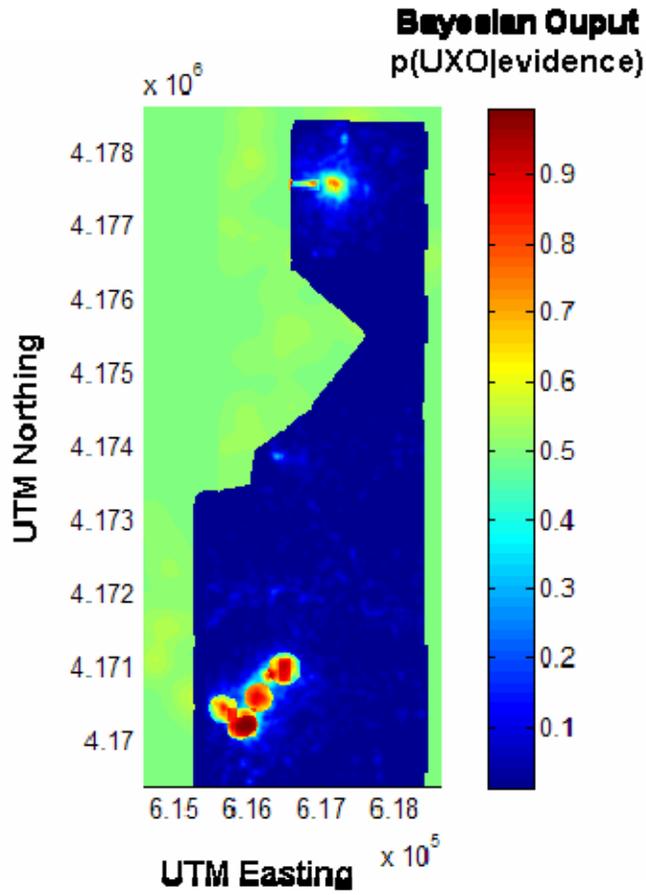
**Figure 32.** Heuristic-based data fusion output utilizing magnetic anomaly density, manually detected crater density, bombing target, and munitions area feature layers. (Map is color scaled from blue=0 to red=1).

*Bayesian approach.* A Bayesian algorithm implementation was tested utilizing similar subsets of feature layers: manually delineated bombing targets and munitions areas, automatically detected crater density, and magnetic anomaly density for the Pueblo site, and manually delineated bombing targets and munitions areas, manually detected crater density, and magnetic anomaly density for the Kirtland site. Each feature layer required the specification of conditional probabilities as a function of intensity value. Initial assignments to the conditional probability distributions were made, reflecting a subjective assessment of the strength of association between each feature layer and the presence of UXO. Assignments were made to the extremal values of feature intensity (0 and 1) as shown in Figure 33. Intermediate values were calculated through linear interpolation.

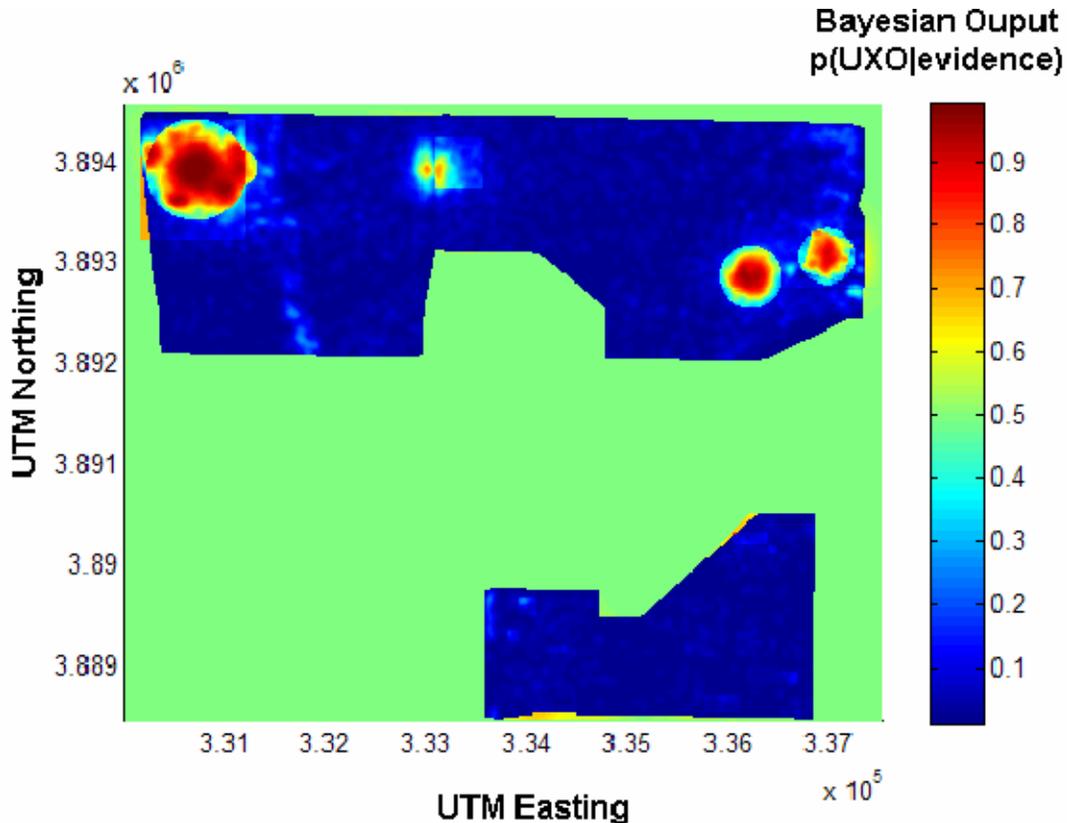
	<b>Feature Intensity</b>	<b><math>p(f UXO)</math></b>	<b><math>p(f \neg UXO)</math></b>
<b>Crater Density</b>	<b>0</b>	<b>0.1</b>	<b>0.1</b>
	<b>1</b>	<b>0.7</b>	<b>0.2</b>
<b>Helimag Anomaly Density</b>	<b>0</b>	<b>0.01</b>	<b>0.8</b>
	<b>1</b>	<b>0.7</b>	<b>0.01</b>
<b>Target Features</b>	<b>0</b>	<b>0.5</b>	<b>0.5</b>
	<b>1</b>	<b>0.8</b>	<b>0.2</b>
<b>Munitions Areas</b>	<b>0</b>	<b>0.5</b>	<b>0.5</b>
	<b>1</b>	<b>0.7</b>	<b>0.3</b>

**Figure 33.** Initial conditional probability assignments for Bayesian data fusion combining helimag anomaly density, crater density, known bombing targets, and manually delineated munitions areas.

The prior probability of UXO,  $p(UXO)$ , was set to 0.5, and reflected complete ignorance as to the presence or absence of UXO at the site before the inclusion of any observational evidence. This prior assignment was updated serially with each new feature layer, as described in the Method section. In this process, the specified conditional probability distribution was used to convert the feature intensity values of each layer into corresponding values of  $p(f | UXO)$  and  $p(f | \neg UXO)$ , and on a point-by-point basis, a posterior probability of UXO,  $p(UXO | f_1, f_2)$  was calculated using the prior and the two conditional probabilities. The resulting output maps are shown in Figures 34 and 35.



**Figure 34.** Bayesian data fusion output for the Pueblo site utilizing magnetic anomaly density, automatically detected crater density, bombing target, and munitions area feature layers.



**Figure 35.** Bayesian data fusion output for the Kirtland site utilizing magnetic anomaly density, manually detected crater density, bombing target, and munitions area feature layers.

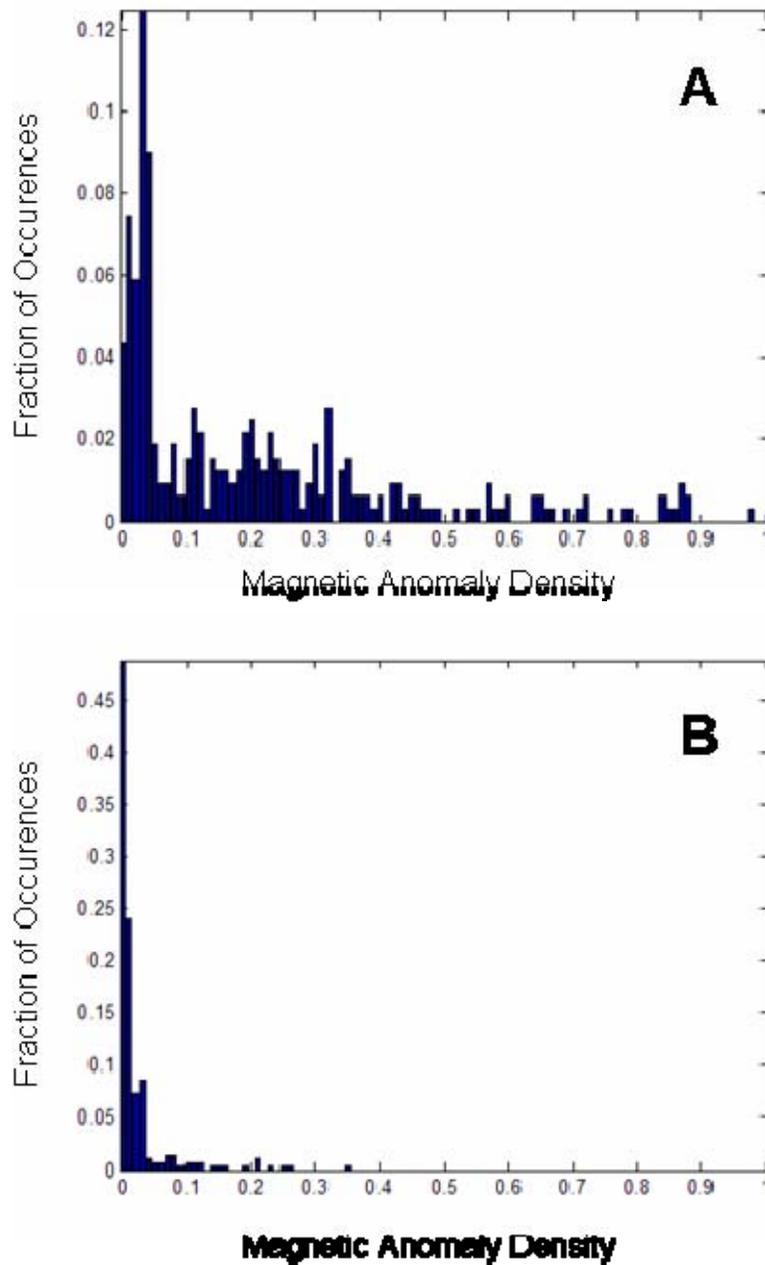
While the output appears reasonable, the conditional probability assignments are problematic for two reasons. First, conditional probability assignments for the UXO application are difficult to estimate in an intuitive fashion based on subjective knowledge. Second, meaningful objective conditional probability distributions are impractical to obtain through empirical means.

The first problem arises when a feature layer has an indirect or poorly defined relationship to UXO. As an example, consider the implications of the probability assignments made to the bombing target feature layer. Subjectively, one supposes that grid points within bombing target features are more predisposed to contamination by UXO or UXO-related scrap than grid points outside of target features. Assuming that this is the extent of our knowledge regarding the relationship between bombing targets and UXO, it is not unreasonable to expect that, in the absence of any other information, grid points outside of target features are as likely to be contaminated with UXO as not. That is, the fact that a grid point is *not* located within a bombing target feature does not convey any information regarding UXO contamination.

Unfortunately, it is difficult to render the subjective knowledge presented by the bombing target feature layer into assignments for conditional probabilities. Grid points within the

feature layer are restricted to one of two values: 1 or 0. The former indicates a pixel is within a bombing target region, the later indicates that it is not. The corresponding conditional probability functions thus have only two assignments, one for each possible feature layer value. For  $p(f | UXO)$ , these values should correspond to the proportion of the total number of UXO-contaminated grid points at the site that are within the bombing targets and those that are outside of bombing targets, respectively. For  $p(f | \neg UXO)$ , these values should correspond to the proportion of non-contaminated grid points that are within and without bombing targets, respectively. To make conditional probability assignments in a principled fashion, a great deal more information is required regarding the nature of the relationship between the feature and UXO. Attempting to estimate values that force behavior agreeing with subjective knowledge results in nonsensical assignments. In order to bring about the proper behavior for grid points outside of target areas,  $p(f | UXO)$  must equal  $p(f | \neg UXO)$  where the feature layer equals zero. Unfortunately, this means that  $p(f | UXO)$  must also equal  $p(f | \neg UXO)$  where the feature layer is one, as the conditional probability distributions must integrate to unity over all possible feature values. Assignment of conditional probability values that represent a predisposition of UXO contamination for grid points within bombing targets, which reflects our subjective knowledge of the feature, thus forces either the abandonment of probabilistic tractability (i.e., conditional probability distributions that do not integrate to unity,) or of the desired treatment of grid points outside of bombing targets (i.e., that such information represents complete ignorance as to the presence or absence of UXO.)

For feature layers with a more direct relationship to UXO, it is more likely that meaningful conditional probability functions can be estimated through empirical evidence. For example, at the Pueblo site, a limited amount of ground truth data were obtained by SERDP in which UXO detections via surface magnetometry were physically examined by digging at each detect site. These surveys included regions suspected to contain UXO, as well as one region suspected to be free of UXO. The resulting 621 detects were characterized as UXO or UXO-related scrap, non-UXO scrap, or geologic (i.e., no object was located during the dig.) Extracting the magnetic anomaly density at each dig location yields the histograms shown in Figure 36.



**Figure 36.** Relationship between ground truth survey data from the Pueblo site and magnetic anomaly density feature layer. (A) depicts the distribution of density values for locations where UXO or UXO-related objects were found while (B) depicts the distribution of density values for locations determined to be free of UXO.

Such data could be used to generate estimates of  $p(f | UXO)$  and  $p(f | \neg UXO)$ , but these implicitly assume that the sampling employed in the ground truth survey represented an unbiased sampling of the contaminated and non-contaminated regions of the site. Such an assumption is likely to prove false, due to the limited scope of the survey and due to the fact that only regions resulting in a “detect” from surface magnetometry were dug up,

resulting in a bias of the data towards UXO-contaminated regions. A truly randomized sampling of the site with the intent to produce more accurate estimates of  $p(f | \text{UXO})$  and  $p(f | \neg\text{UXO})$  is likely to be prohibitively expensive in terms of time and cost, negating the utility of an empirical approach.

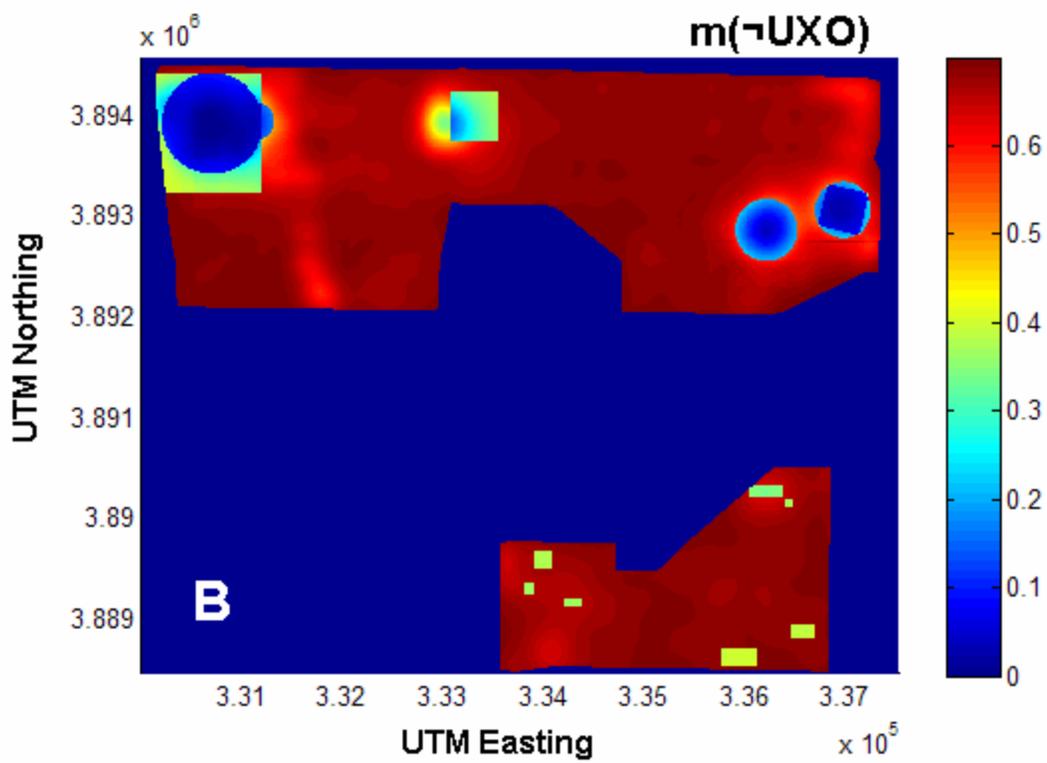
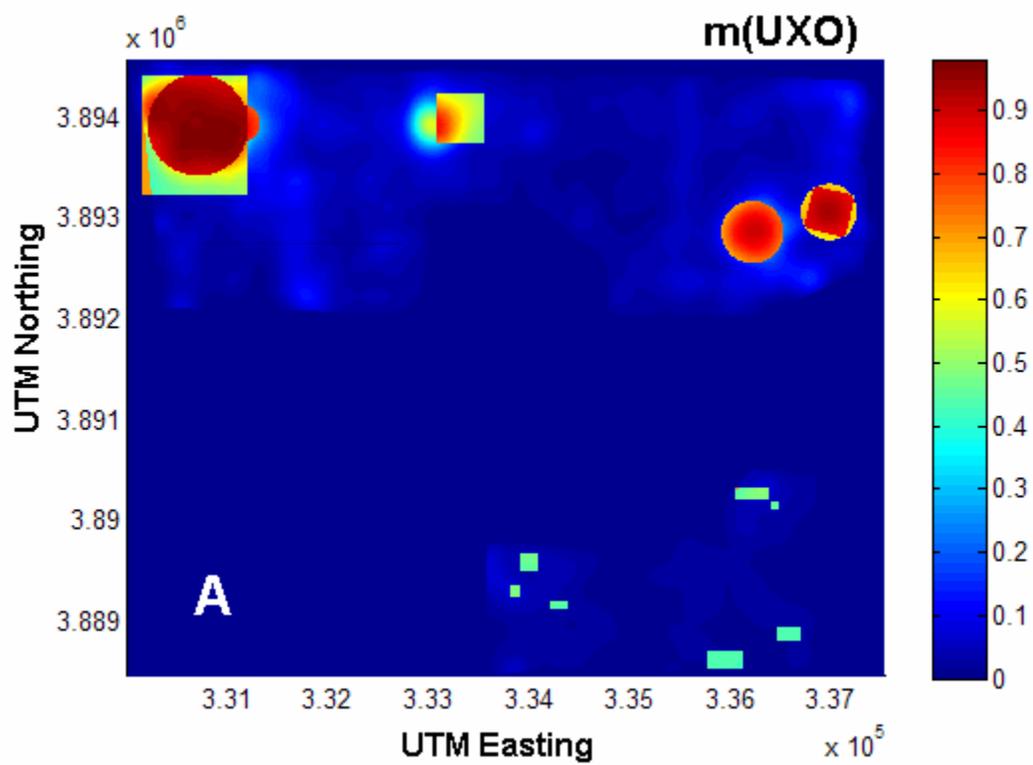
*Dempster-Shafer approach.* In order to test the Dempster-Shafer data fusion algorithm implementation, a subset of feature layers were chosen for input and assigned preliminary probability mass functions,  $m()$ , for the UXO frame of discernment,  $\{\text{UXO}\}$ ,  $\{\neg\text{UXO}\}$ , and  $\{\text{UXO}, \neg\text{UXO}\}$ . As detailed probability mass functions were unknown, probability masses were assigned to the extremal feature intensity values (0 and 1) for each layer. Probability mass assignments for intermediate feature intensity values were then generated through linear interpolation from the extremal values, as described earlier.

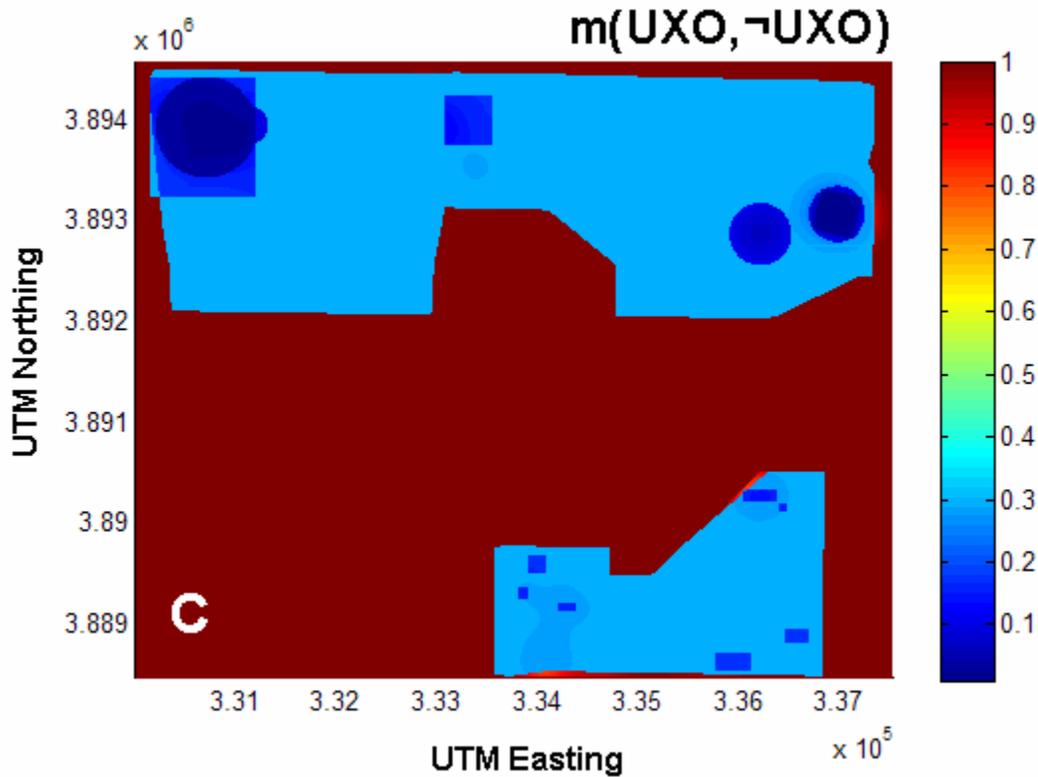
The assignments for the first set of selected feature layers are summarized in Figure 37. For zero crater density, all probability mass was assigned to uncertainty. This reflected the fact that an absence of craters provides neither positive nor negative evidence for the presence of UXO-related objects. At the maximum crater density observed (i.e., an intensity value of 1), a probability mass of 0.75 was assigned to the proposition that UXO was present and 0.25 to uncertainty, reflecting the assertion that relatively high crater densities are associated with the presence of UXO-related objects, although not unequivocally so. The probability mass assignments for each of the two manually-delineated feature layers followed a similar pattern. Where these features were not present, no information was provided, and thus the probability mass assignment went entirely to uncertainty. Where the features were present, an association with UXO was indicated by a non-zero assignment to  $m(\text{UXO})$  that was proportionate to the degree of belief that feature provided to the proposition that UXO was present. The remainder of the mass was assigned to uncertainty. Helicopter magnetometry uncertainty assignments reflected the fact that a lack of anomaly density provided partial evidence for an absence of UXO-related objects, while a high anomaly density provided partial evidence for the presence of UXO-related objects.

	Feature Intensity	$m(UXO)$	$m(\neg UXO)$	$m(UXO, \neg UXO)$
Crater Density	0	0	0	1
	1	0.75	0	0.25
Helimag Anomaly Density	0	0	0.7	0.3
	1	0.7	0	0.3
Target Features	0	0	0	1
	1	0.8	0	0.2
Munitions Areas	0	0	0	1
	1	0.7	0	0.3

**Figure 37.** Probability mass assignments for Dempster-Shafer data fusion for extreme feature intensity values (0 and 1) of layers: helimag anomalies, manually detected craters, known bombing targets, and manually delineated munitions areas.

The probability mass assignments were input along with the corresponding feature layers to the Dempster-Shafer algorithm implementation. Figure 38 depicts the resulting output site-wide assessments of  $m(UXO)$ ,  $m(\neg UXO)$ , and  $m(UXO, \neg UXO)$ .





**Figure 38.** Dempster-Shafer data fusion output combining helimag anomaly density, manually detected crater density, known bombing targets, and manually delineated munitions areas at the Kirtland site.

Figure 38(A) is a color-scaled map depicting the output assessment of  $m(\text{UXO})$ , that is, the degree of belief the four lines of evidence assigned to the proposition that UXO was present. Dominating this map are the regions corresponding to known bombing targets and munitions areas, which generally coincide with relatively high densities of magnetic anomalies and craters. The potential munitions areas in the southern region of the Kirtland site exhibit a much smaller value assigned to  $m(\text{UXO})$  in the data fusion output. The smaller value was due to the lack of corroborating evidence provided by the other feature layers. Figure 38(B) is a color-scaled map depicting the output assessment of  $m(\neg\text{UXO})$ , or the proposition that UXO-related objects are not present. Since the magnetic anomaly feature layer was the only one providing evidence corroborating this proposition, we see high values of  $m(\neg\text{UXO})$  wherever magnetic anomaly density was low, and no other evidence contradicted it. Finally, Figure 38(C) is a color-scaled map depicting the output assessment of  $m(\text{UXO}, \neg\text{UXO})$ , or uncertainty as to whether UXO is present or not. As expected, the data fusion algorithm assigned a high value to this parameter in areas where survey data were missing or unavailable. Elsewhere, mass assigned to uncertainty was much less, reflecting the increased certainty with which the presence or absence of UXO was supported by available evidence.

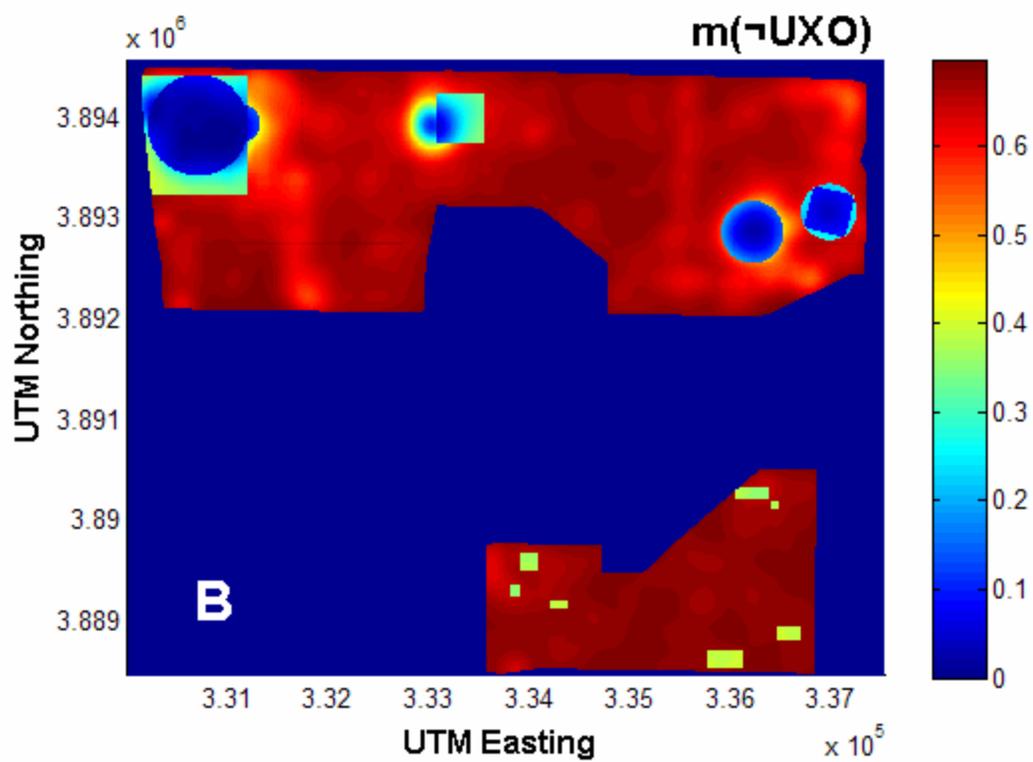
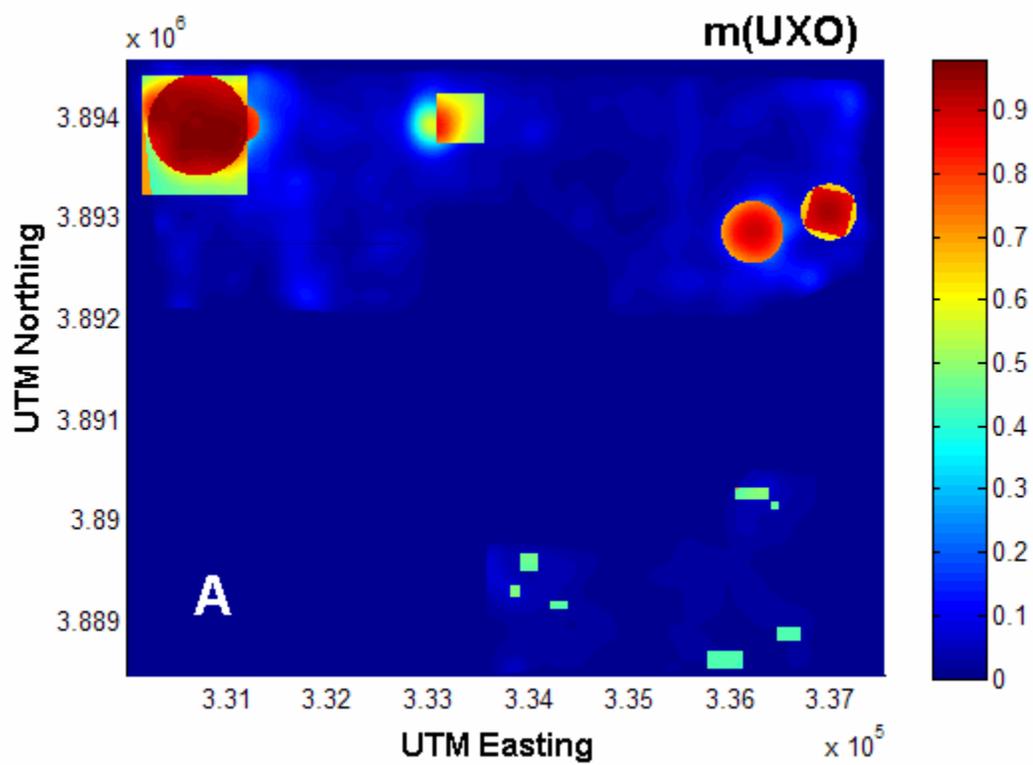
A comparison of Figures 32, 35, and 38 illustrates the difference between the data fusion approaches. Although each output depicts more or less the same areas of interest within the Kirtland site, the amount of information contained is very different. The weighted overlay shown in Figure 32, for instance, provides only limited indication as to the degree of UXO likelihood. The Bayesian output shown in Figure 35, however, indicates a relative propensity for or against the presence of UXO-related objects, with an output of 0.5 indicating complete uncertainty. Finally, the Dempster-Shafer output depicted in Figure 38 provides three output maps, collectively indicating the proportion of the available evidence that supports each of the three propositions of the frame of discernment at each point in the survey area.

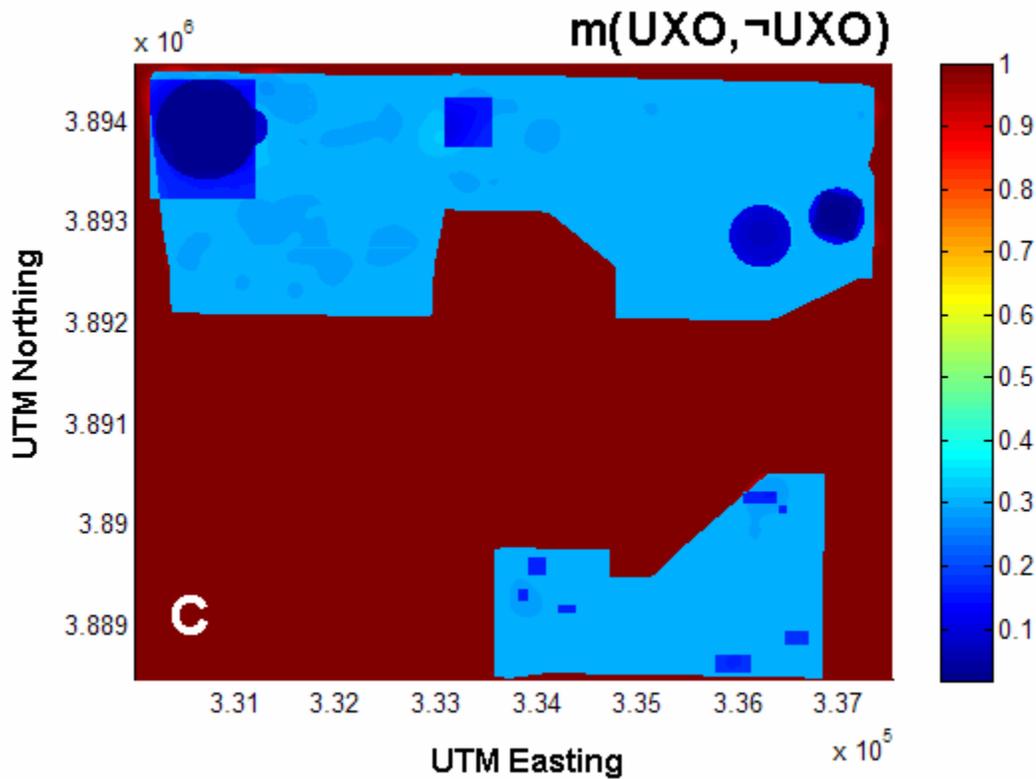
Next, an alternate scenario was tested where a different set of feature layers was utilized. In this case, it was simulated that manually detected craters and magnetic anomaly features were not available, and in their place, automatically detected craters derived from LiDAR data and threshold-applied, morphologically filtered, helimag signal were used. The probability mass assignments were adjusted accordingly to reflect an increased degree of uncertainty in their ability to correctly indicate UXO-related objects, as shown in Figure 39.

	<b>Feature Intensity</b>	<b><math>m(UXO)</math></b>	<b><math>m(\neg UXO)</math></b>	<b><math>m(UXO, \neg UXO)</math></b>
<b>Auto-detected Crater Density</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>1</b>
	<b>1</b>	<b>0.5</b>	<b>0</b>	<b>0.5</b>
<b>Threshold Helimag Density</b>	<b>0</b>	<b>0</b>	<b>0.7</b>	<b>0.3</b>
	<b>1</b>	<b>0.65</b>	<b>0</b>	<b>0.35</b>
<b>Target Features</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>1</b>
	<b>1</b>	<b>0.8</b>	<b>0</b>	<b>0.2</b>
<b>Munitions Areas</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>1</b>
	<b>1</b>	<b>0.7</b>	<b>0</b>	<b>0.3</b>

**Figure 39.** Probability mass assignments for Dempster-Shafer data fusion for extreme feature intensity values (0 and 1) of layers: thresholded helimag signal with morphological filtering, automatically detected craters, known bombing targets, and manually delineated munitions areas.

Utilizing these assignments and the corresponding feature layers, the output shown in Figure 40 was obtained from the Dempster-Shafer based data fusion algorithm.

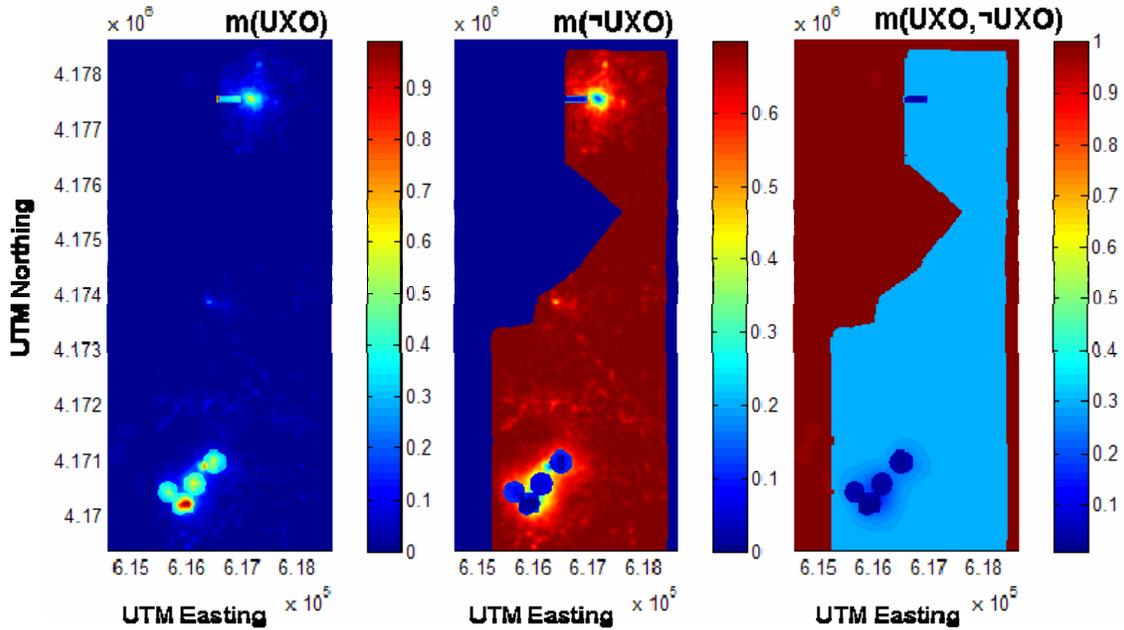




**Figure 40.** Dempster-Shafer data fusion output combining thresholded helimag signal with morphological filtering, automatically detected craters, known bombing targets, and manually delineated munitions areas at the Kirtland site.

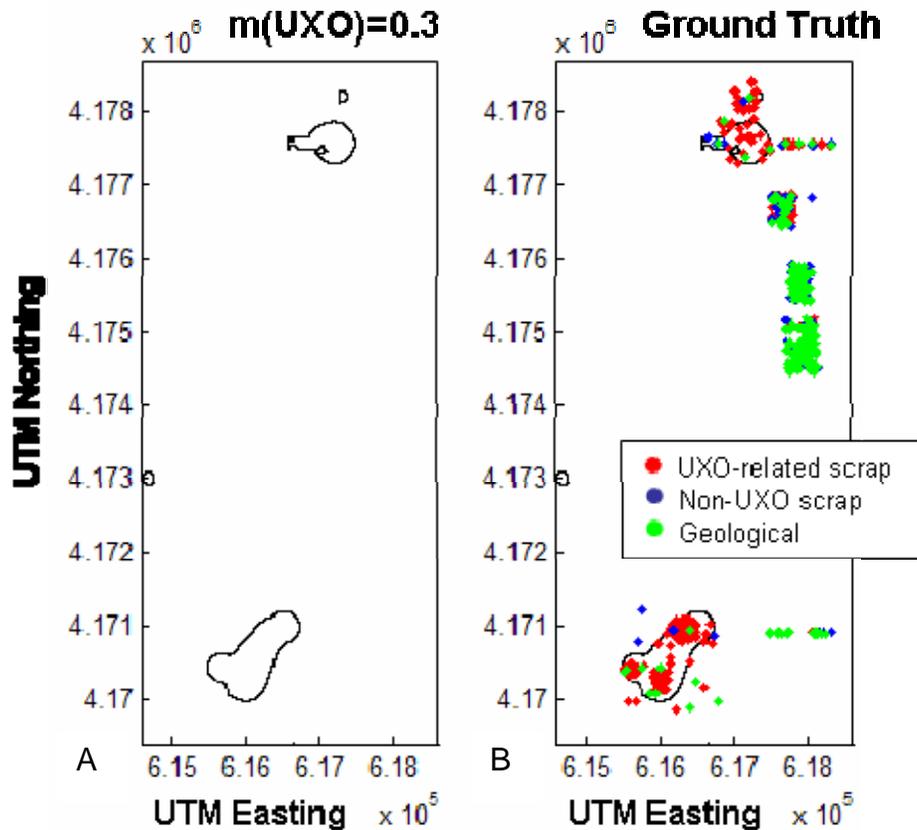
As compared to the previous feature layer set, essentially the same output was obtained, although in Figure 40(B) slightly lower values were assigned to  $m(\neg\text{UXO})$  in many areas, reflecting the relatively higher densities observed throughout the site with this magnetometry feature layer.

A similar exercise was performed with an analogous set of feature layers available for the Pueblo site, utilizing the same probability mass assignment functions shown in Figures 37 and 39. In this example, magnetic anomaly density and automatically-detected crater location density feature layers were utilized. The results of Dempster-Shafer data fusion with these Pueblo feature layers are shown in Figure 41.



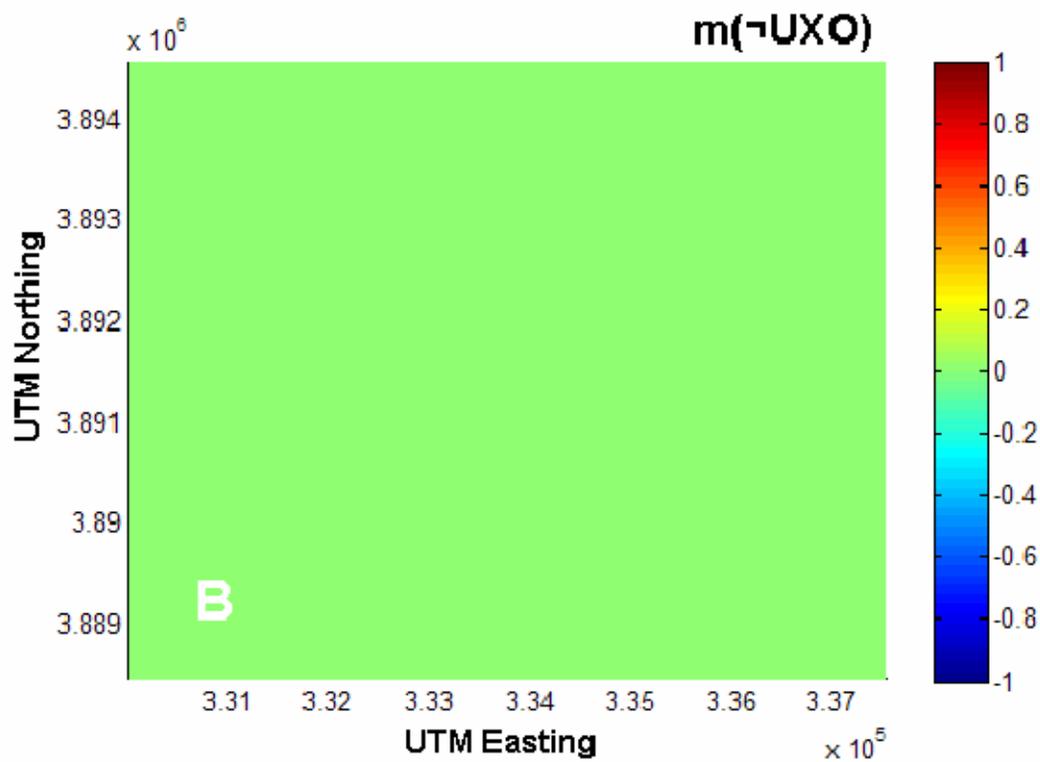
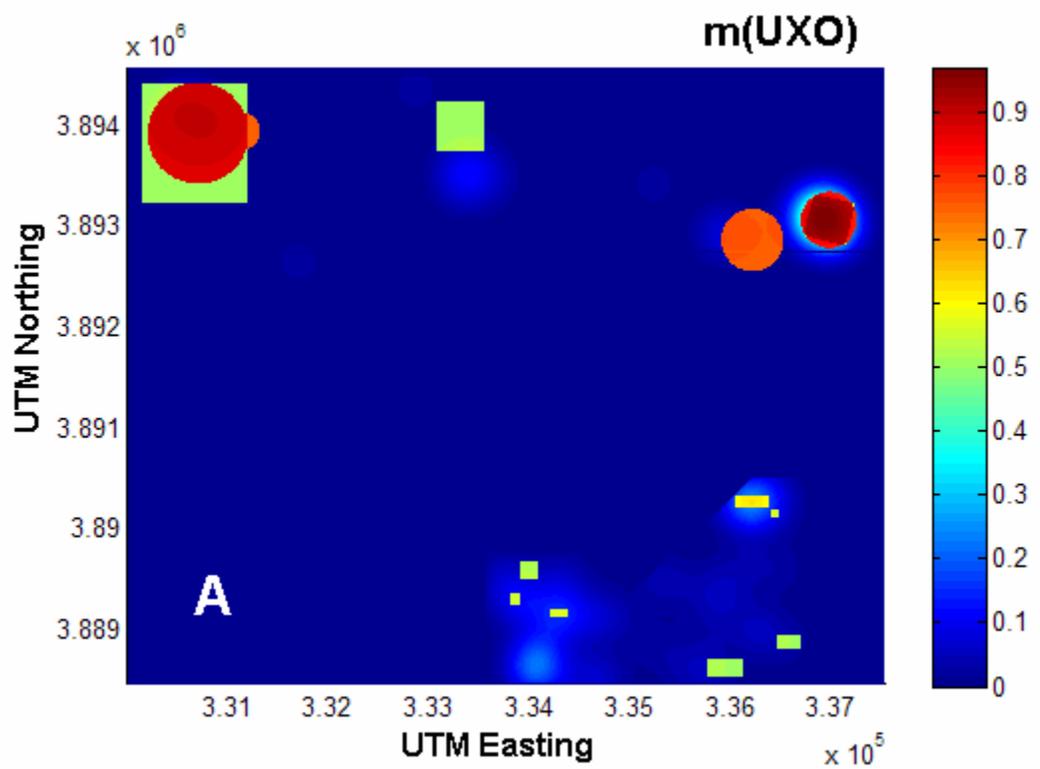
**Figure 41.** Dempster-Shafer data fusion output combining helimag anomaly density, automatically detected crater density, known bombing targets, and manually delineated munitions areas at the Pueblo site.

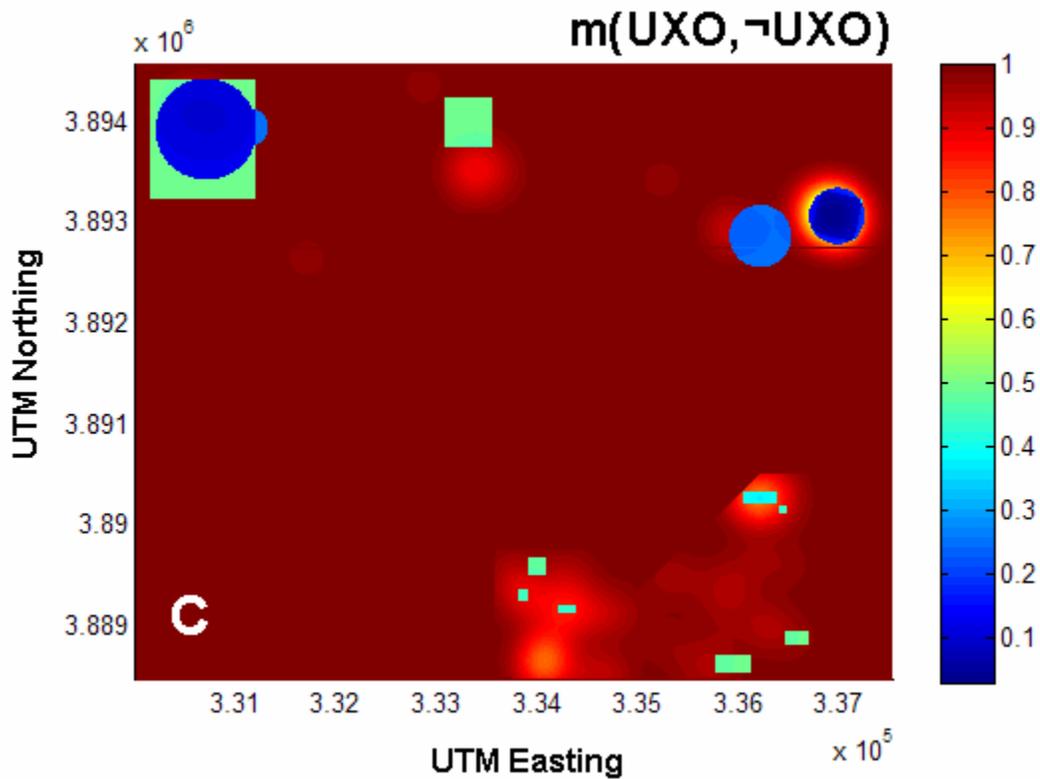
As with the Kirtland data fusion results, the highest values assigned to  $m(\text{UXO})$  occurred where multiple lines of corroborating evidence supported the proposition that UXO-related objects were present, while the highest values assigned to  $m(\text{UXO}, \neg\text{UXO})$  occurred where no data were available. By setting a threshold of 0.3 for  $m(\text{UXO})$ , as shown in Figure 42(A), a potential delineation of UXO-contaminated areas was generated. When overlaid with limited ground truth data available at the Pueblo site, as in Figure 42(B), it can be seen that the delineated regions were consistent with ground truth.



**Figure 42.** A preliminary delineation of areas of likely UXO contamination utilizing Dempster-Shafer data fusion output for the Pueblo site (A), threshold set at 0.3. Overlaid on this plot are limited truth data showing agreement with this assessment (B).

The effects of adding and subtracting feature layers were examined by recalculating the data fusion algorithm output with a reduced feature layer set. Figures 43 and 44 demonstrate the Dempster-Shafer algorithm's ability to function when input data streams are missing, but a comparison to Figure 38 shows the toll subtracting an information source can have on data fusion output. The effect of the removal of the magnetic anomaly density feature layer is shown in Figure 43, which depicts the output of Dempster-Shafer fusion of manually-located crater density with manually identified bombing targets and manually delineated potential munitions remediation areas.

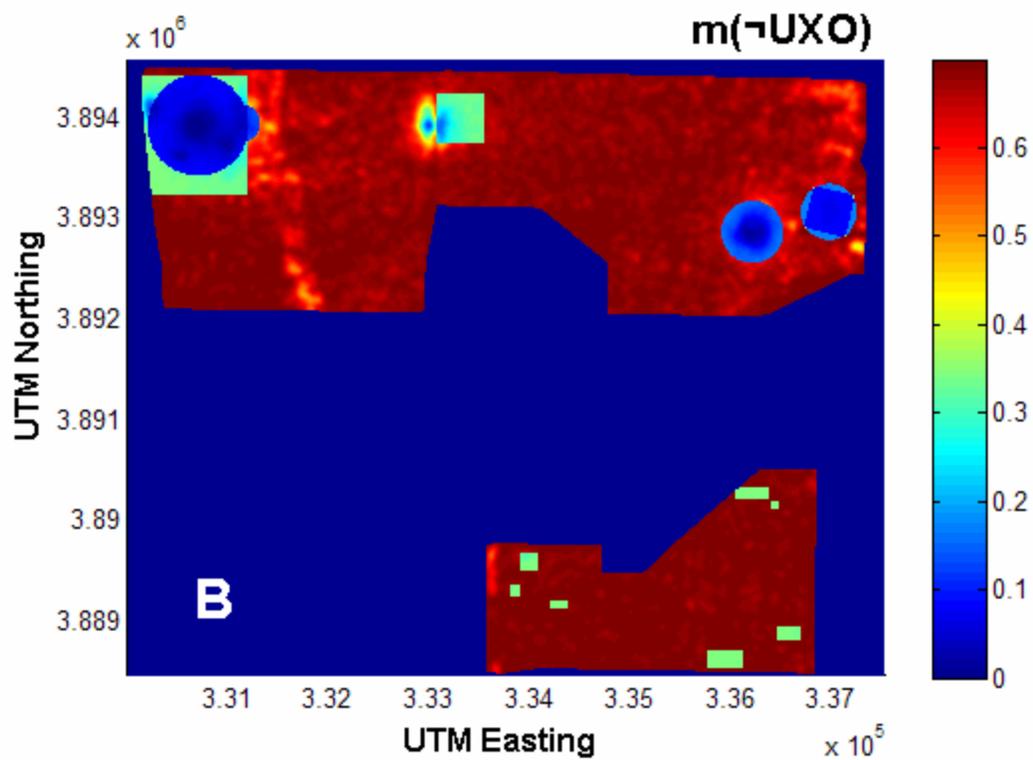
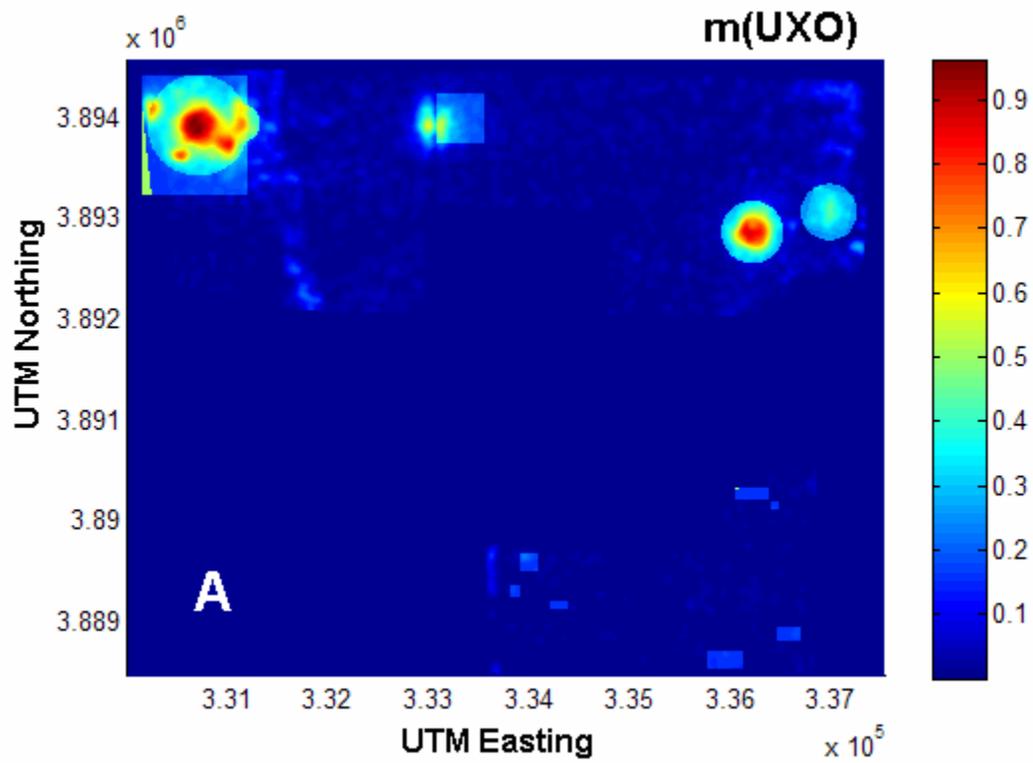


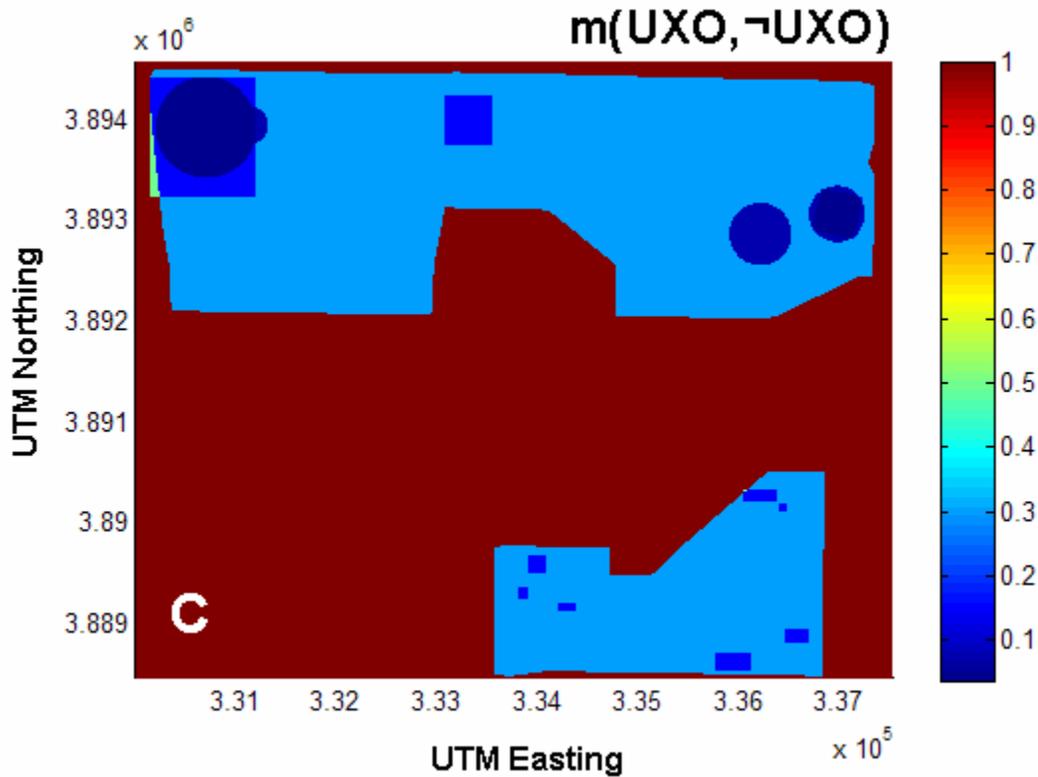


**Figure 43.** Dempster-Shafer data fusion output combining manually detected craters, known bombing targets, and manually delineated munitions areas.

The removal of the magnetic anomaly density feature layer resulted in a zero assignment of belief to  $m(\neg\text{UXO})$ , as well as a relative increase in the belief assigned to  $m(\text{UXO})$  for the areas delineated as potential munitions areas. In both cases, the changes were due to the absence of evidence that a lack of magnetic anomaly density provided to the proposition the UXO was not present.

Alternatively, removing the crater density feature layer resulted in the output shown in Figure 44, which depicts the output of Dempster-Shafer fusion of helicopter magnetometry density with manually identified bombing targets and manually delineated potential munitions remediation areas.





**Figure 44.** Dempster-Shafer data fusion output combining helimag anomalies, known bombing targets, and manually delineated munitions areas.

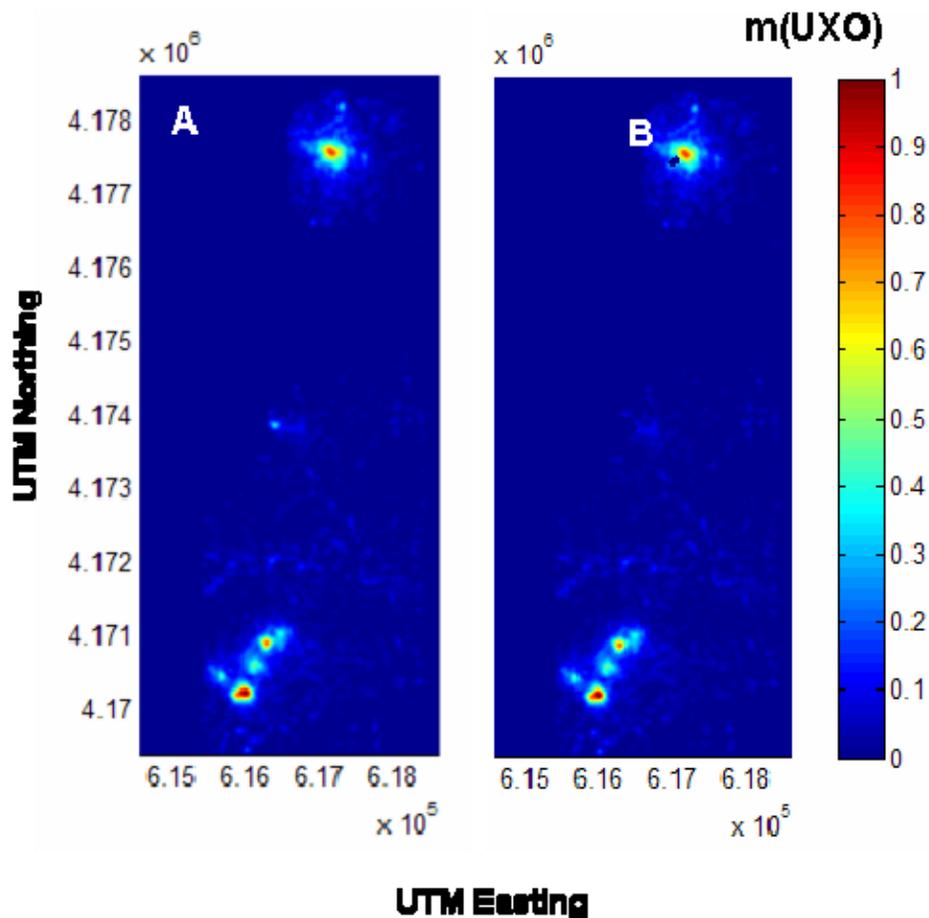
The removal of the crater density feature layer resulted in a relatively smaller assignment to  $m(\text{UXO})$  for the manually-located bombing targets at the eastern and western edges of the Kirtland survey area, as these areas contained relatively high densities of manually identified craters.

*Hybrid heuristic/Dempster-Shafer approach.* As data fusion algorithm development progressed, it became clear that in certain instances heuristic rules were required to capture specific feature-to-feature interdependencies. For example, it is possible that manually delineated regions of interest in the feature set may describe locations or objects associated with significant non-UXO magnetic signal, such as fence lines, man-made structures, and pipelines. Accordingly, the impact of magnetic signals recorded at these locations on the output UXO assessment should be minimized or blocked. A second example reflects the incorporation of measures of data quality. The helicopter magnetometry data layer was estimated from point measurements that were acquired at non-regular spatial intervals and provided non-homogenous coverage of the survey site. Areas with sparse data point coverage in the magnetometry data provided less certain information than those with dense coverage. Knowledge of the quality of data point coverage needed to be reflected in the output UXO assessment. Thus, in a general sense it became important to explore means by which these kinds of knowledge could be

incorporated into the data fusion framework. The following implementations were developed for the two heuristic rules just described.

A simple rule incorporating the concept of blocking known non-UXO magnetic signals was implemented as follows. First, feature layers associated with significant non-UXO magnetic signal were flagged as such. Second, feature layers susceptible to this form of magnetic interference were flagged as such. Finally, a general rule was added to the data fusion algorithm indicating that whenever a feature layer belonging to the latter set was being used to update a current UXO assessment, the regions of that feature layer coinciding with those identified in the former (blocking) set were shrunk to a value of zero.

The effect of this heuristic rule can be observed in the data fusion output generated with and without the rule in place, as shown in Figure 45.



**Figure 45.** UXO assessment (A) before and (B) after incorporation of a heuristic rule blocking magnetic signal in regions containing known interferences.

The principal difference lies in the reduction of belief assigned to  $m(\text{UXO})$  in an area in the center of the survey region, near the point (4.174, 6.16). This area, which was manually identified as containing man-made structures, resulted in a relatively high density of magnetic anomalies that were presumably not due to UXO-related objects and an artificially high assignment of probability mass to the region. After application of the heuristic rule, the feature density from magnetic anomalies in this region was no longer considered in the final data fusion output.

Implementation of a rule incorporating data density was accomplished as follows. The first step was the development of a density metric for the helicopter magnetometry data feature layer. As part of the year one effort, a layer containing data point density was calculated as the number of data points observed within a square meter on 1m common grid. A metric was derived from this value as  $1/(1+\text{density})^2$  and ranged in value from 1 (corresponding to zero coverage) to close to 0 as coverage increased. The resulting map of density metric values was then convolved with a Gaussian kernel for smoothing to allow for neighborhood effects. The final data density metric is shown in Figure 46.

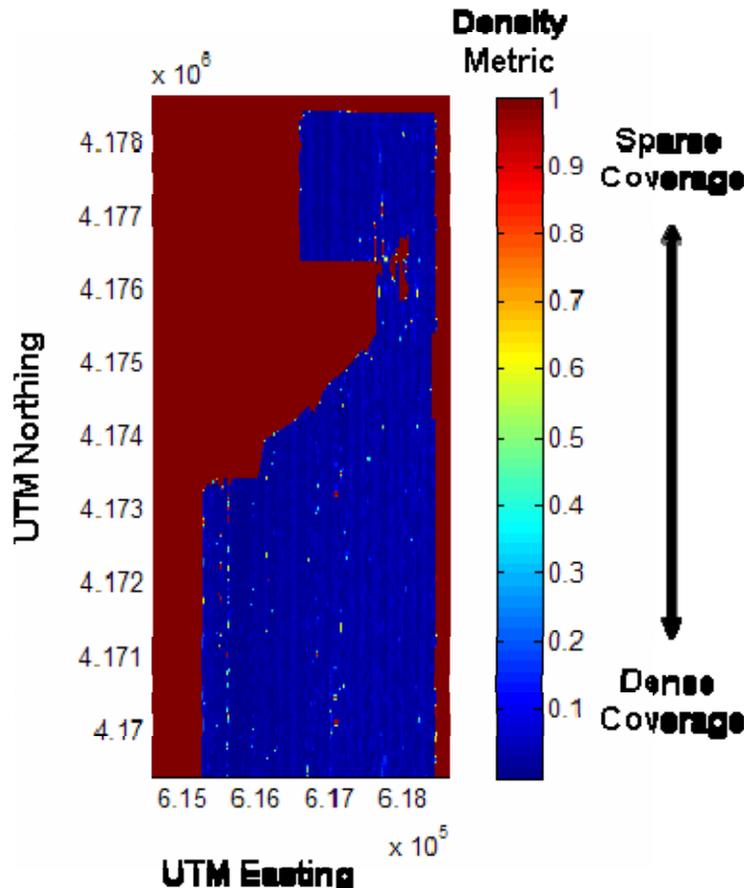
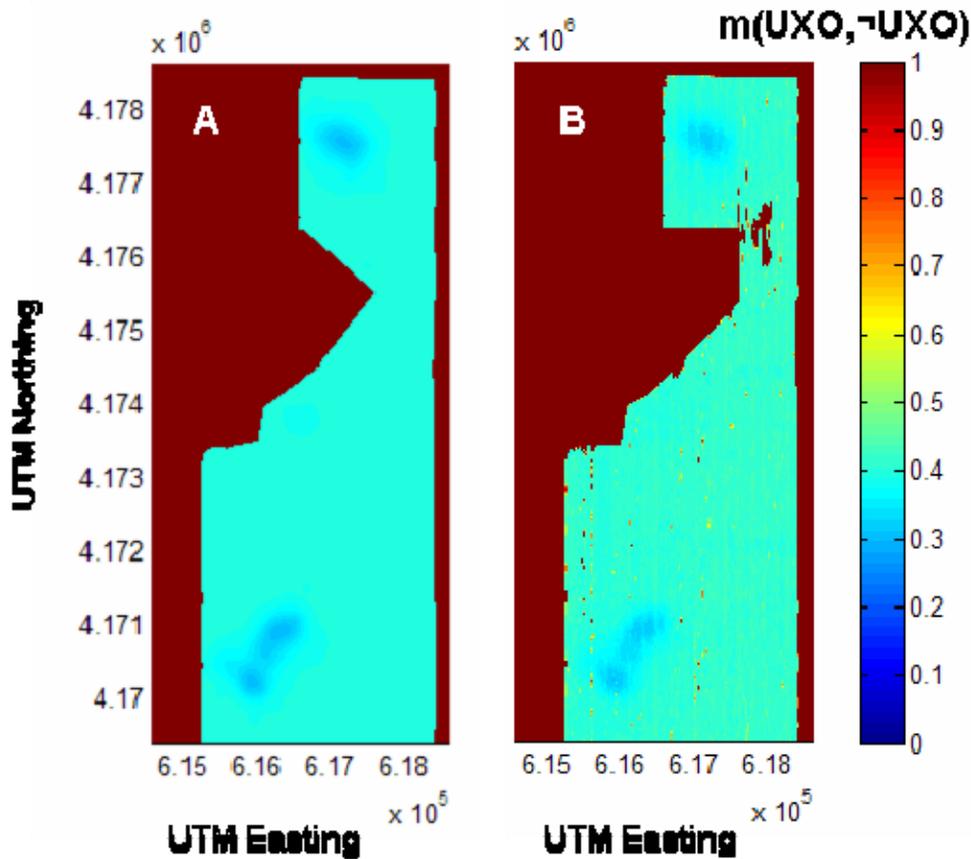


Figure 46. Site-wide values for the helicopter magnetometry data density metric.

Once calculated, the data density metric was used to modify probability mass assignments for magnetometry feature layers. This was accomplished by, again flagging the set of feature layers susceptible to this rule. When such a feature layer was being used to update an output UXO assessment, the data density metric layer was loaded and used to modify the probability mass assignments as follows: grid points with a metric of one were assigned a probability mass associated with complete uncertainty. Grid points associated with a density metric close to zero were assigned the probability masses indicated by the appropriate probability mass functions for that layer. Finally, grid points with a data density metric intermediate between one and zero were assigned probability masses equal to a linear combination of these two extremes, scaled according to the value of the density metric. The effect of this heuristic rule can be observed in the data fusion output generated with and without the rule in place, as shown in Figure 47.

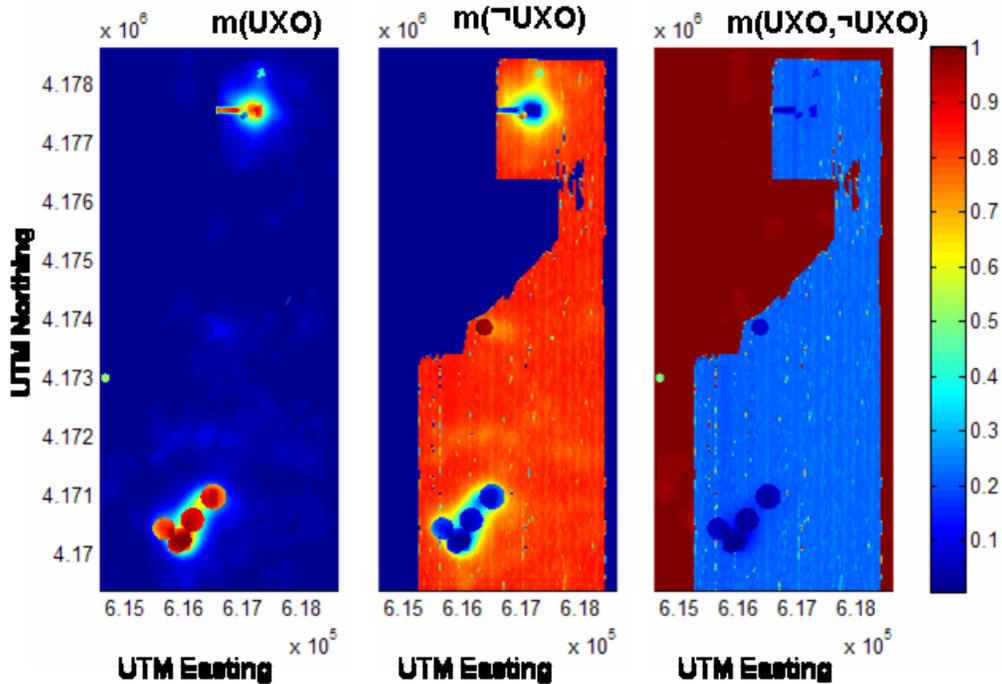


**Figure 47.** UXO assessment (A) before and (B) after incorporation of a heuristic rule incorporating a data density metric for helicopter magnetometry feature layers.

The impact of the heuristic can be seen in the increase in assignment to  $m(\text{UXO}, -\text{UXO})$ , which is most pronounced in areas with zero data density where the assignment of belief goes almost completely to uncertainty. A revised data fusion output for the Pueblo site,

incorporating both of the developed and implemented heuristics rules is shown in Figure 48.

Heuristic rules such as these allow for the incorporation of specific knowledge regarding relationships between feature sets. In the first example, the incorporation of a rule blocking magnetometry signal in regions with man-made structures led to a reduction of false positive indications of UXO. In the second example, the incorporation of a rule attenuating the impact of magnetometry signal in areas of sparse data density allowed for a more accurate assessment of the uncertainty present in the resulting Dempster-Shafer data fusion output maps.



**Figure 48.** Dempster-Shafer data fusion output combining thresholded helimag signal with morphological filtering, automatically detected craters, known bombing targets, and manually delineated munitions areas at the Pueblo site. Heuristic rules involving feature blocking and data point density metrics are implemented.

## Summary of Results

The major accomplishments for Project MM-1510 are listed below:

- A common geo-referenced database for registering data, features, and associated meta-data resulting from wide-area surveys.
- The evaluation of standard feature extraction strategies on wide-area survey data for their potential in assessing UXO contamination.
- Automatic feature extraction algorithms for identification of craters in LiDAR data and anomalies in helicopter magnetometry data.
- A wavelet filtering algorithm was developed to improve the automatic crater identification algorithm. Filtering reduced false positive crater identifications by removing surface texture artifacts.
- A statistical nearest-neighborhood clustering filter algorithm was developed to improve the magnetic anomaly data by locating anomalies that were part of non-uniform groupings. The clustering algorithm reduced background noise and false positive assessments of UXO in the final data fusion output.
- A prototype Data Fusion Framework based on a hybrid Dempster-Shafer theory.
- A methodology for the input and registration of disparate data and feature streams from wide-area sensing technologies and other site-specific knowledge.
- A methodology allowing the input of meta-information regarding the relationship between these lines of evidence and the presence (or absence) of UXO or UXO-related objects.
- A method by which heuristic rules can be incorporated into the Data Fusion Framework in order to take advantage of specific known interdependencies between feature layers.
- Demonstration of the Data Fusion Framework prototype's ability to successfully provide useful output assessment of UXO likelihood from an assortment of WAA assessment data and features.

The principal accomplishment of Project MM-1510 was the development of a prototype Data Fusion Framework suitable for wide-area assessment of UXO contamination. A key enabling technology was the development of a generalized method for processing input data feature streams from UXO WAA survey efforts. The development of this method is significant, as it requires only a limited number of specifications to be imposed on input features, allowing a wide range of feature sets and relationships to be formatted and input for data fusion. Such flexibility is crucial, as the disparate nature of the data and features available from potential WAA survey techniques presents a significant impediment towards adoption of more basic data fusion approaches.

Features relevant for wide-area assessment of UXO contamination were successfully obtained from the various data sources. Automatic crater extraction from the LiDAR

data was successful. An algorithm based on the circular Hough transform was able to extract a value of four meters as the characteristic diameter of craters at the Pueblo site. Further, a pattern recognition algorithm based on the morphology of the craters was developed to locate them in the LiDAR data. This information was then converted into a feature map describing the density aspect of craters. Feature maps describing the intensity and quality aspects to craters provide data fusion algorithms with additional discriminatory information.

Feature extraction algorithms were also developed for magnetometry data. Due to the minimal geomagnetic features at the Pueblo site, a simple threshold proved effective in eliminating geologic background, which is expected to be more significant at other ESTCP WAA Pilot Program sites. A pattern recognition algorithm was developed to separate ordnance-related signal from the ferromagnetic background of man-made structures. This information was then converted to a feature map describing the density aspect of ordnance-related material. Separate feature maps describing the intensity and morphology aspects of the ordnance-related and man-made components of data are also expected to provide additional discriminatory information for data fusion.

Successful feature extraction algorithms developed in year one were further optimized and validated against new data, including an automatic crater detection algorithm and a nearest-neighbor clustering filter to selectively remove false-positive magnetometry anomalies. Methods for generating feature layers from extracted features for input to data fusion were developed and implemented. For each input feature set, a corresponding feature intensity map and specification of a functional relationship between a feature's intensity and the likelihoods for or against the presence of UXO that are supported by the feature's intensity values. Heuristic, Bayesian, and Dempster-Shafer theoretic algorithms for combining evidence presented in feature layers were investigated as possible engines for a UXO WAA data fusion framework prototype. These were implemented as MATLAB code and evaluated with feature layers generated from both the Pueblo and Kirtland site data acquired by performers in the ESTCP WAAPP.

The Dempster-Shafer approach, with its ability to quantify uncertainty about evidence, was shown to be the most appropriate data fusion strategy for the UXO problem and proved to be the most successful of the three. The ability to incorporate heuristic rules regarding specific dependencies between input feature layers into the Dempster-Shafer based data fusion framework prototype was described and demonstrated utilizing two specific examples. The first demonstrated a reduction of false positive indications of UXO by utilizing a feature layer comprised of manually identified man-made structures to selectively block magnetometry-derived features. The second demonstrated an adjustment of the impact of magnetometry-derived features on the output assessment of UXO to accurately reflect the uncertainty associated with increased magnetometry data sparseness in some areas of the helimag survey. The prototype data fusion framework developed was able to delineate areas of likely contamination while providing reasonable estimates of the likelihood of that contamination given supporting observational evidence and a priori knowledge. Preliminary results were compared with limited ground truth data available at the Pueblo site and agreed well.

## Conclusions

Automatic feature extraction algorithm development was an early focus of the project, due to the fact that wide-area assessment data are, by nature of their size, difficult to analyze manually in a comprehensive fashion. Successful feature extraction algorithms were developed in year one of the project, including an automatic crater detection algorithm and a method of selecting UXO-related signals in helicopter magnetometry survey data. These algorithms were further optimized and validated against new data in year two of the project and a nearest-neighbor clustering filter was developed and implemented to selectively remove false-positive magnetometry anomalies. As an initial part of this work, a means of importing and organizing large data sets into a common storage format for subsequent viewing and analysis was designed and implemented on the MATLAB computational platform.

Following the feature extraction work, a generalized method for processing input data feature streams from UXO WAA survey efforts was developed. The method requires the generation of a corresponding feature intensity map and the specification of a functional relationship between a feature's intensity and the hypotheses (i.e., the presence or absence of UXO as well as the possibility that at certain values or in certain areas the feature doesn't support a determination either way) that are supported by the feature's intensity values. The development of a general input method is significant, as it requires only a limited number of specifications to be imposed on these two inputs, allowing a wide range of feature sets and relationships to be formatted and input for data fusion. This flexibility is crucial, as the disparate nature of the data and features available from potential WAA survey techniques presents a significant impediment towards adoption of more basic data fusion approaches. As part of this work, various feature sets derived from both ESTCP performers and from customized feature extraction algorithms developed at NRL were utilized as inputs to data fusion, demonstrating the flexibility of the approach.

Heuristic, Bayesian, and Dempster-Shafer theoretic algorithms for combining evidence presented in feature layers were investigated as possible engines for a UXO WAA data fusion framework prototype. These were implemented as MATLAB code and evaluated with feature layers generated from both the Pueblo and Kirtland site data acquired by performers in the ESTCP WAAPP. The Dempster-Shafer approach, with its ability to quantify uncertainty about evidence, was shown to be the most appropriate approach for the UXO problem and proved to be the most successful of the three. The ability to incorporate heuristic rules regarding specific dependencies between input feature layers into the Dempster-Shafer based data fusion framework prototype was described and demonstrated utilizing two specific examples. The first demonstrated a reduction of false positive indications of UXO by utilizing a feature layer comprised of manually identified man-made structures to selectively block magnetometry-derived features. The second demonstrated an adjustment of the impact of magnetometry-derived features on the output assessment of UXO to accurately reflect the uncertainty associated with increased magnetometry data sparseness in some areas of the helimag survey. The prototype data fusion framework developed was able to delineate areas of likely contamination while

providing reasonable estimates of the likelihood of that contamination given supporting observational evidence and a priori knowledge. Preliminary results were compared with limited ground truth data available at the Pueblo site and agreed well.

While successful data fusion requires complementary data sets for input, the key theoretical advantage for wide-area assessment is the ability to reduce false positives while retaining high detection rates. The framework described is flexible, tolerating missing data and allowing multiple configurations of potential input data streams, as well as scalable, allowing new data streams to easily be included in the assessment. Further, the impact of available and new data streams on the output can be readily quantified. When presented with reduced sets of input feature layers, the prototype provided reasonable, although less accurate, assessments of UXO contamination, demonstrating both the robustness of the approach and the improvement provided by data fusion of feature layers containing complementary information. One challenge is that the structured input methodology requires the specification of each feature layer's relationship to the presence or absence of UXO. However, the input methodology allows specification to be accomplished in a highly flexible manner. The user has the ability to input specifications that vary from simple, intuitive estimations based on expert knowledge to detailed functional relationships based on empirical evidence of sensor performance. Thus, the data fusion framework is capable of utilizing all the information and observation evidence available, without necessarily requiring that the exact same inputs be present for assessment. This flexibility is an important feature of the data fusion approach as it is expected that, for a number of reasons, it will rarely be the case that exactly the same types or quality of data will be available for analysis each time a wide-area UXO assessment is performed.

Future work is needed to further the development, evaluation, and optimization of a prototype Dempster-Shafer based data fusion framework with the intended result being a prototype software implementation of the data fusion framework suitable for demonstration and evaluation at various sites of potential UXO contamination. A final, production-grade data fusion framework that is well-suited for independent use by site administrators would most likely require close coordination with SERDP and potential end-users and partnership with an existing GIS software vendor.

## References

1. Available at URLs <<http://www.estcp.org/ux/#Wide>> and <http://www.serdp.org/research/UXO.html>
2. The Mathworks, Inc., “MATLAB Numerical Analysis Software Suite,” <<http://www.mathworks.com>>.
3. Available at URL < <http://www.estcp.org/links/WAA-Pilot-Video.cfm>>
4. B.J. Johnson, T.G. Moore, B.J. Blejer, C.F. Lee, T.P. Opar, S. Ayasli, and C.A. Primmerman, “A Research and Development Strategy for Unexploded Ordnance Sensing,” SERDP project UX-860 final report, April 1996. Available at <http://www.serdp.org/research/UXO.html>
5. Yan Zhang; Collins, L.M.; Carin, L., “Unexploded ordnance detection using Bayesian physics-based data fusion,” *Integrated Computer-Aided Engineering* **10**(3), 231-47, (2003)
6. Collins, L.M.; Zhang, Y.; Carin, L., “Model-based statistical sensor fusion for unexploded ordnance detection,” *Proceedings of IEEE International Geoscience and Remote Sensing Symposium. IGARSS 2002*, 24-28 June 2002, Toronto, Ont., Canada, vol. 3, pg. 1556-9
7. I. Shamatava, F. Shubitidze, c.c. Chem, H.S. Youn, K. O’Neil, K. Sun, “Potential benefits of combining EMI and GPR for enhanced UXO discrimination at highly contaminated sites,” *Proceedings of SPIE* Vol. 5415, 1201-1210, (2004)
8. L.M. Collins, Y. Zhang, J. Li, H. Wang, L. Carin, S.J. Hart, S.L. Rose-Pehrsson, H.H. Nelson, and J.R. McDonald, “A Comparison of the performance of Statistical and Fuzzy Algorithms for Unexploded Ordnance Detection,” *IEEE Transactions on Fuzzy Systems*, **9**(1), 17-30,( 2001)
9. “Intelligent Data Fusion for Wide-Area Assessment of UXO Contamination. SERDP Project MM-1510. 2006 Annual Report” Rose-Pehrsson, S.L., Johnson, K.J., Minor, C.P, Guthrie, V.N. NRL Memorandum Report, NRL/MR/6180—07-9039, April 20, 2007.
10. KDE Toolbox for MATLAB authored by Alexander Ihler and available at URL < <http://www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=7800&objectType=File>>
11. S. James Press, “Bayesian Statistics: Principles, Models, and Applications,” John Wiley & Sons, New York 1989.
12. D.S. Sivia, “Data Analysis: A Bayesian Tutorial,” Oxford University Press Inc., New York 1996
13. G. Shafer, “A Mathematical Theory of Evidence,” Princeton University Press, Princeton, NJ 1976
14. Lawrence A. Klein, “Sensor and Data Fusion: A Tool for Information Assessment and Decision Making,” SPIE Press, Bellingham, WA 2004.
15. Kari Sentz and Scott Ferson, “Combination of Evidence in Dempster-Shafer Theory,” Sandia National Laboratories Report SAND 2002-0835, April, 2002.
16. Z. Yi, Y. Khing, C.C. Seng, Z.X. Wei, “Multi-ultrasonic sensor fusion for autonomous mobile robots,” *Proceedings of SPIE*, Vol. 4051, 314-321, (2000)
17. A. Sarkar, A. Banerjee, N. Banerjee, S. Brahma, B Kartikeyan, M. Chakraborty, K.L. Majumder, “Landcover classification in MRF context using Dempster-Shafer fusion for multisensor imagery,” *IEEE transactions on image processing* **14**(5), 634-45 (2005)
18. M. Raza, I. Gondal, D. Green, R.L. Coppel, “Fusion of FNA-cytology and gene-expression data using Dempster-Shafer Theory of evidence to predict breast cancer tumors,” *Bioinformatics* **1**(5), 170-5, (2006)

19. Duda, R.O., and Hart, P.E., "Use of Hough transformation to detect lines and curves in pictures" *Communications of the ACM*, **15(1)**, 11-15 (1972)
20. Li Bai, Linlin Shen, Yan Wang, "A Novel Eye Location Algorithm based on Radial Symmetry Transform," *Pattern Recognition, 2006 18<sup>th</sup> Intl. Conf. on Pattern Recog. (ICPR'06)*, **3**, 511-514, (2006)
21. Dr. Herb H. Nelson, personal communication (2007)

## **Appendix A: Supporting Data**

The common grid and feature maps are on the attached DVD. Examples of the MATLAB implementations of the feature extraction and data fusion algorithms are also on the attached DVD.

## **Appendix B: List of Technical Publications**

### **Presentations:**

Minor, C., Johnson, K. and Rose-Pehrsson, S., “Intelligent Data Fusion for Wide-Area Assessment of UXO Contamination,” Partners in Environmental Technology Technical Symposium & Workshop, Washington, DC, November 28-30, 2006. (Poster)

Minor, C., Johnson, K. and Rose-Pehrsson, S., “Intelligent Data Fusion for Wide-Area Assessment of UXO Contamination,” Partners in Environmental Technology Technical Symposium & Workshop, Washington, DC, December 4-6, 2007. (Poster)

### **Referred Journal Article:**

Johnson, K.J, Minor, C.P, Guthrie, V.N., and Rose-Pehrsson, S.L., “Intelligent Data Fusion for Wide-Area Assessment of UXO Contamination,” *Stochastic Environmental Research and Risk Assessment (SERRA)*, *in press*.

### **Patent Disclosure:**

Johnson, K.J and Minor, C.P, “Feature Extraction Methods for Data from Wide-Area Assessment of Buried Unexploded Ordnance,” Submitted January 22, 2008.

Johnson, K.J and Minor, C.P, “Data Fusion Framework for Wide-Area Assessment of Buried Unexploded Ordnance,” Submitted January 22, 2008.

### **NRL Memorandum Reports:**

Rose-Pehrsson, S.L., Johnson, K.J., Minor, C.P., Guthrie, V.N., “Intelligent Data Fusion for Wide-Area Assessment of UXO Contamination. SERDP Project MM-1510. 2006 Annual Report,” NRL Memorandum Report, NRL/MR/6180—07-9039, April 20, 2007.

Rose-Pehrsson, S.L., Johnson, K.J., Minor, C.P., Guthrie, V.N., “Intelligent Data Fusion for Wide-Area Assessment of UXO Contamination. SERDP Project MM-1510. 2007 Annual Report,” NRL Memorandum Report, NRL/MR/6180—08-XXXX, *in press*.