# Implementing Classification on Munitions Response Projects

**Herb Nelson**
Naval Research Laboratory
SERDP/ESTCP

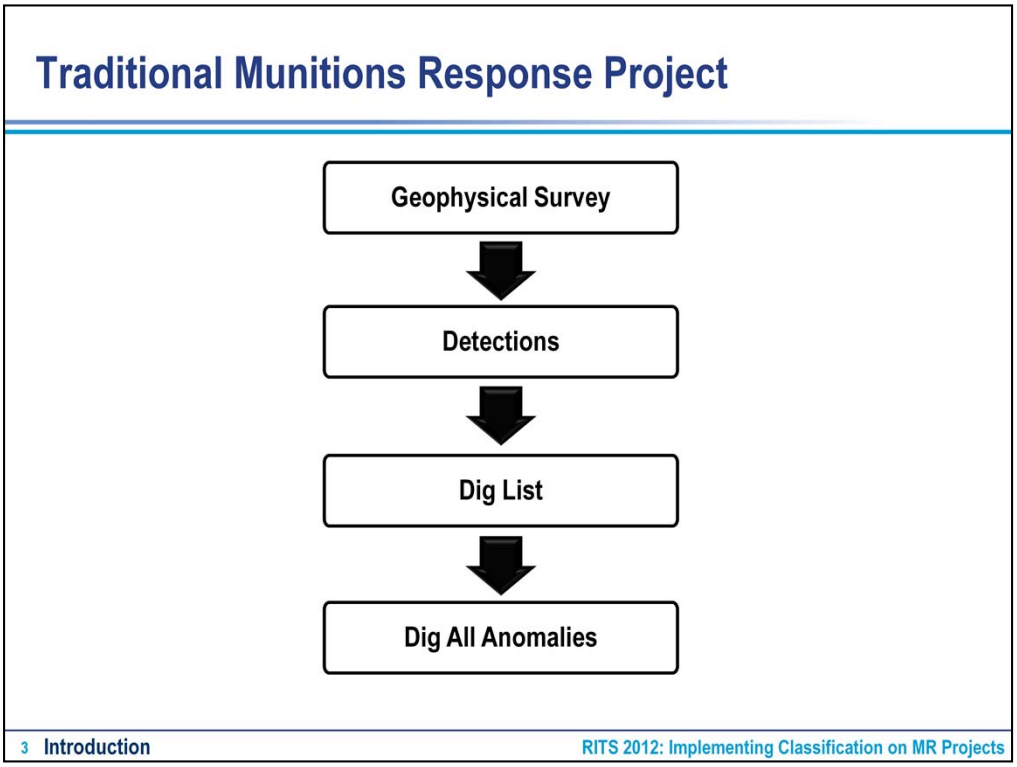
**Bryan Harre**
NAVFAC ESC

**Stacin Martin**
NAVFAC LANT

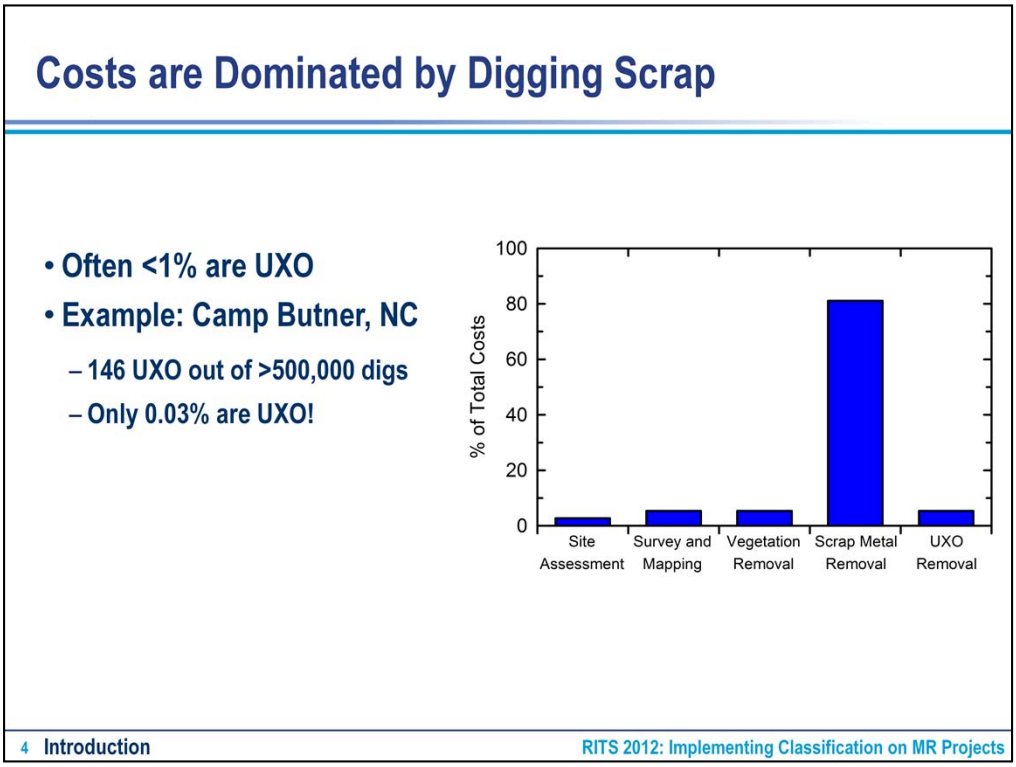
RITS 2012

## Overview of Presentation

- ► Introduction
- Classification Basics
- Workflow and Classification Planning
- Case Study: Pole Mountain, WY
- Summary

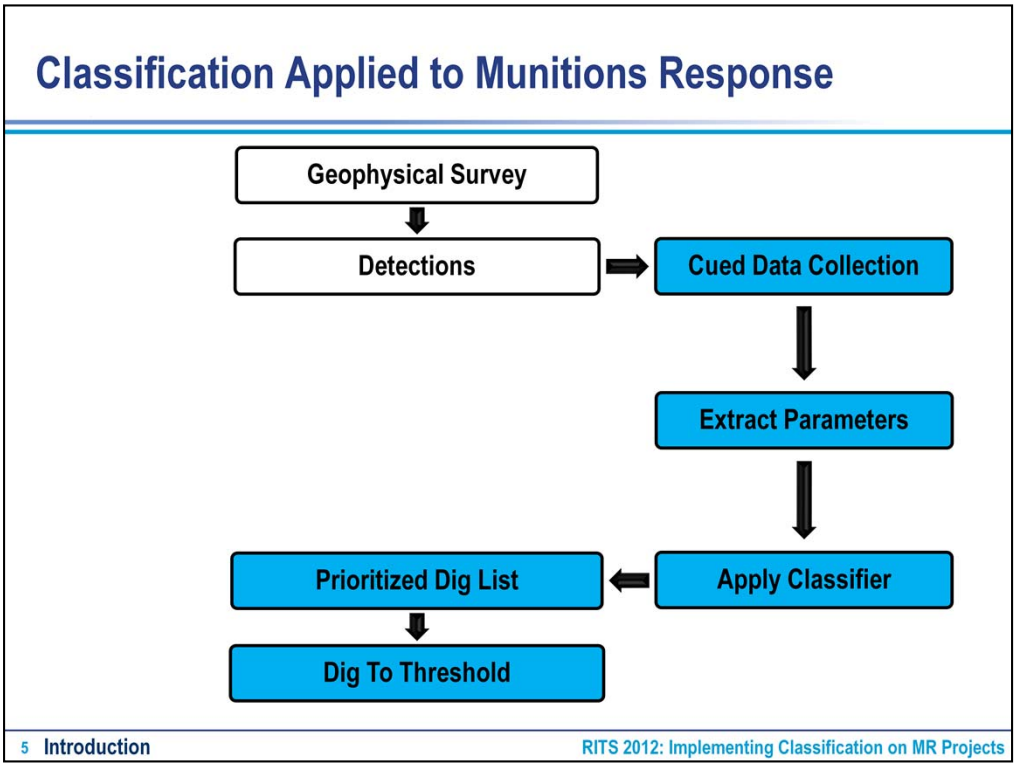



The traditional approach to munitions response includes a geophysical survey and then intrusive investigation of all detected anomalies.




This is a very expensive and inefficient process. The data in the chart are from the 2003 Defense Science Board UXO Task Force report (http://www.denix.osd.mil/mmrp/upload/uxo.pdf). They show that over 70% of the resources on a typical munitions response project are spent removing non-hazardous items from the site. This is re-enforced by the example from former Camp Butner in NC.

Since the cleanup programs have fixed budgets, this inefficiency delays the completion of planned remediations. This is especially true for the FUDS program; the currently forecast completion dates of some of these projects are as late as 2060.




One way to solve this mismatch of resources and acreage to be remediated is to employ classification. In this approach, we add an extra step to the traditional process. We gather additional data about each anomaly that allows us to confidently classify it as arising from harmless scarp and frag or UXO. We then only have to remove the hazardous items allowing the available resources to clean more area.

## Classification

- **Attempts to assign an input value to one of a given set of *classes* based on some *attributes* of the input**
  - **Is the incoming e-mail "spam" or "not-spam"?**
  - **Is the buried object that caused this anomaly "UXO" or "not-UXO"?**



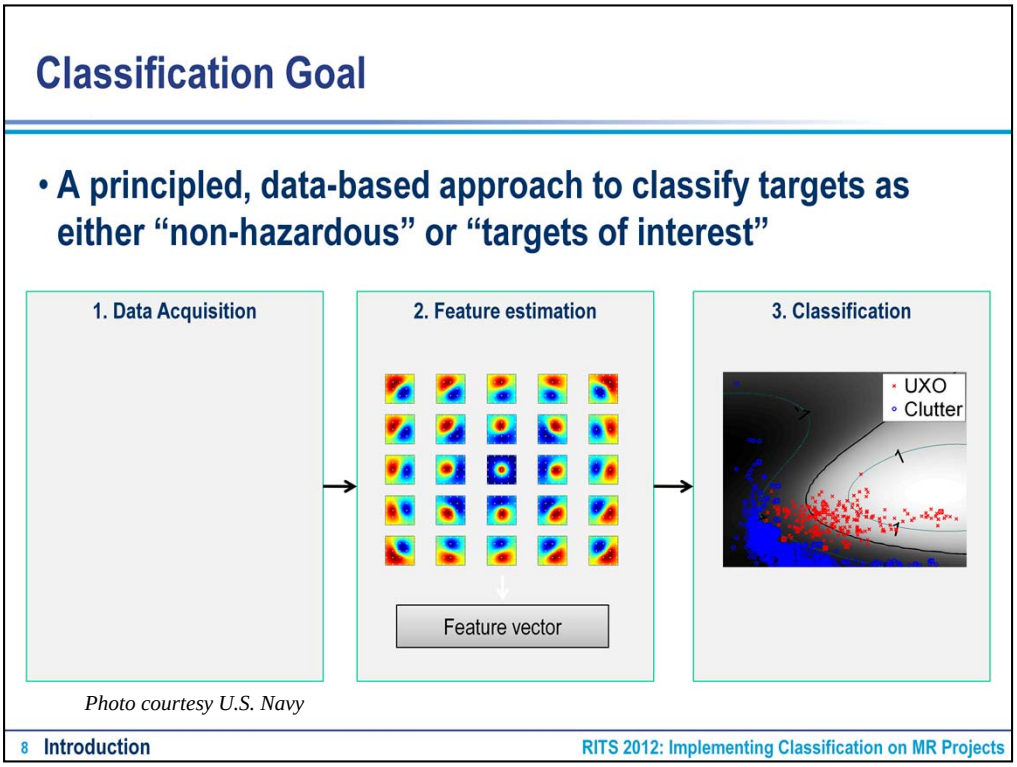
Classification (or discrimination) is the process of using attributes of the anomaly to assign it to the class of UXO or scrap. A good example is your e-mail spam filter which uses attributes of the incoming message to assign the message to your junk mail folder or your in-box.

## How Do We Classify Munitions?

*Photo courtesy ESTCP*

7 Introduction RITS 2012: Implementing Classification on MR Projects

Visually, we classify munitions by attributes such as shape, length-to-diameter aspect ratio, presence of fins or rotating bands, and wall thickness. The objects on the left are clearly munitions and those on the right frag. When the items are buried, we have to use geophysical sensors to extract these very same parameters to allow us to classify the targets.




The standard approach to classification is 1) collect geophysical data over the target, 2) use those data to derive parameters of the buried target (sometimes called features), and 3) use those parameters to classify the buried target as UXO or not.

## History

- Technology Development Objectives
  - Advanced Processing for Commercial Sensors
  - Purpose-built Sensors for Classification
- Testing and Demonstration Timeline
  - Late 1990s: Careful data collection with commercial sensors, mostly improves detection
  - 2000: Demonstrations of crude classification ability with commercial sensors on controlled sites
  - 2000-2001: SEED project to design sensor for UXO classification (BUD)
  - 2005: build and demonstrate BUD on controlled site
    - Successful demo of developmental system
  - 2007-present: Develop and demonstrate transitionable systems, including HW, analysis procedures, SOPs

9 Introduction RITS 2012: Implementing Classification on MR Projects

These technologies have been in development for 20 years. The big breakthrough came in 2005 when the Berkeley UXO Discriminator (BUD) was demonstrated for the first time. The BUD was designed specifically for the UXO classification task and showed what was achievable.

## Why Should a Site Team Consider Classification?

- Classification offers a better understood and documented product
  - Rigorous
  - Tied to Targets of Interest
  - Reviewable at every step

**We can't afford not to use classification**

10 Introduction RITS 2012: Implementing Classification on MR Projects

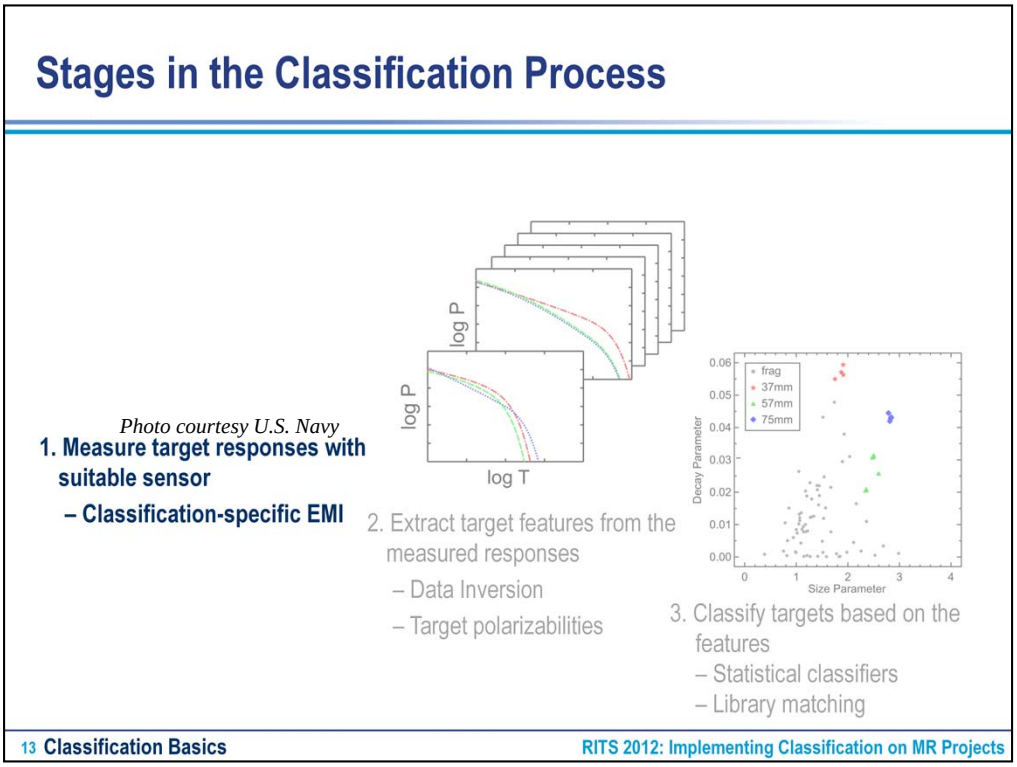
Every site team should consider using classification on their MR projects. The economics are compelling.

## Overview of Presentation

- Introduction
- ► Classification Basics
- Workflow and Classification Planning
- Case Study: Pole Mountain, WY
- Summary

## Outline

- Stages in the classification process
- Sensors
  - Electromagnetic induction (EMI) fundamentals
  - Conventional vs. classification-specific sensor technology
- EMI Response Features
  - Data Inversion
  - Principal axis polarizability
- Classification
  - Statistical classifiers
  - Library/template matching




As above, the first step in classification is to collect data with an appropriate geophysical sensor over the target.

## EMI Sensors

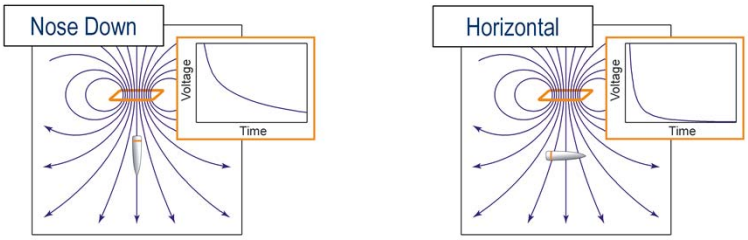
**Electromagnetic Induction (EMI) sensors measure the response of nearby metal objects to magnetic fields created by currents running through a loop of wire**



Classification relies on advanced processing of signals measured using Electromagnetic Induction (EMI) sensors which are just metal detectors. Their basic elements are a transmit coil and a receive coil.

## EMI Signals

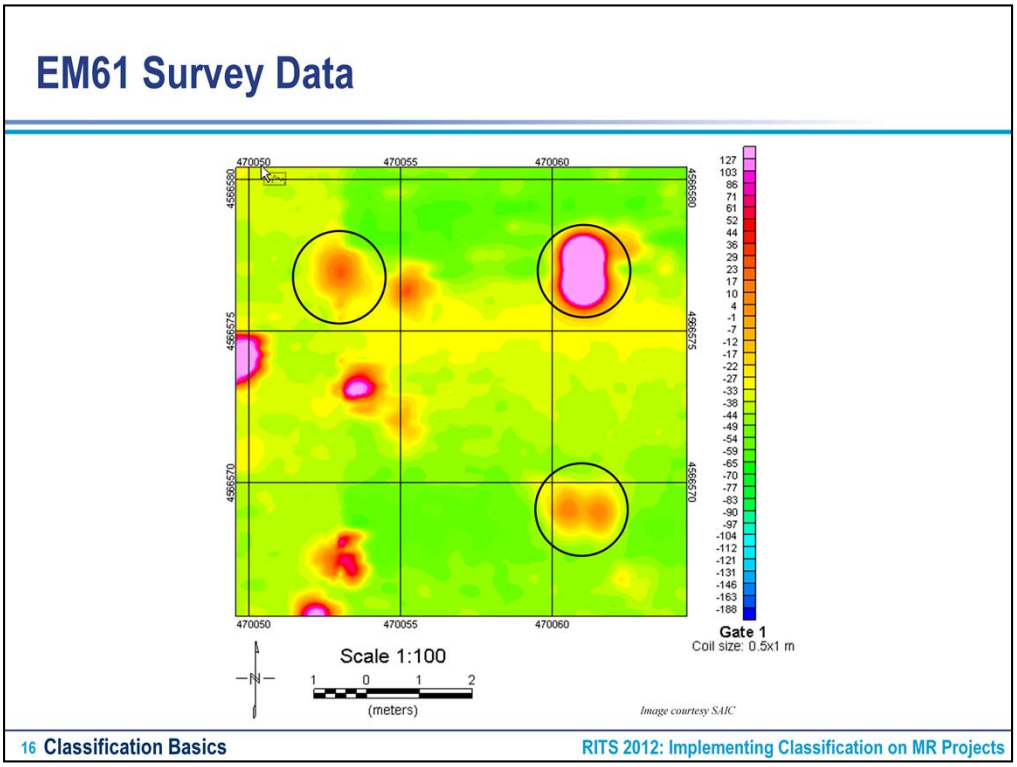
- EMI response signal determined by target properties
  - Size and shape
  - Material type and thickness




- Muddled by response variation with target location and orientation relative to primary field

15 Classification Basics RITS 2012: Implementing Classification on MR Projects

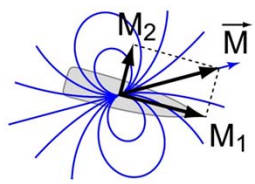
The signal from an EMI sensor is a function of the targets size, shape, and material (all things we want to use for classification) but is muddled by a strong dependence on the location and orientation of the target. A target oriented vertically gives a very different signal from the same object oriented horizontally.




Here is another example of this using EM61-MK2 survey data. The three anomalies that are circled look very different but were all caused by 2.36-in rockets at different orientations and depth.

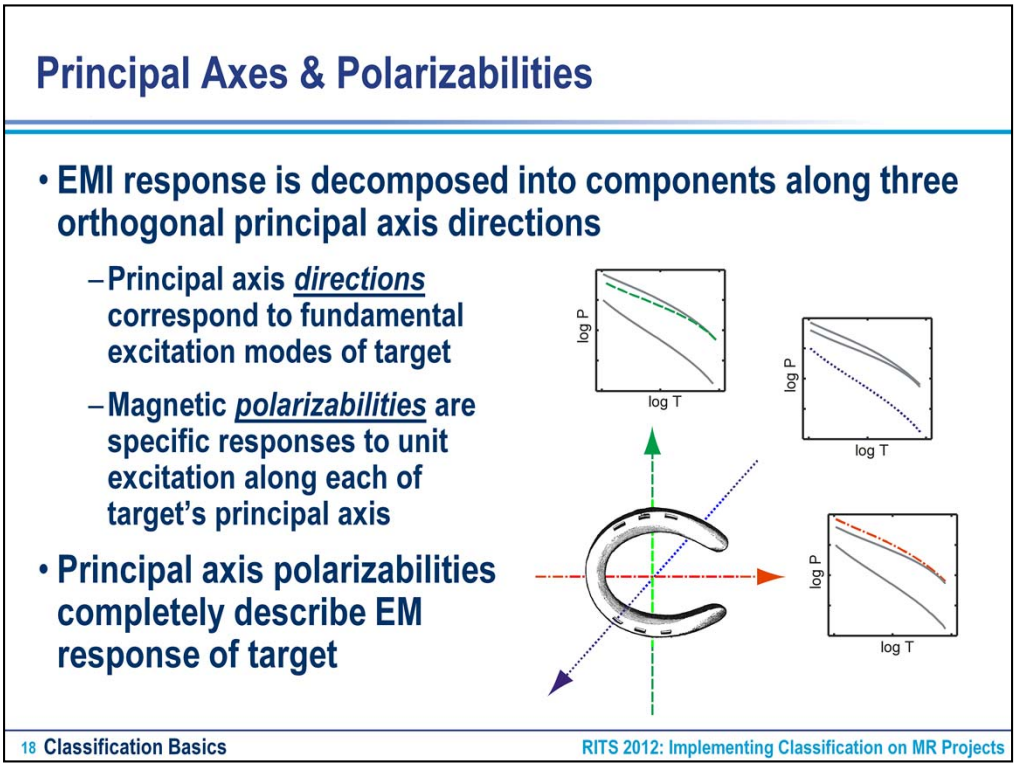
## EMI Response – Dipole Model

- **Observed EMI response is modeled as a function of:**
  - **Induced Magnetic Moment (M) along the three principal axes of the target**
  - **Position of the target relative to the sensor (x, y, z)**
  - **Orientation of the target relative to the sensor (θ, φ, ψ)**




17 Classification Basics RITS 2012: Implementing Classification on MR Projects

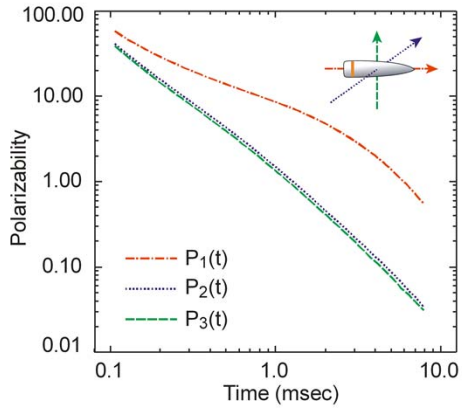
To take care of this, we fit the observed data to an EMI response model (typically a dipole model) that is a function of the intrinsic properties of the target (here called the induced magnetic moment) which is the part we will use to make decisions and the orientation and location of the target which will be useful when we go back to dig it up.




We decompose these induce magnetic moments along the three principal axes of the target. For the horseshoe example shown, there are two relatively large responses in the plane of the slide (up-down and left-right) because the horseshoe is big in those directions and one relatively small response in and out of the plane of the slide because the horseshoe is smaller in that axis. All plate-like objects will look like this – two big responses and one smaller response.

## Polarizability Properties

- Generally simple form
  - Power laws trailing off to exponential decay
- Decay rate depends mainly on material thickness
- Relative strengths for different principal axes indicative of target shape
  - 2 secondary polarizabilities equal for axisymmetric objects
- Magnitude scales with target size (volume)

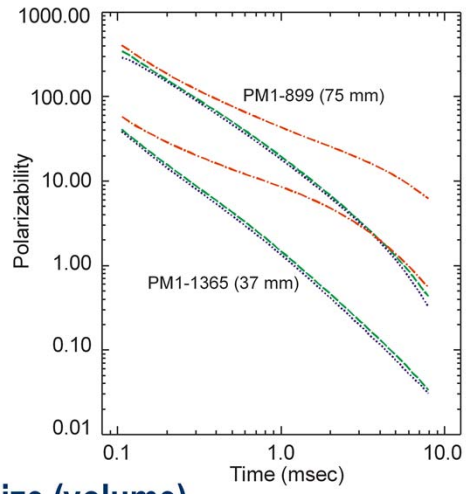



19 Classification Basics RITS 2012: Implementing Classification on MR Projects

By contrast, a cylindrical object (like a munition) will have one large response corresponding to the long axis of the cylinder and two smaller, but equal, responses corresponding to the short axes. The rate of decay of the responses with time is mainly a function of wall thickness of the target. Most munitions are thick-walled so decay slower than scrap.

## Polarizability Properties (cont.)

- Generally simple form
  - Power laws trailing off to exponential decay
- Decay rate depends mainly on material thickness
- Relative strengths for different principal axes indicative of target shape
  - 2 secondary polarizabilities equal for axisymmetric objects
- Magnitude scales with target size (volume)




20 Classification Basics RITS 2012: Implementing Classification on MR Projects

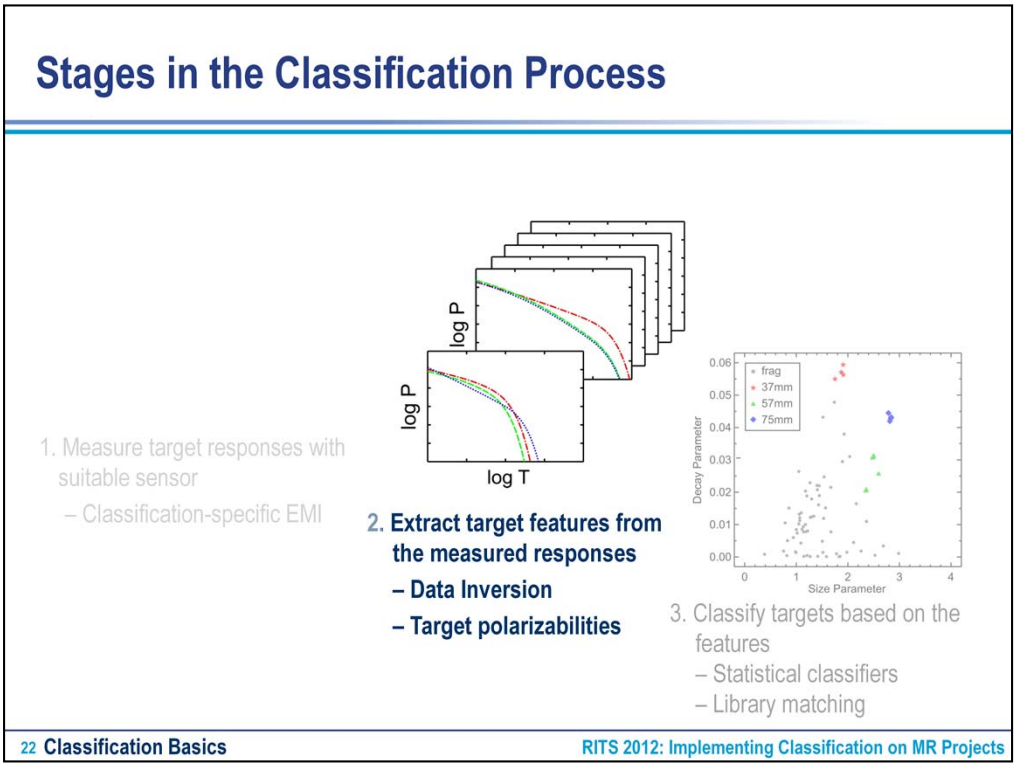
The amplitude of these “principal axis polarizabilities” scales with volume. Here we see that the response of a 75-mm projectile is about ten times larger that a 37-mm projectile.

## Advanced Sensors

- Designed for classification
  - Measure complete decay signal
  - Fixed arrays for precise positioning
  - Multi-axis transmit/receive coils for complete target illumination



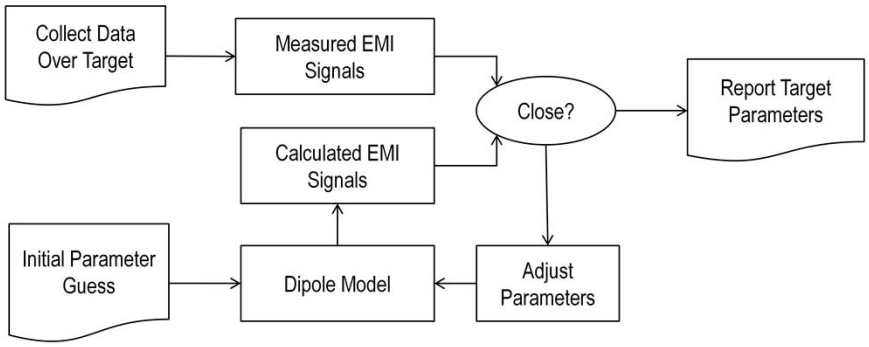
To efficiently acquire the data required for classification we use multi-axis, multi-coil sensors designed for this purpose.




Once we have the required geophysical data, the next step is to extract the target parameters (or features) from those data.

## Conventional Dipole Inversion (Feature Extraction)

- **Determines dipole response model parameters (x,y,z location, θ,φ,ψ orientation and $\beta_1$, $\beta_2$, $\beta_3$ polarizabilities) which produce closest match to measured response**

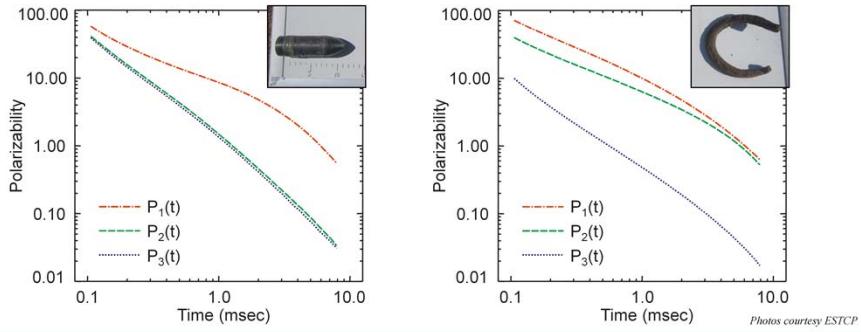



23 Classification Basics RITS 2012: Implementing Classification on MR Projects

Target parameters are estimated using an iterative process called geophysical inversion. An initial guess is made for these parameters, the signal that would result is calculated, and this calculated signal is compared to the measured data. If they do not match the parameters are adjusted slightly and the cycle repeats. After thousands of cycles (easy for a computer) we achieve a good match between the calculated and measured signals and we can be confident we have the target parameters.

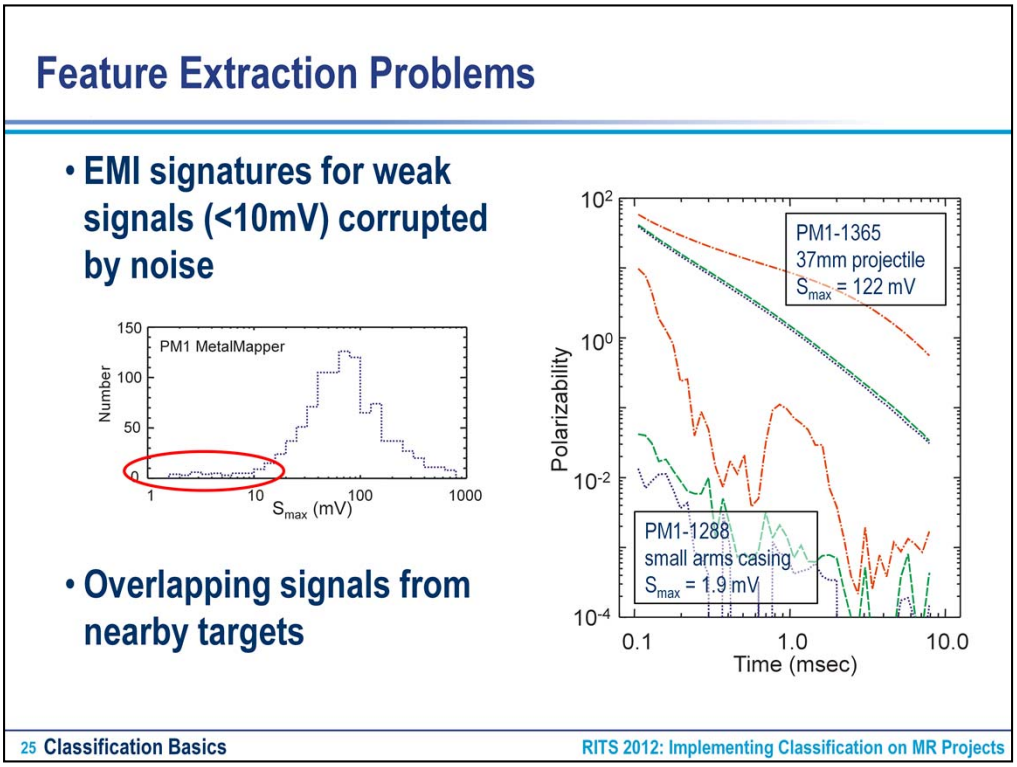
## Principal Axis Polarizability

- **Principal axis polarizability curves completely specify target's EMI response characteristics**
  - **Independent of sensor/geometry**
  - **Contain all information useful for classification**





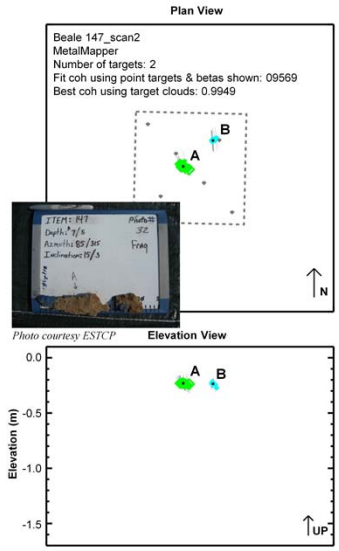

As we saw before, everything we need to know is in these polarizability curves. The 37-mm projectile on the left looks cylindrical and, as we will see later, matches what we expect for a 37mm. The horseshoe on the right looks plate-like – nothing like a munition.




Most times we are able to get a successful inversion but not always. The plot on the left is a histogram of the observed signal amplitudes from a recent demonstration at Pole Mountain. A small number of the measurements had such low amplitude that nothing useful could be extracted – see PM1-1288 on the right. Since we can not say anything confidently about these targets, we have to dig them.

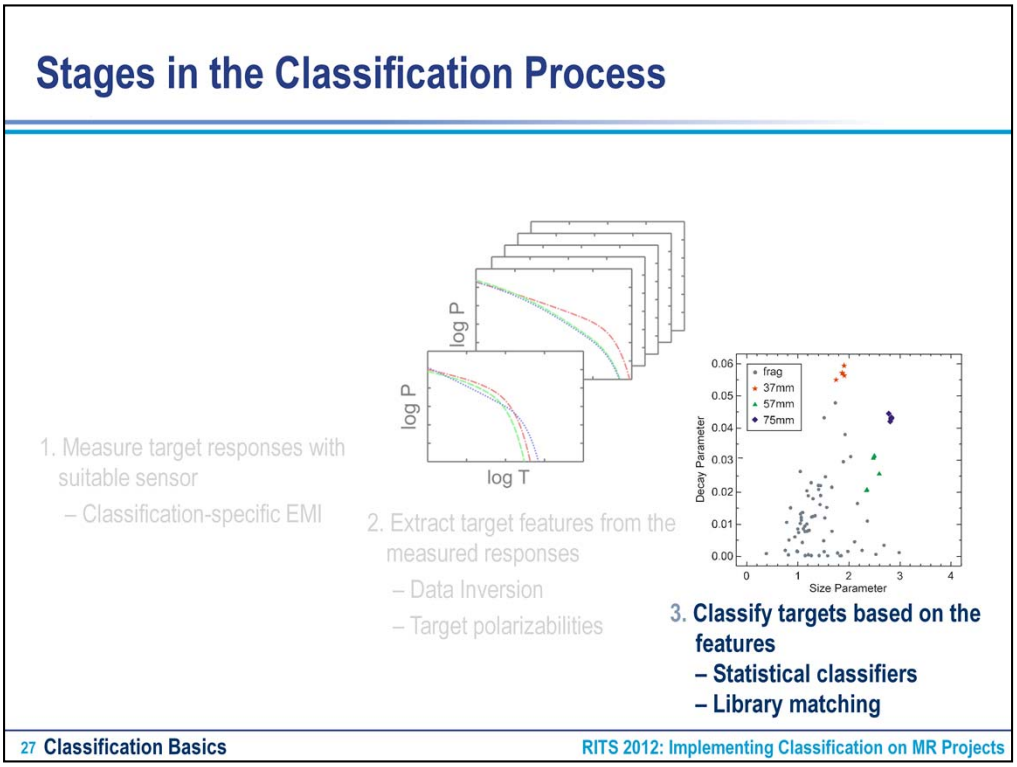
## Multiple/Overlapping Targets

- **Interfering signals from multiple targets in field of view**
  - **Edit out competing signals from small surface clutter**
  - **Use multi-target solvers for larger overlapping targets**
    - Determine where contributions to signal are coming from
    - Perform multi-dipole inversion centered on potential target locations (few targets)
  - **Preemptive dig if anomaly has too many interfering targets**




26 Classification Basics RITS 2012: Implementing Classification on MR Projects

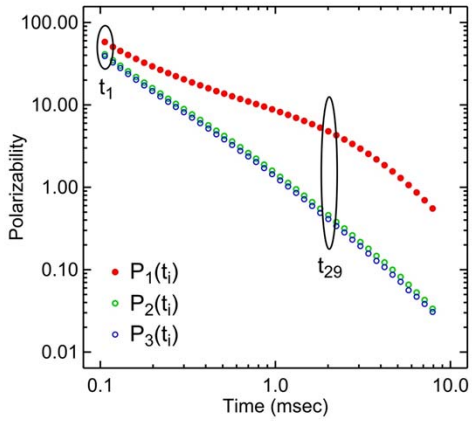
Another possible problem is the presence of more than one object in the field-of-view of the sensor. Two or three years ago this would have been a big problem but there have been enormous advances in the last couple of years. This slide shows an example from former Camp Beale where there were two pieces of frag in the same hole. The analysis had no problem separating them and reporting that they were at the same depth. This also works for fragments buried near a munition. The limit these days is three or four objects under the sensor.




The final stage in classification is to use the parameters we have derived to classify the target.

## Simple Response Features

- Simple properties of principal axis polarizabilities can distinguish between some types of objects
- Large vs. small
  - Overall strength of response at early time (~0.1 ms) scales with size of object
- Thin- vs. thick-walled
  - Strength of response at later time (~2 ms) relative to early time scales with thickness of material






Trying to get your hands around the full polarizability curves (3 curves x 42 time gates each) is sometimes a little difficult. We often would like to use a lower-dimensionality representation for convenience. One way that you often see is to derive something related to the size of the target from the amplitude of the polarizabilities and something related to the wall thickness from the decay.

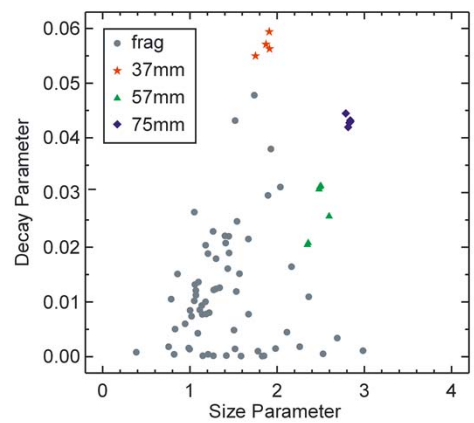
## Simple Feature Spaces

- Simple two dimensional "feature spaces" like size/decay can be useful for visualizing data and identifying clusters of similar objects
  - "Feature vector" is a point (size, decay) in the feature space

$$Size = \log\sqrt{\{p_1^2(t_1) + p_2^2(t_1) + p_3^2(t_1)\}}$$

$$Decay = \frac{\sqrt{\{p_1^2(t_{29}) + p_2^2(t_{29}) + p_3^2(t_{29})\}}}{\sqrt{\{p_1^2(t_1) + p_2^2(t_1) + p_3^2(t_1)\}}}$$

$t_1 \sim 0.1$ msec, $t_{29} \sim 2$ msec




29 Classification Basics RITS 2012: Implementing Classification on MR Projects

Here we have constructed a "feature space" using these two simple parameters or features. On a plot like this, we expect munitions to be toward the upper right since they are larger than most fragments and have thicker walls (decay slower) than most of the scrap.

The Feature Space plot is analogous to an Executive Summary in a 126-page report.

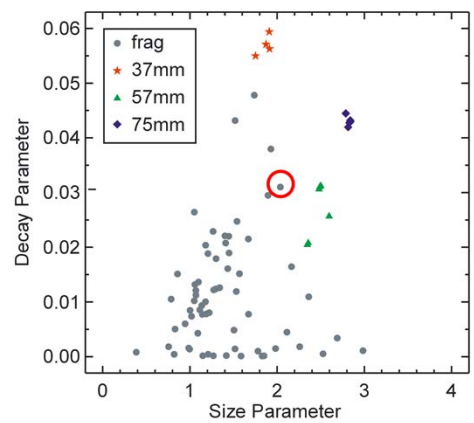
## Simple Feature Spaces (cont.)

- **Simple two dimensional "feature spaces" like size/decay can be useful for visualizing data and identifying clusters of similar objects**
  - **"Feature vector" is a point (size, decay) in the feature space – *e.g., (2, 0.03)***

$$Size = \log\sqrt{\{p_1^2(t_1) + p_2^2(t_1) + p_3^2(t_1)\}}$$

$$Decay = \frac{\sqrt{\{p_1^2(t_{29}) + p_2^2(t_{29}) + p_3^2(t_{29})\}}}{\sqrt{\{p_1^2(t_1) + p_2^2(t_1) + p_3^2(t_1)\}}}$$

$t_1 \sim 0.1$ *msec*, $t_{29} \sim 2$ *msec*

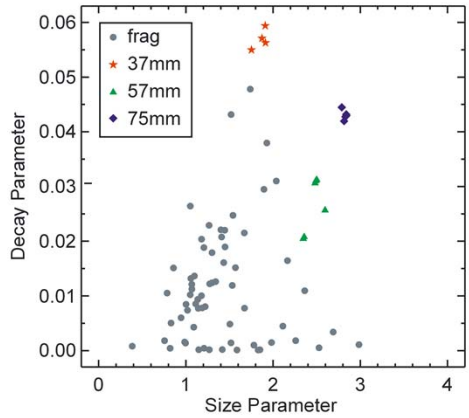



30 Classification Basics RITS 2012: Implementing Classification on MR Projects

Each measurement we make results in a point on this plot. In this 2-dimentsional space, the "feature vector" for each target is its size,decay pair.

## Simple Feature Spaces (cont.)

- Simple two dimensional "feature spaces" like size/decay can be useful for visualizing data and identifying clusters of similar objects
  - "Feature vector" is a point (size, decay) in the feature space
- Limited classification power relative to complete set of polarizability curves




31 Classification Basics RITS 2012: Implementing Classification on MR Projects

When people try to do classification using an older sensor like the EM61-MK2 they can only use simple features like those shown on this plot. With the data from the modern sensors, we have much more information to work with (3 curves x 42 gates each) and only use these plots as a convenient way to look at the data.

## Classification Techniques

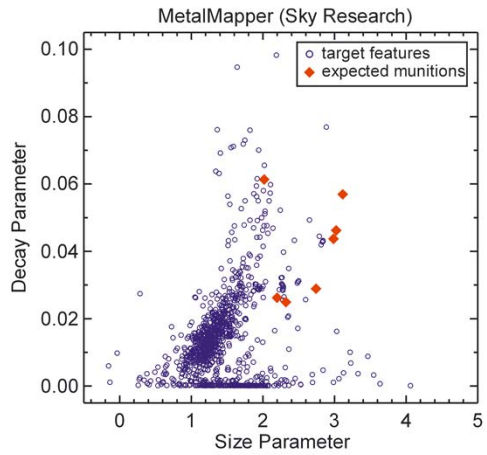
- **Statistical Classifiers**
  - Input features include all $3_{axes} x N_{gates}$ polarizabilities
  - Machine learning – support vector machines, neural nets, etc.
  - Trained on prior target information and labeled training data
- **Library Matching**
  - Asks what an unknown target "looks like" in EMI sense
  - Compares polarizability against bank of signatures for expected munitions and other training objects
- **Both approaches are based on signal matching**
  - Statistical classifiers create their own library
  - Both can have problems with unexpected munitions types

32 Classification Basics RITS 2012: Implementing Classification on MR Projects

There are a number of ways to use the polarizabilities to classify. The two most common in munitions work are statistical classifiers and library matching methods. Both use the polarizability curves to find matches.

## Statistical Classification Approach

- **Locate expected munitions item signatures in feature space**

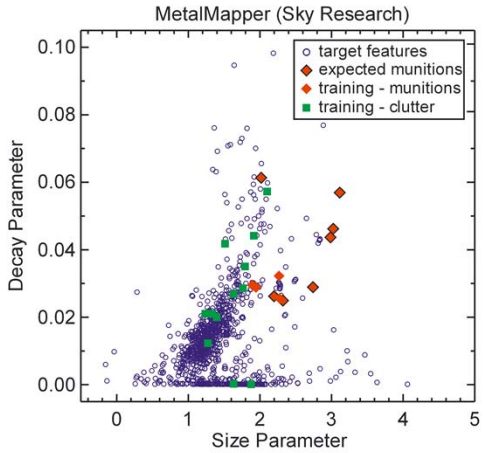



33 Classification Basics RITS 2012: Implementing Classification on MR Projects

A statistical classifier starts by teaching the computer where known munitions are located in feature space. Here we have plotted the time decay points for the munitions we measured in the Instrument Verification Strip (IVS) and a small test pit we had on site. As expected, the munitions are in the large, slow decay quadrant of the plot.

## Statistical Classification Approach (cont.)

- Locate expected munitions item signatures in feature space
- Sample feature space (dig targets & identify/label)
  - Regions around munitions
  - Target clusters

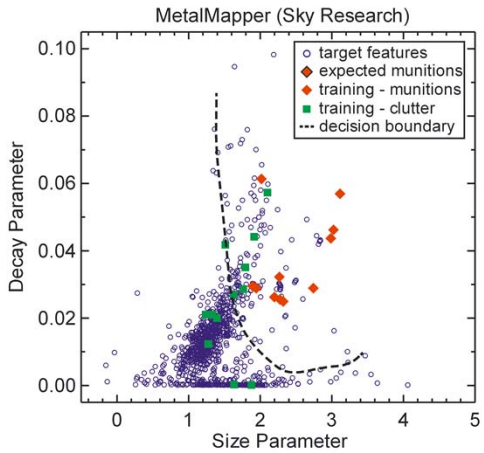



34 Classification Basics RITS 2012: Implementing Classification on MR Projects

Our goal is to have the classifier divide the space into munitions and not-munitions. So we request the identity of some other points in this space to train the classifier.

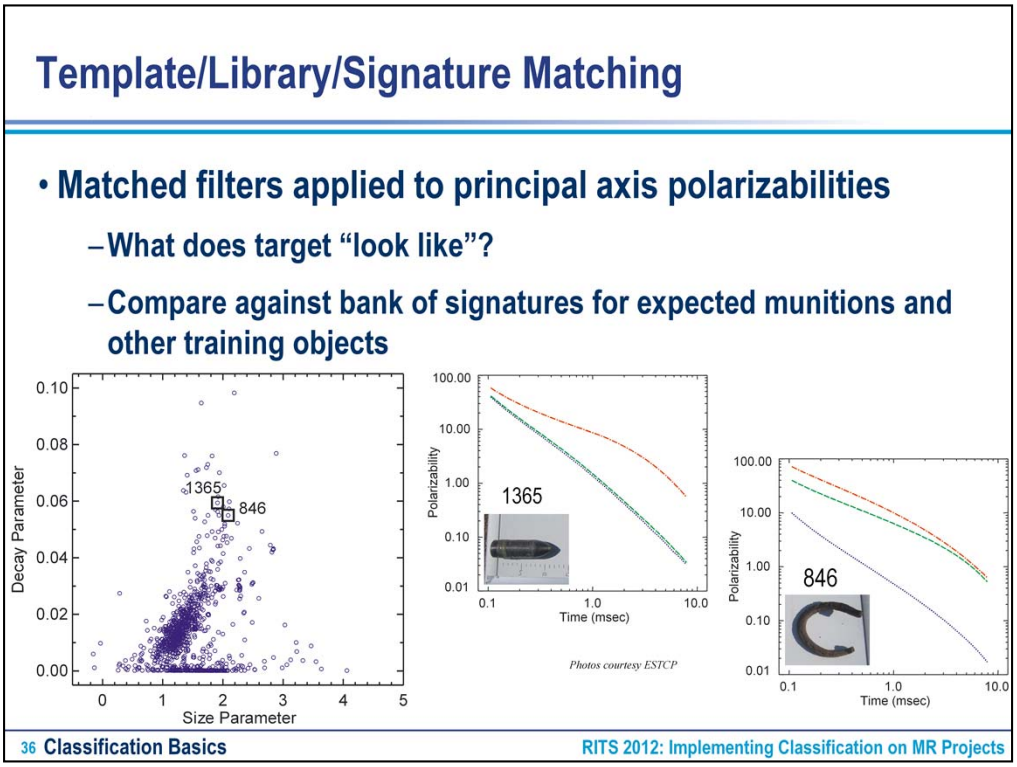
## Statistical Classification Approach (cont.)

- Locate expected munitions item signatures in feature space
- Sample feature space (dig targets & identify/label)
  - Regions around munitions
  - Target clusters
- Train classifier with labeled features
  - Set decision boundary to exclude high confidence clutter




35 Classification Basics RITS 2012: Implementing Classification on MR Projects

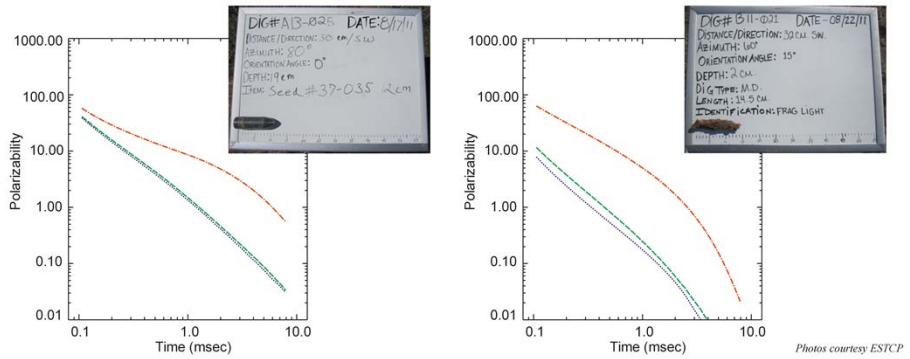
After training, the classifier can draw a boundary separating the munitions region from the not-munitions region. Note that this is not very precise in this two-dimensional space; we'll see later how easy this becomes when we take advantage of all the information the advanced sensors provide.




Another approach is called library matching. Here we have a library of munitions signatures and we compare each unknown item to the library. Anything that matches we call a munition, anything that does not, we call clutter.

## Template/Library/Signature Matching (cont.)

- How much alike do a pair of EMI signatures look?
- Frag looks somewhat different than 37mm




*Photos courtesy ESTCP*

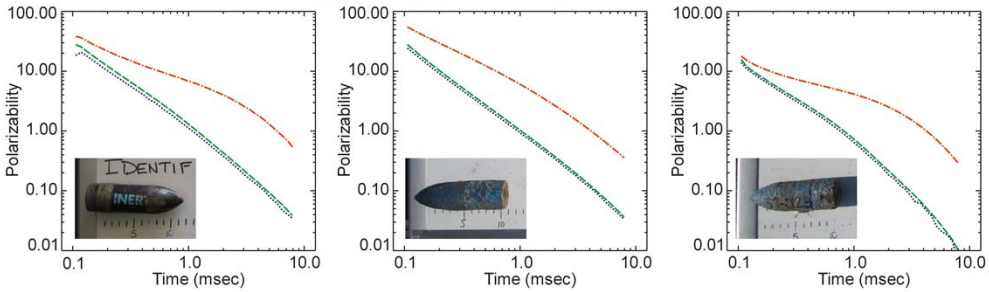
37 **Classification Basics** RITS 2012: Implementing Classification on MR Projects

Calling a set of polarizabilitiy curves a match to the library involves defining what a match means. The fragment on the right looks something like the 37mm on the left (roughly cylindrical, about the same size) but differs in detail. These two are not a match.

## Signature Variability

- EMI signatures of nominally identical items can differ
  - Different subtypes, damage, inversion errors due to noise
- Matching procedures must tolerate some variability

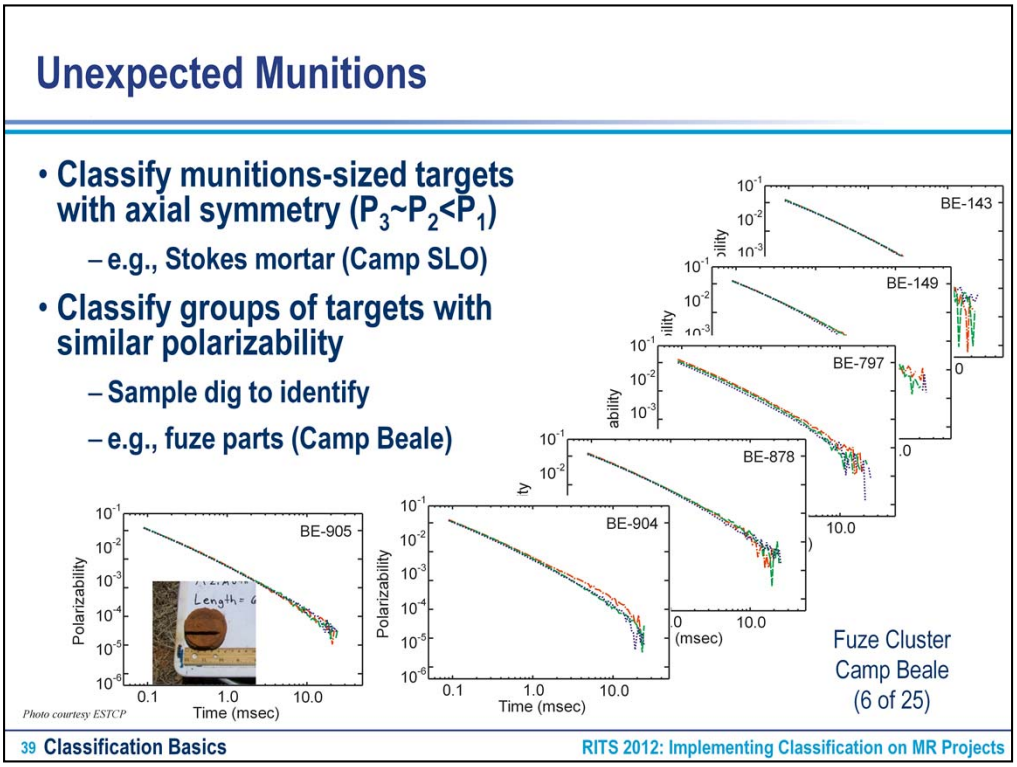
e.g., 37mm projectile rotating band effects




*Photos courtesy ESTCP*

38 Classification Basics RITS 2012: Implementing Classification on MR Projects

All three of these curves correspond to different versions of a 37-mm projectile. We have all of these in the library so we won't miss any of these items.




Using library match methods, there is always the worry that we will encounter something on our site that is not in the library. At the former Camp Beale demonstration, a number of items were encountered that all looked alike but did not match anything in the library. The analyst requested that one of them be excavated to see what they were. In this case, they were expended fuzes that the site team decided were safe to leave in the ground. This provides a good example of how clusters (or groups) of unexpected items are handled.

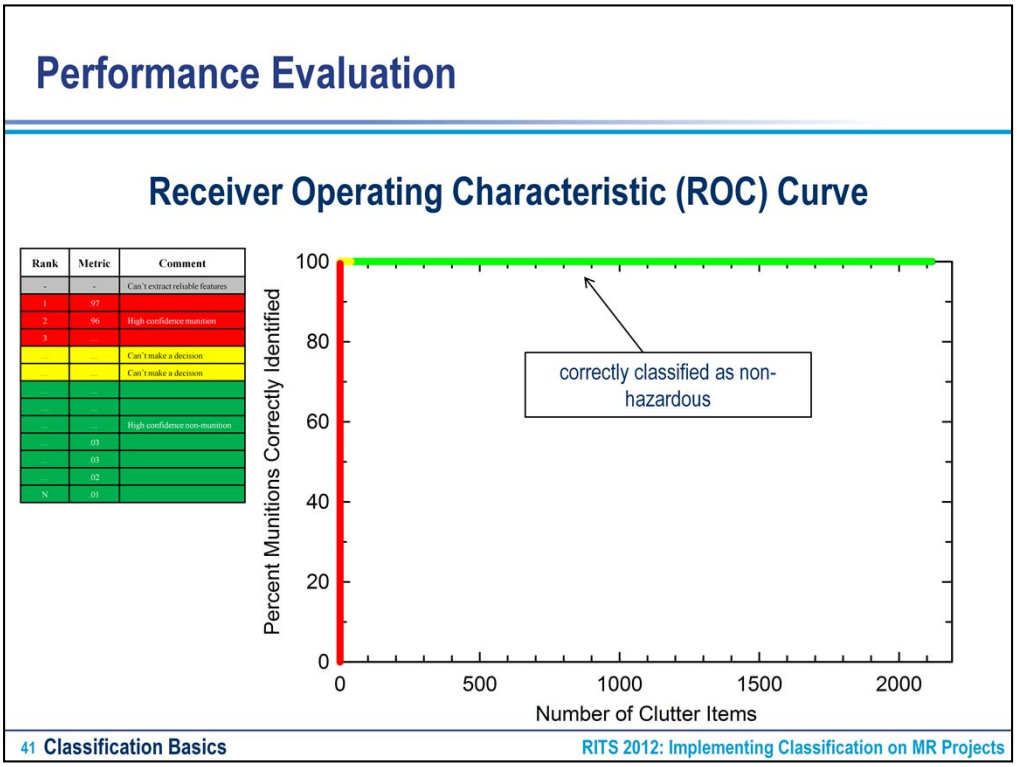
## Classifier Output

- Ranked anomaly list
  - Likely munitions
  - Likely clutter
  - Can't tell
- "Stop dig" threshold →
  - Set to exclude only high confidence clutter
  - Dig all uncertain targets (and likely munitions)

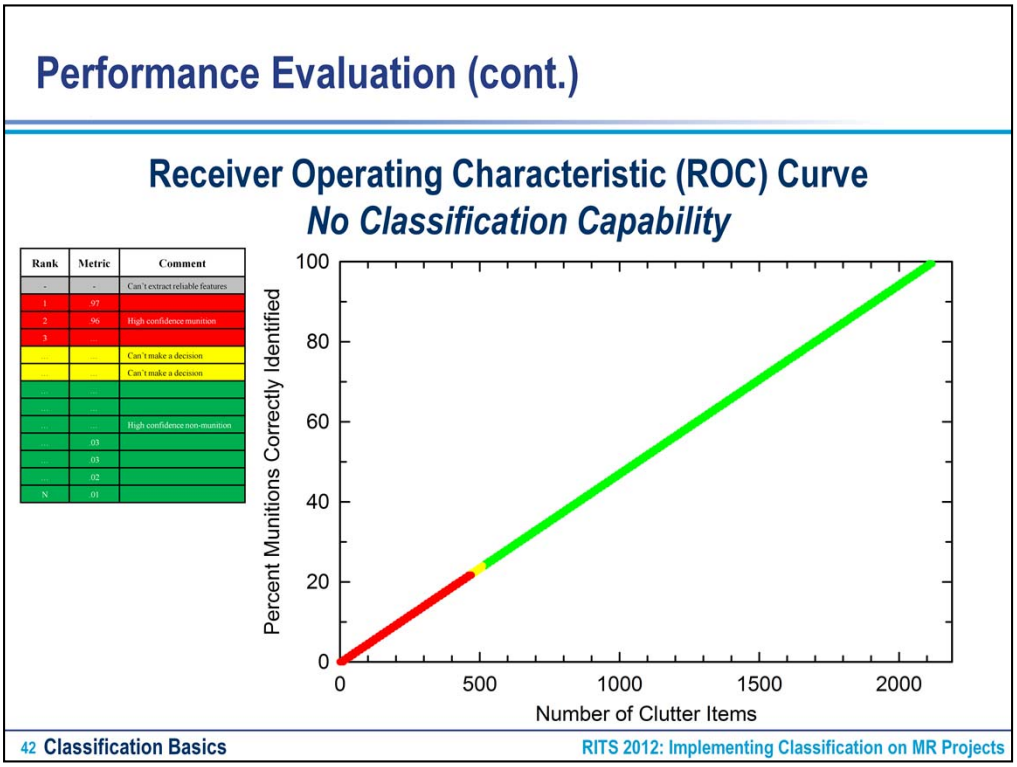
| Rank | Decision Metric | Comment |
|---|---|---|
| -9999 | -9999 | Can't extract reliable features |
| 1 | .97 | |
| 2 | .96 | High confidence munition |
| 3 | … | |
| … | … | Can't make a decision |
| … | … | Can't make a decision |
| … | … | |
| … | … | |
| … | … | High confidence non-munition |
| … | .03 | |
| … | .03 | |
| … | .02 | |
| N | .01 | |

40 Classification Basics RITS 2012: Implementing Classification on MR Projects

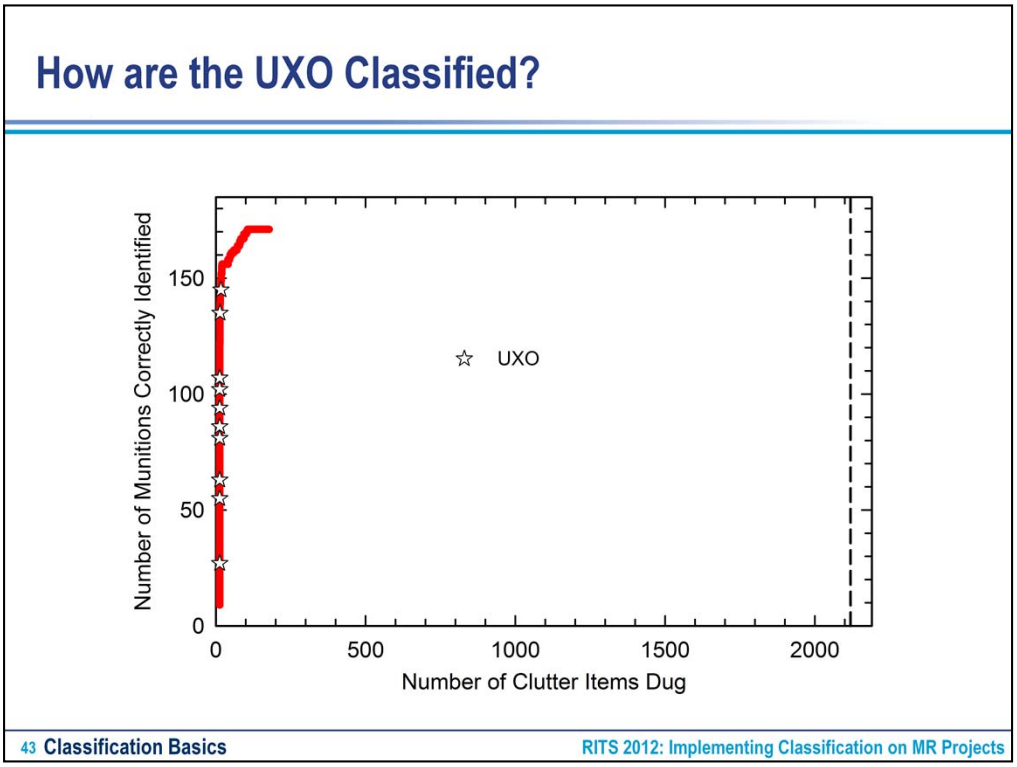
The final output of the classifier is a ranked anomaly list constructed as shown here. There are always a few targets for which the data are bad or there was some analysis problem. Since we can say nothing useful about these targets, they have to be dug. We show them here in gray. Next come the items we are confident are munitions (remember the 37-mm projectiles from a few slides back). These also must be dug. There may or may not be some targets that we can't decide about based on our current information n (the fuzes from the last slide are a good example of this). The only things we can leave in the ground are the targets we are confident are not munitions. These are shown in green on the list.




Once we have the list and have excavated all the targets (which we do as part of our demonstrations), we can construct a ROC curve. We build this curve by digging each item on the list in turn. If the item is a munition we go up on the plot, if it is clutter we go to the right. The example here is a near-perfect ROC. All the things we called high-confidence munitions (colored red like the anomaly list) were, in fact, munitions. Everything we weren't sure about or called high-confidence clutter was clutter.




This ROC curve results from a method with no classification ability. The things called high-confidence munitions (red points) has some chance of being munitions and some chance of being clutter but the things called high-confidence clutter (green points) had the same chance of being munitions. This performance is sometimes called the chance diagonal. If our results look like this, we are a failure.




One other check we might do is see where any existing UXO (other than our QC seeds) appear on the ranked anomaly list. On this slide we have reproduced the portion of the ROC curve through the items classified as high-confidence munitions. The eleven UXO found on this site are marked with stars. This is a good result; the UXO are neither easier nor harder to find than the QC seeds. We chose the seeds well and can have confidence that we found all the UXO.

## Classification Basics – Summary

- Modern sensors accurately measure EMI response information needed for reliable classification
- Analysis procedures remove sensor/geometry effects to extract target's intrinsic EMI signature
  - Depends only on size, shape and material properties of target
- Statistical and library-based classifiers can reliably distinguish between munitions and clutter items
  - Match unknown targets with other objects with similar EMI signatures (i.e., things they "look like")
  - Presumptive UXO (dig) if EMI "vision" is fuzzy or obscured

## Overview of Presentation

- Introduction
- Classification Basics
- ► Workflow and Classification Planning
- Case Study: Pole Mountain, WY
- Summary

45 RITS 2012: Implementing Classification on MR Projects

We are now going to discuss how classification can be implemented on a MRS, what some of the changes in process are, and what some of the planning and QC considerations will include.

## MR Project Work Elements that Classification Significantly Changes

- GIS setup
- Document management and control
- Subcontracting
- Technical and operational approach
- Work Plan preparation and approval
- Site prep and mobilization
- Site survey/grid layout
- Vegetation removal
- Surface removal
- Geophysical System Verification (GSV)
- Geophysical survey, data collection, and processing
- Anomaly reacquisition and investigation
- MEC/MPPEH management
- Demobilization
- Final report
- Archiving
- Project closeout

46 Workflow and Classification Planning RITS 2012: Implementing Classification on MR Projects

This slide summarizes all the MR project work elements for a typical MRS. The elements classification significantly changes are in blue and underlined.

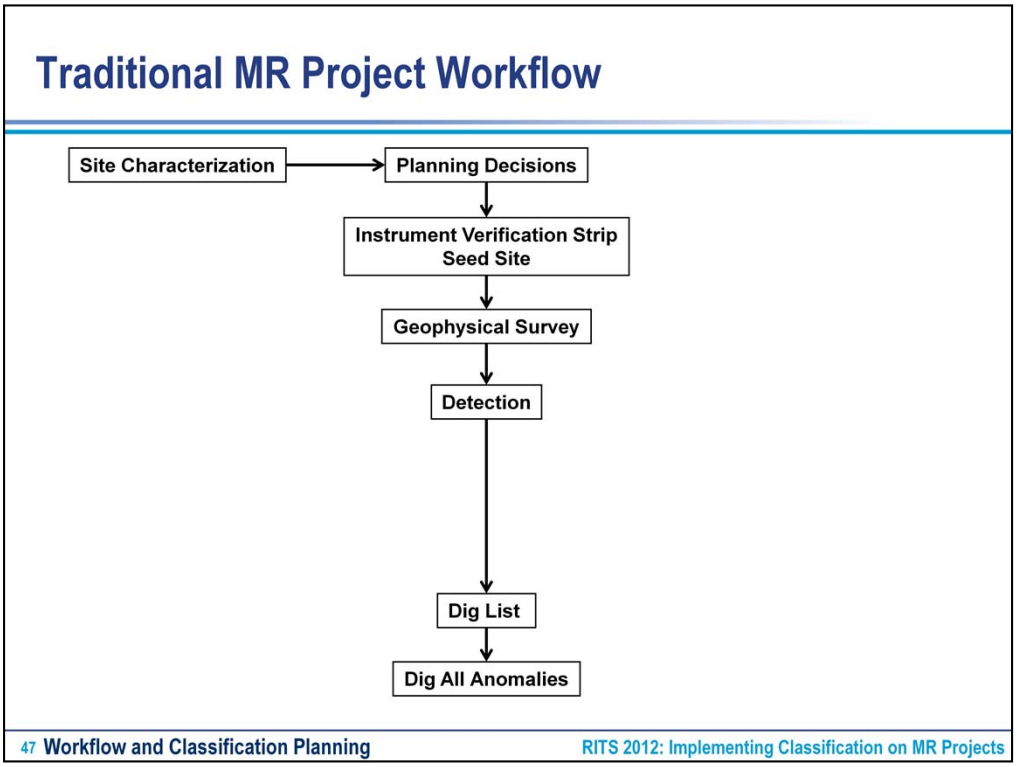
1) The GSV QC products are different

2) The survey is conducted twice

3) Additional data processing and analysis is required

4) Only a small fraction of anomalies will be investigated

5) The UXO techs will encounter MEC at much higher frequency than ever before, but the number of digs will be much smaller. This will impact donor explosives delivery/handling.




Here are the basic steps/workflow for addressing a typical MRS site with today's standard DGM equipment.

You start with site characterization data from a site inspection or a remedial investigation. This data is then used to develop plans for the site.

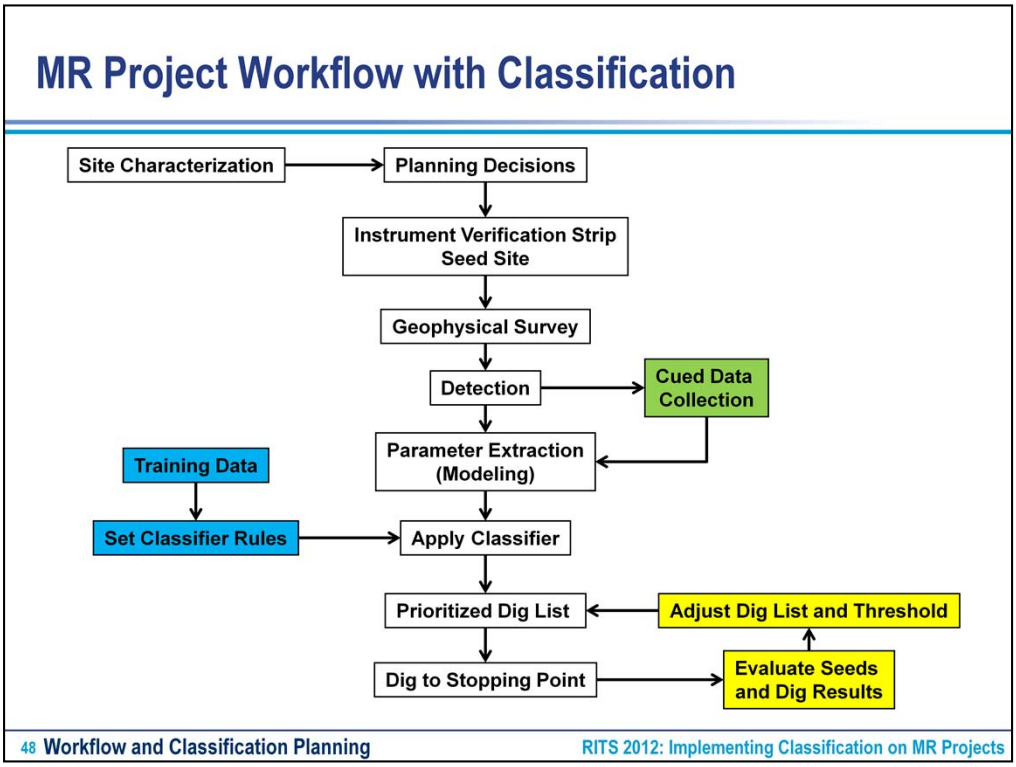
Planning considerations are developed on Navy workplans thru the project quality objective 7 step process. These decisions and considerations are documented in a QAPP.

The IVS and seeding of the site are the primary QC tools used to evaluate how well a contractor is performing

The geophysical survey is completed w/DGM equipment

The detections are then rudimentarily "classified" based upon an agreed upon anomaly selection criteria (mV threshold, etc)

All the anomalies meeting the criteria are then sent to the UXO techs and dug




Here are the basic steps/workflow for addressing a typical MRS site with the advanced sensors capable of performing classification.

You start with site characterization data from a site inspection or a remedial investigation. This data is then used to develop plans for the site.

Planning considerations are developed on Navy workplans thru the project quality objective 7 step process but now must include QA and QC considerations for the advanced sensors. These decisions and considerations are documented in a QAPP.

The IVS and seeding of the site are the primary QC tools used to evaluate how well a contractor is performing and as mentioned before are different because such things as polarizibility curves are used to evaluate the contractor.

The geophysical survey is completed w/DGM equipment and currently a second survey is performed on the selected anomalies (two-pass system). This allows for the collection of the cued data, which is basically the richer data needed to classify.

Parameters are extracted that will help classify an item in a comparison of the actual data to modeled data.

Training data can be used to develop the rules for classifying an item as MEC or not

The classifier is then applied to the dataset. After that the process enters a feedback loop (results, additional training data, etc) until the evaluation of the seeds/dig results meets project quality standards.

Again, only a small number of anomalies meeting the classifier criteria are sent to the UXO techs to be dug.

## Is Classification Applicable at this Site?

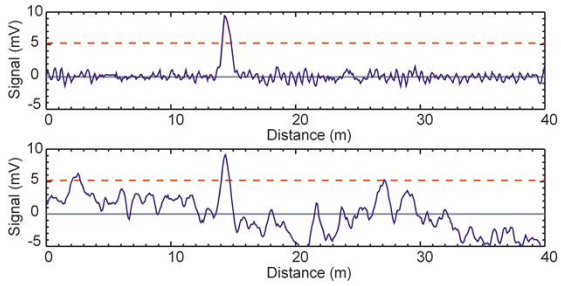
- What are the targets of interest at this site?
  - Historical research
  - Recovered munitions and fragments/scrap
  - Depth and density

49 Workflow and Classification Planning RITS 2012: Implementing Classification on MR Projects

One question you maybe asking is “Is classification applicable at my site?” To answer that question you need to first determine the targets of interest at your site. Discuss the factors on the slide that help you determine the Targets of Interest (TOIs). Note that items buried too deeply for today’s traditional sensors also cannot be classified and neither can the “purple blob” of high density metal sites be classified.

## Is Classification Applicable at this Site? (cont.)

- **What is the appropriate threshold for detection?**
  - **Clutter environment and geology affect detection threshold and the ability to classify**
  - **If project objectives require "picking into the noise," you need a new plan**




**Good**

**Bad**



You also need to determine the appropriate threshold for detection, which is affected by clutter and geology. The two graphs illustrate a good and bad place to attempt to classify an anomaly. Point out the variability in the noise/instrument response and the clearly defined peaks and the dubious peaks.

## Classification Planning

- What classification sensor is appropriate for the site?
  - Benign terrain
    - MetalMapper
  - Rougher terrain with steeper slopes and trees
    - TEMTADS 2x2, MPV, Handheld BUD



Once you have decided to use an advanced sensor you need to determine which one. Different MRS's have different requirements for the type and size of advanced sensors. This is primarily a function of terrain and vegetation. MetalMapper is large and needs large equipment to move it. The smaller, more man-portable sensors can work under trees and in more demanding terrain.

## Classification Planning (cont.)

- **What parts of the site have an anomaly density that will allow classification to be successful and at what cost?**
  - **Small sites and densities greater than 1,000 per acre may not be appropriate**
  - **For sites with few anomalies, the costs of remediation must justify the extra expense required for data collection**
    - Wetlands, chemical sites, etc.

**Not Appropriate**

**Landfill Surprise**

**Appropriate**

**40-mm Anti-Aircraft Projectile**

*Photos courtesy U.S. Navy*

52 **Workflow and Classification Planning** RITS 2012: Implementing Classification on MR Projects

Another consideration in planning to use classification at a site includes anomaly density and cost. There maybe portions of the MRS that can perform classification and portions that cannot. The two sub bullets highlight some considerations based upon size of site, anomaly density and the reality that current systems are two-pass. The pictures illustrate some extreme examples where classification would not be appropriate and where it would be totally appropriate.

## IVS and QC Seeds

- What are the appropriate seed items and how deeply should they be buried?
  - Surrogates
  - ISOs similar to known munitions
  - Surrogates of unknown munitions
- How many production area seeds?
  - Encountered daily
- How often should the IVS and noise strip be surveyed?
  - Twice daily

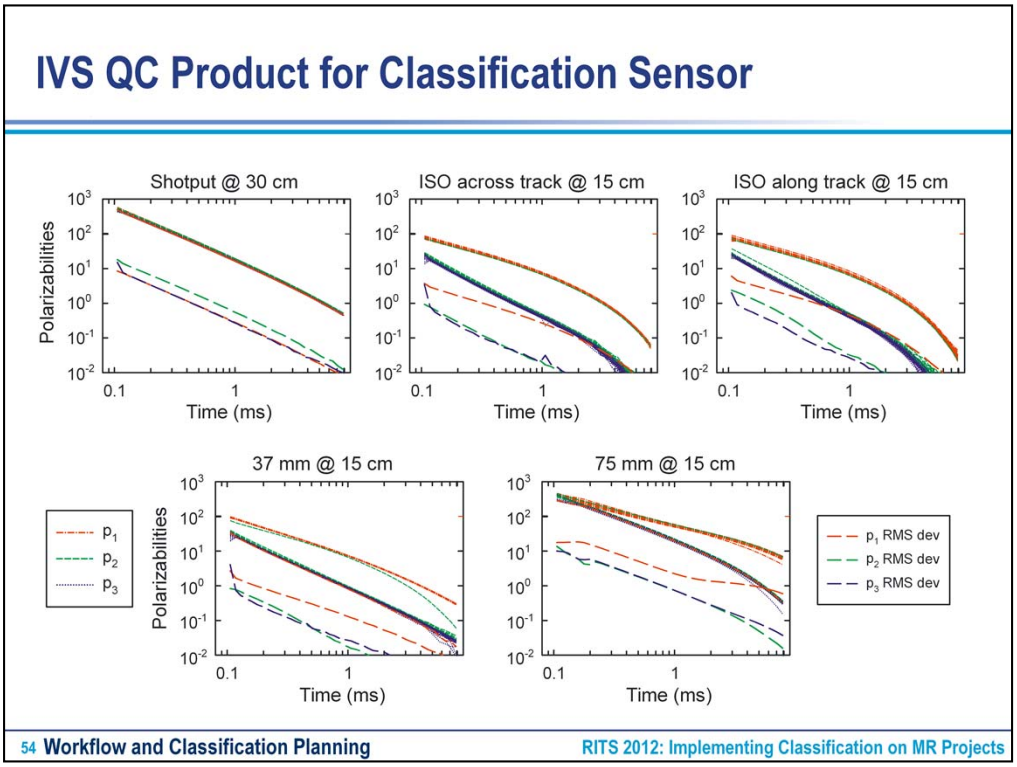
Blind seed bagged and buried

*Photo courtesy of ITRC*

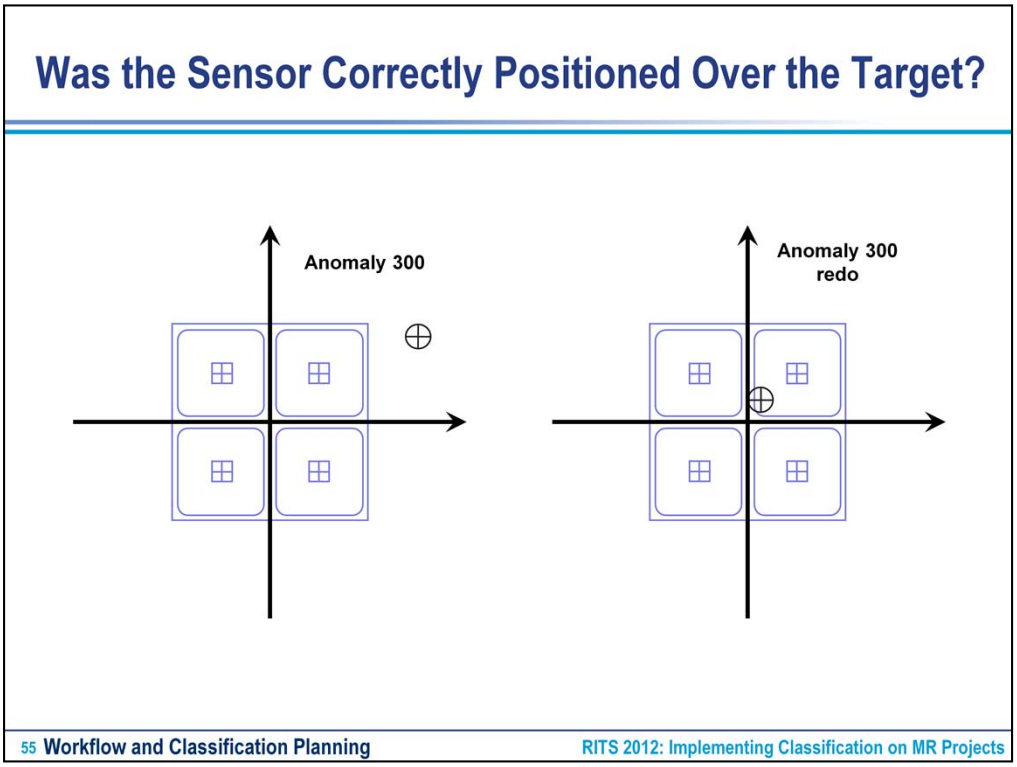
*Photos courtesy ESTCP*

53 Workflow and Classification Planning RITS 2012: Implementing Classification on MR Projects

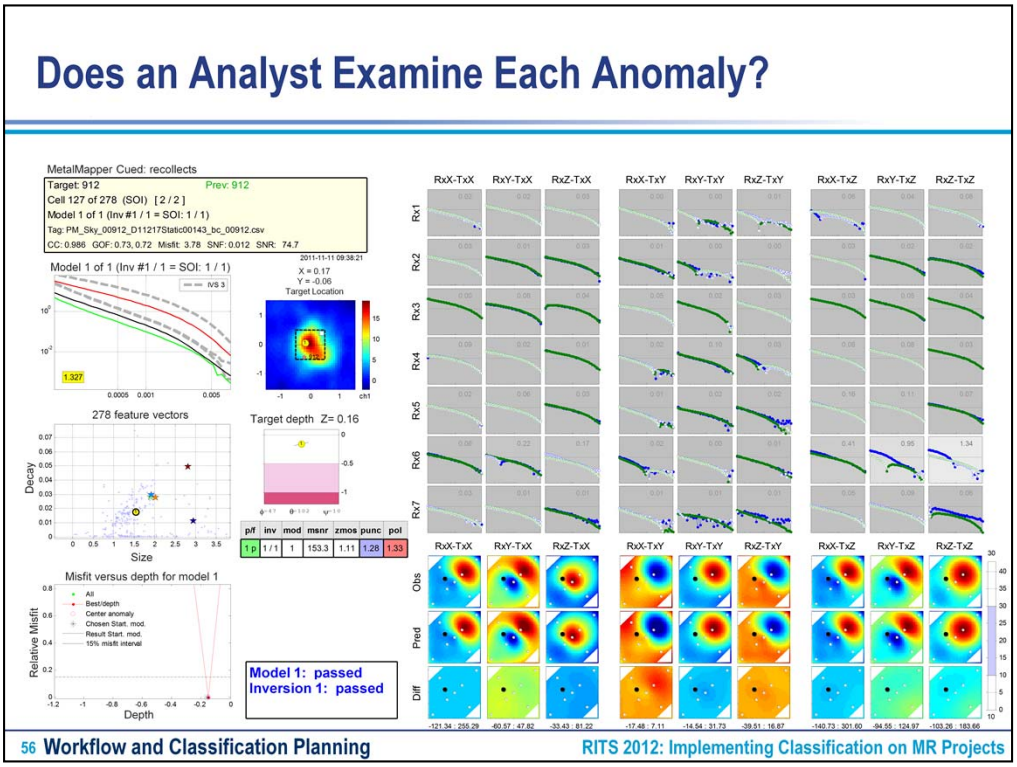
Appropriate seed items and how deeply they should be buried are determined by the project team. These are some of the items that have been buried to check the quality of the contractor. Note that we don't have perfect site knowledge and one additional check on the classification process could be the planting of unknowns. The production area seeds should be encountered daily and the strips should be run twice daily.



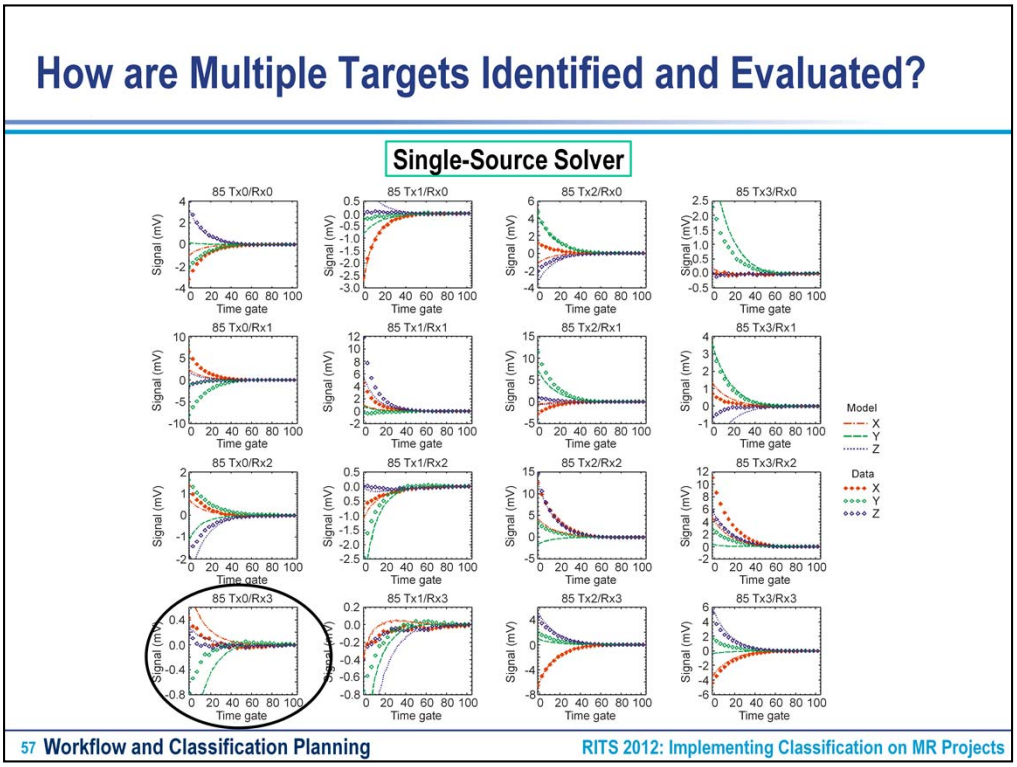

Here are the polarizability curves for different items in the IVS. Note the repeatability of the measurements over time and the QC measure of less than 10% delta



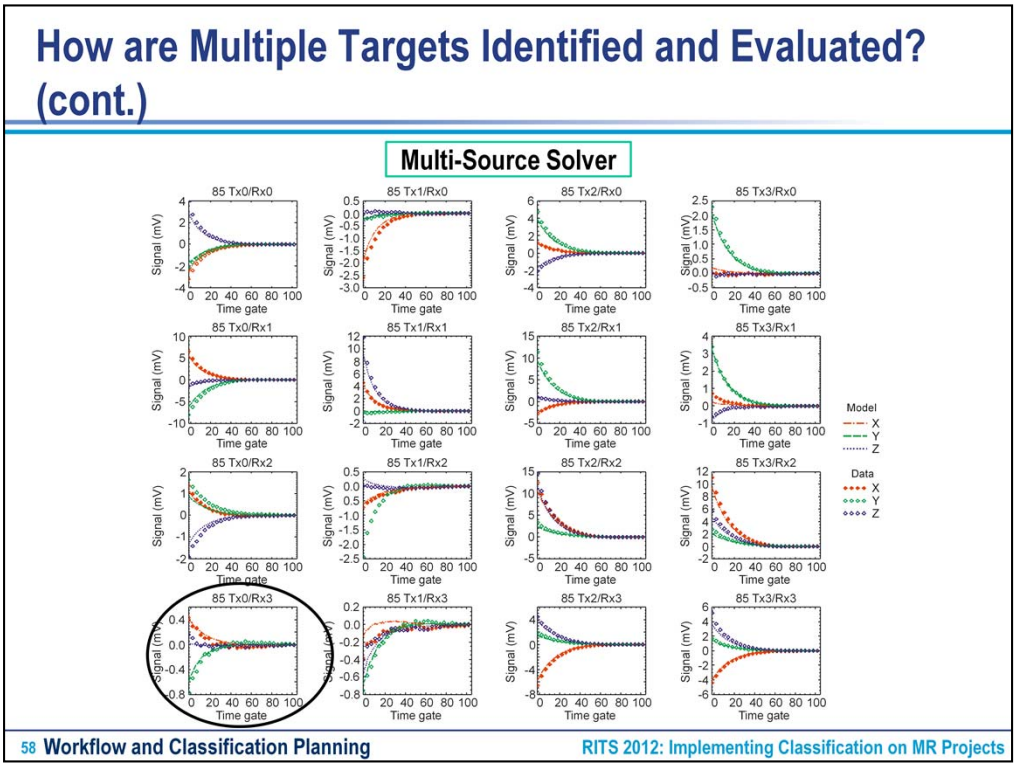

This slide illustrates the importance of being correctly positioned over the target. The first example may get enough data for the transmit/receiver coil #2 but clearly the other pairs will suffer in data quality. The second figure shows a correctly positioned sensor that will allow enough data to be capable to perform classification. Note that the standard for the MM is 0.4m and that teams have had to redo in the past 10 to 15% based upon post processing that night.




Another major consideration is whether an analyst reviews the data on each anomaly. This should be a systematic process of looking at the polarizability curves, the estimated position of the target, the function of the transmit/receive coil pairs, and the comparison of the model data to the actual data.




You also should discuss with your contractor how multiple targets will be identified. On this slide a single source solver is used to model the data. As you can see in the circled graphic, there isn't a very good overlap between the model and the measured data.




This slide shows a much better fit using the multi source solver and thus there should be multiple targets under the sensor that will have to be classified individually.

## Overview of Presentation

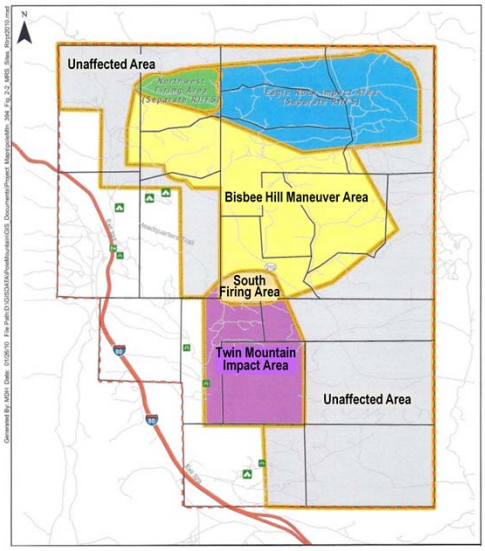
- Introduction
- Classification Basics
- Workflow and Classification Planning
- ► Case Study: Pole Mountain, WY
- Summary

59 RITS 2012: Implementing Classification on MR Projects

Now that we have talked about all the elements of a classification project, let's work through an example to see how this works at a real site.

## Pole Mountain Training and Maneuver Area

- Bisbee Hill Maneuver Area
- South Firing Area
- Twin Mountain Impact Area




*Source: U.S. Army Corps of Engineers, 2010*

60 Case Study: Pole Mountain, WY RITS 2012: Implementing Classification on MR Projects

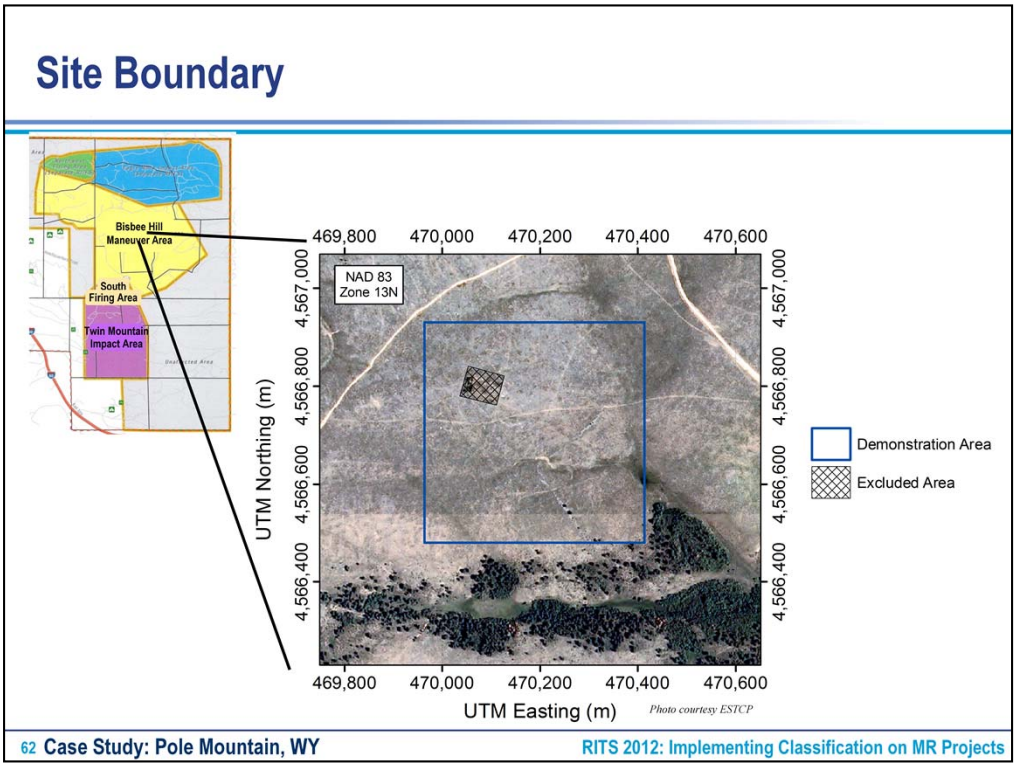
ESTCP performed a demonstration in summer 2011 at the Pole Mountain Training and Maneuver Area in Wyoming. The USACE is currently working at this FUDS site so we could share some logistics and support. The FUDS is comprised of three areas as shown on this map.

## Munitions at the Site

- A variety of munitions have been reported as used at Pole Mountain. Physical evidence for the following items was discovered during the RI:
  - Projectiles containing HE (high explosive) filler (37-mm to 155-mm, and 2.95-inch)
  - Shrapnel projectiles (75-mm and 3-inch)
  - 37-mm projectiles (inert and unfuzed)
  - 3-inch Stokes mortars (practice, fuzed) and
  - 60-mm mortars containing HE filler



A wide variety of munitions have been used at this site so we have a classification challenge.




The area we demonstrated on is a 50-acre parcel in the middle of the Bisbee Hill Maneuver Area. It is in a National Forest so gets a lot of recreational use.

## Project Details

- Seed emplacement
- EM61-MK2 detection survey
  - RTK GPS
- Select anomalies for further investigation
- Collect cued data using MetalMapper
- Intrusive investigation

63 Case Study: Pole Mountain, WY RITS 2012: Implementing Classification on MR Projects

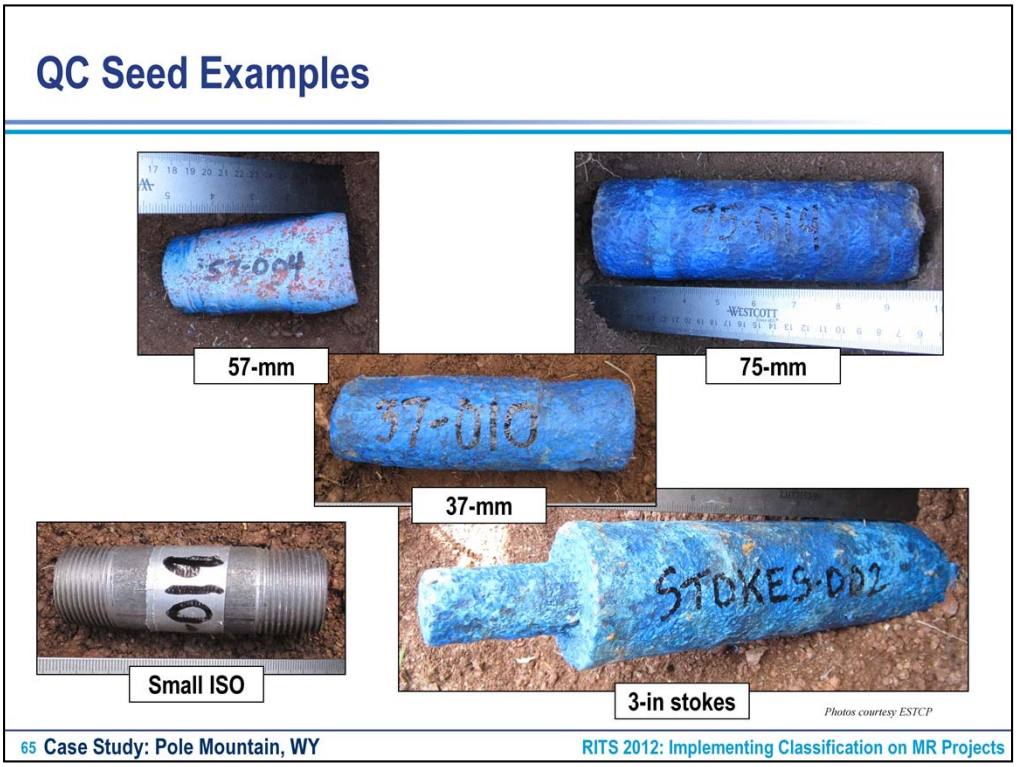
We used our standard approach at this site. We emplaced QC seeds, identified anomalies from an EM61-MK2 survey, then collected classification data over those anomalies using the MetalMapper system. At present, we use this two-step approach of detection then classification. There are several systems in the research program that do both jobs in one pass.

## Training Data

- Twice daily Instrument Verification Strip (IVS) + Static Tests
  - One sphere
  - Two small ISOs
  - One 37-mm projectile
  - One 75-mm projectile
- Training Pit
  - 37-mm, 57-mm, and 75-mm projectiles and 3-in stokes mortar
  - Two depths (at least one with good SNR)
  - Four orientations



We also collect sensor data that will be useful for training our classifier. We use the GSV process at all our sites so we set up an Instrument Verification Strip for twice-daily testing of the sensor system. We also have a shallow training pit on site to collect classification sensor data over a few of the munitions expected at the site.




Examples of the items in the IVS and emplaced as QC seeds.

## The Premise

- The first ~25% of the Pole Mountain site was surveyed in the first field season
  - Anomalies selected from the EM61 survey data
  - Cued MetalMapper data collected
- Data analysis is complete and it is time for the site team to make some decisions
  - Are the data acceptable?
  - Is the analysis acceptable?
  - Can we stop digging at the contractor's stop-dig point?



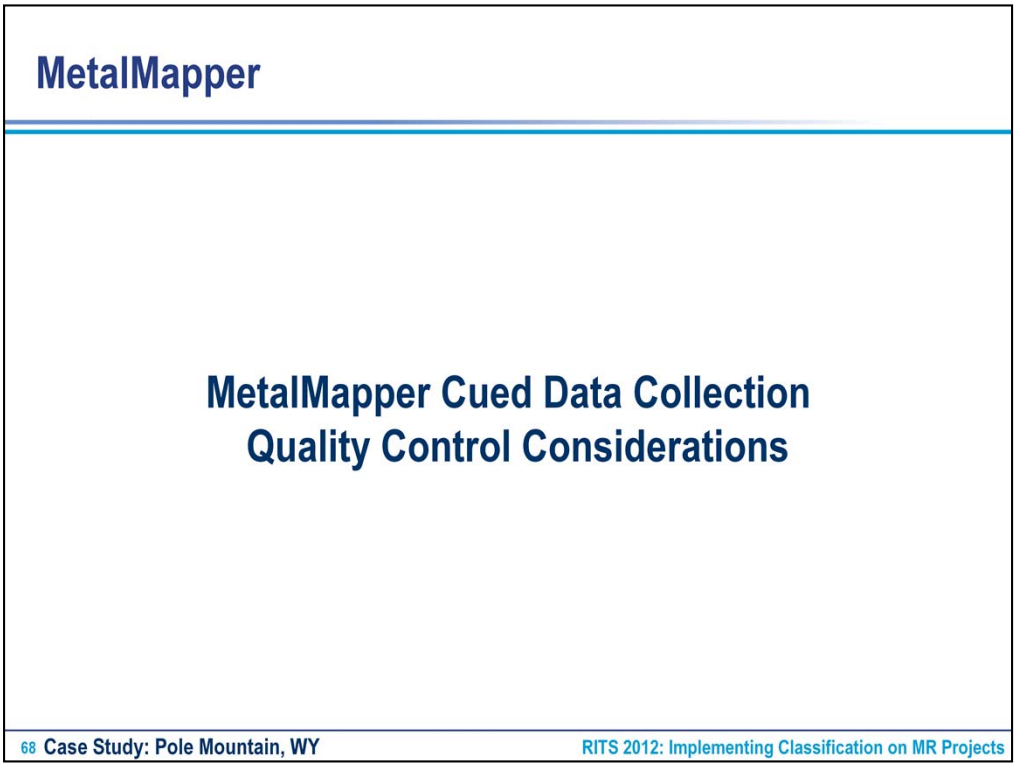
This slide details the premise of the exercise we are going to do. The audience will act as the site team for this site and will decide if the contractor's data and conclusions are acceptable.

## EM61-MK2 Data

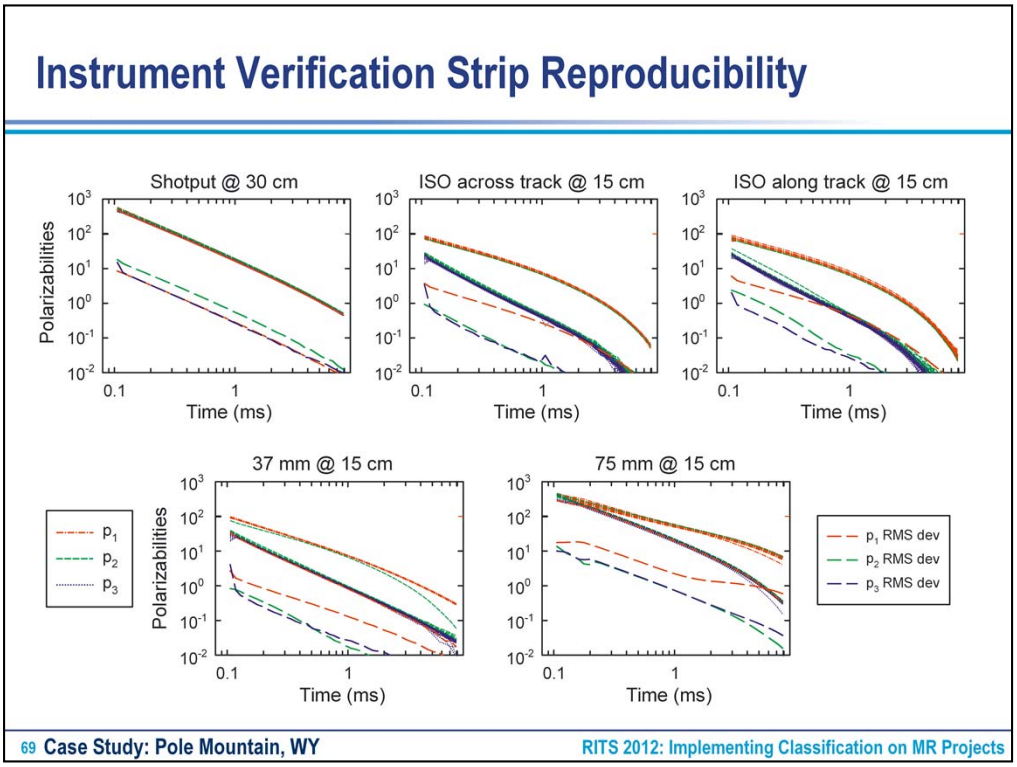
- **Anomaly selection based on 37-mm projectile at 30 cm**
  - **EM61 is appropriate sensor for this job**
    - RMS noise ≈ 0.7 mV in channel 2
    - Selection threshold = 5.2 mV in channel 2
    - 938 anomalies selected
- **All QC seeds detected using this threshold**
  - **Some just inside the 60-cm halo**
- **IVS reproducibility within requirements**
- **Coverage and measurement density within requirements**



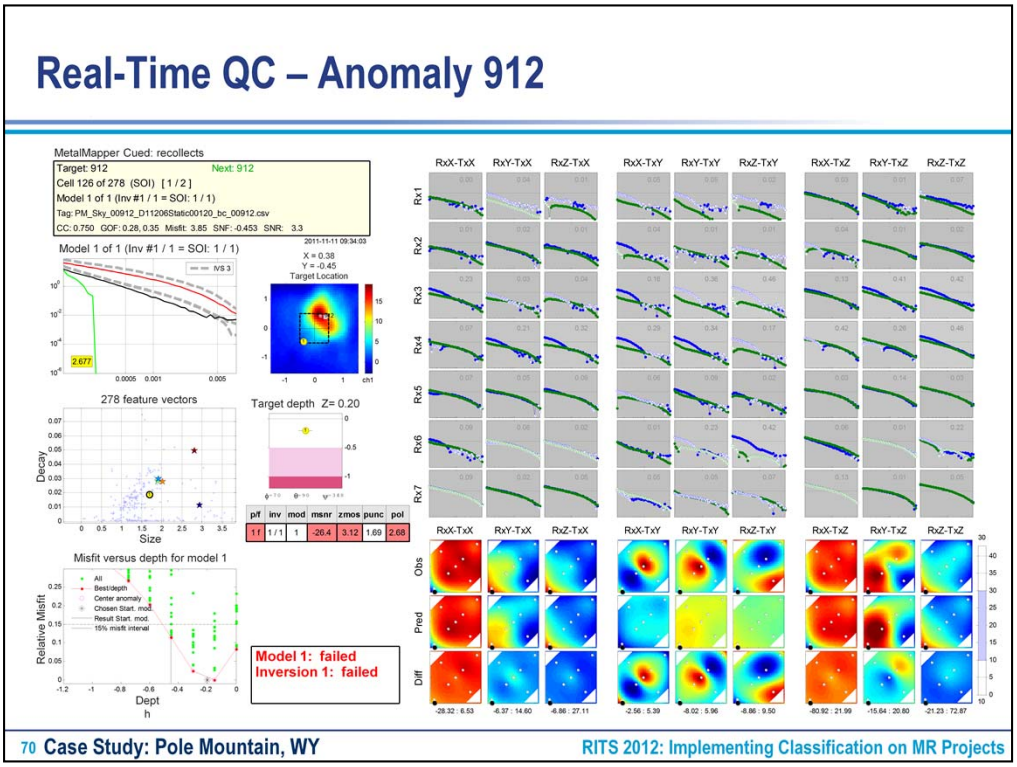
Summary of the EM61 survey data and analysis. All QC targets were met so we accept these data and analysis.



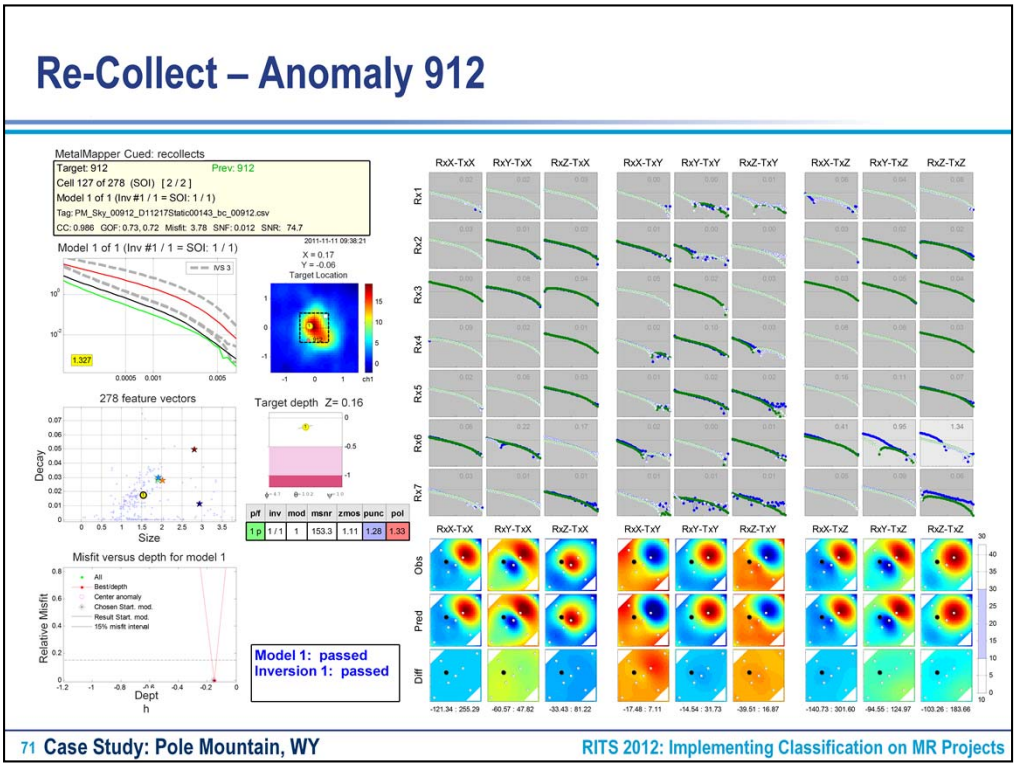

Now let's move on to the MetalMapper classification data.




The IVS is used to make sure the sensor system is working correctly morning and evening of each survey day. In a detection-only survey, this means confirming that the signal amplitude over each item in the IVS is the same each pass. For a classification sensor, we require the polarizabilitiy curves derived from the measurements to be the same each pass. Here we have plotted the curves for each measurement over each item and the resulting standard deviation. The specification is that the results have to be repeatable to 10%. In most cases the standard deviation was a few percent of the measured values and below 10% in all cases. These data are acceptable.




As we discussed above, we also want to make sure that an analyst examines each data set. Here is the QC plot provided by the contractor for the first data set collected over anomaly 912. Notice that the observed and modeled decays (red and blue lines) in the upper right do not match well at all. We see the reason for this when we look at the false color plot in the upper middle of the slide. The position of the MetalMapper when the measurement was made is plotted over the EM61 survey data. As can be seen, the MetalMapper was positioned to the SW of the actual anomaly position. We have a requirement that the center of the MetalMapper be within 40 cm of the anomaly position to ensure good illumination of the target. This anomaly must be remeasured.




The second time the data collection team visits anomaly 912, they position the MetalMapper in the correct location and the data are reliable. This recollect is acceptable.

## MetalMapper Cued Data

- Cued data collected over all anomalies
  - Real-time QC
  - Some re-collections
- IVS reproducibility within specifications
- Data are acceptable



We have seen the QC steps taken for the MetalMapper data. All requirements have been met so we judge the data to be acceptable.

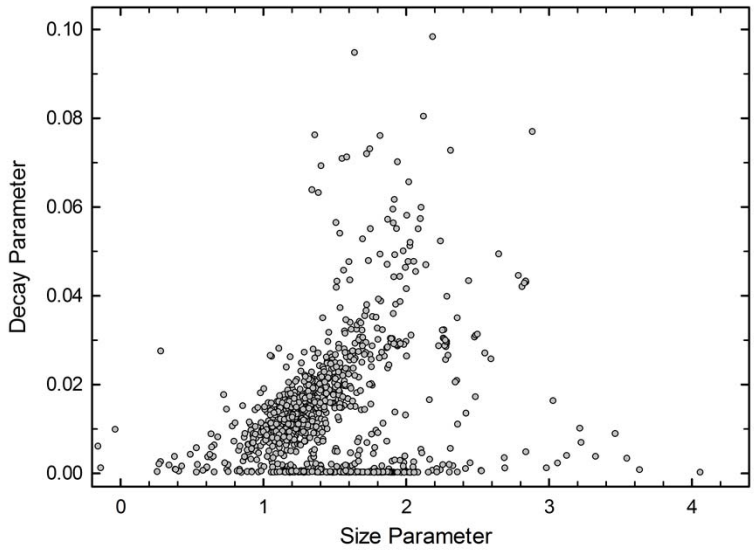
## "Decision Time"

**MetalMapper Analysis and Anomaly Ranking**



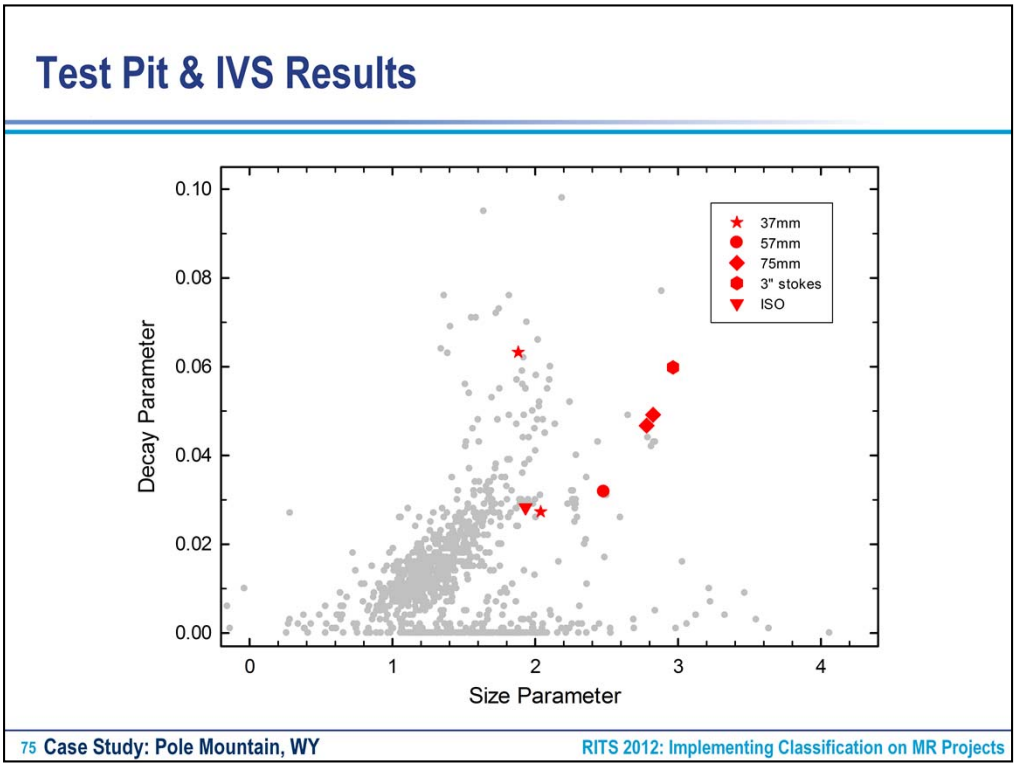
Now comes the hard part – the classification decisions we make based on these data.

## Statistical Analysis – 2D Feature Plot

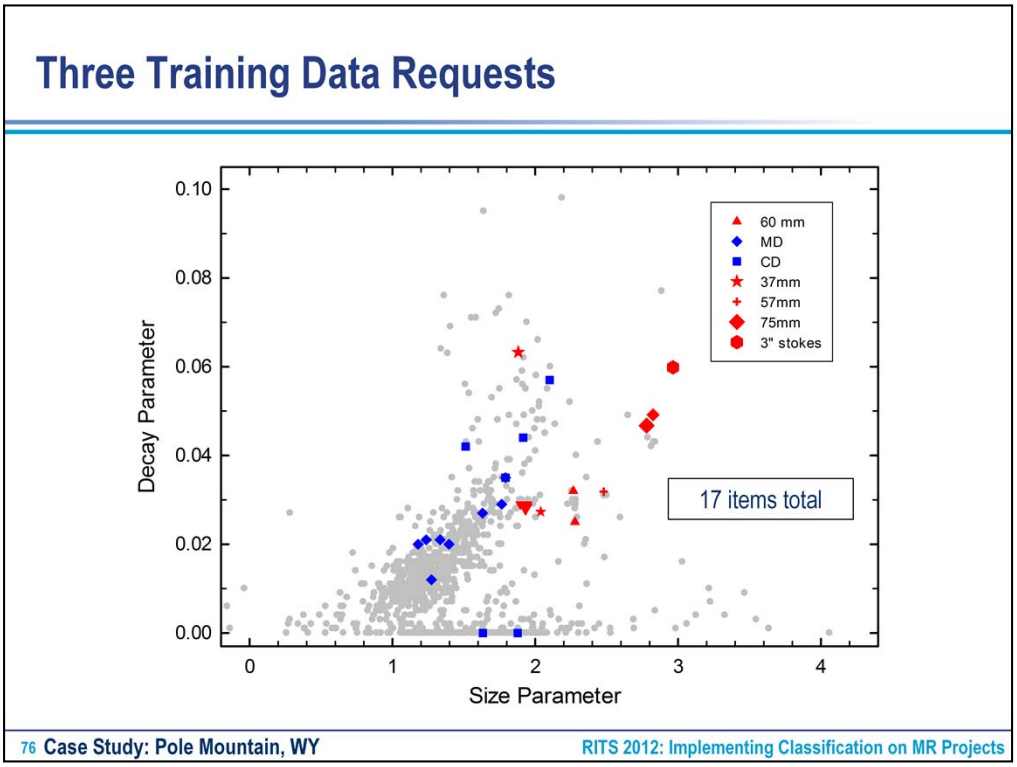



74 Case Study: Pole Mountain, WY RITS 2012: Implementing Classification on MR Projects

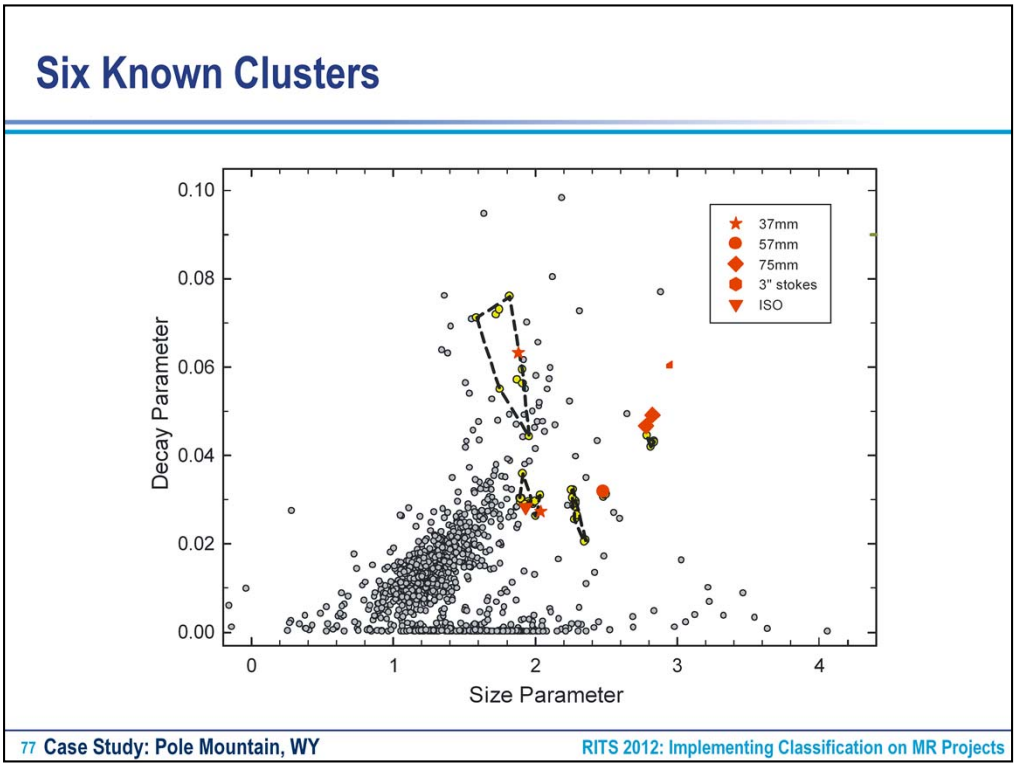
We'll start our consideration of classification by looking at the familiar 2-D feature plot with 938 points representing all the anomalies in the first phase at Pole Mountain. Remember that we are not going to make any decisions based on this simplified presentation, just use it to guide the discussion.



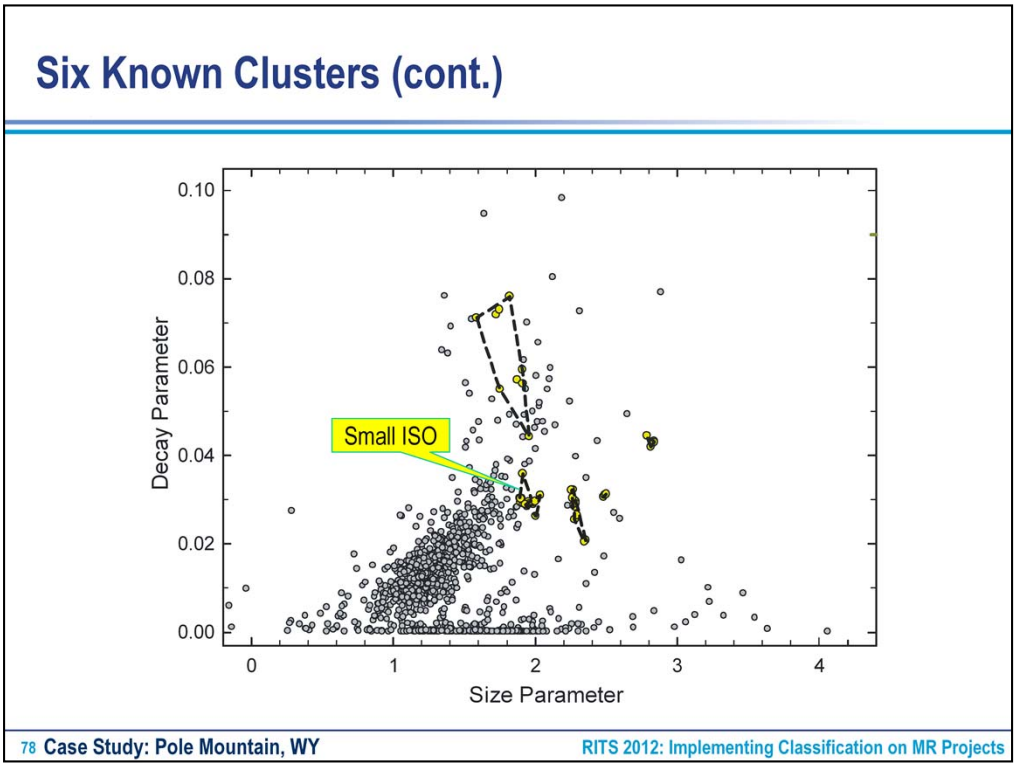

On this slide, we've plotted the size, decay pairs for the munitions items we measured in the IVS and the test pit. As expected, they are in the upper right of the plot. We can reasonably expect many of the buried anomlies that fall close to these points to be munitions also.




To train the classifier so that it can draw a boundary between munitions and not-munitions, we need to give it the identities of some other points on the plot. We don't just choose these training points randomly, we'll walk through the methods we use to obtain the most information from the fewest digs (which cost money).



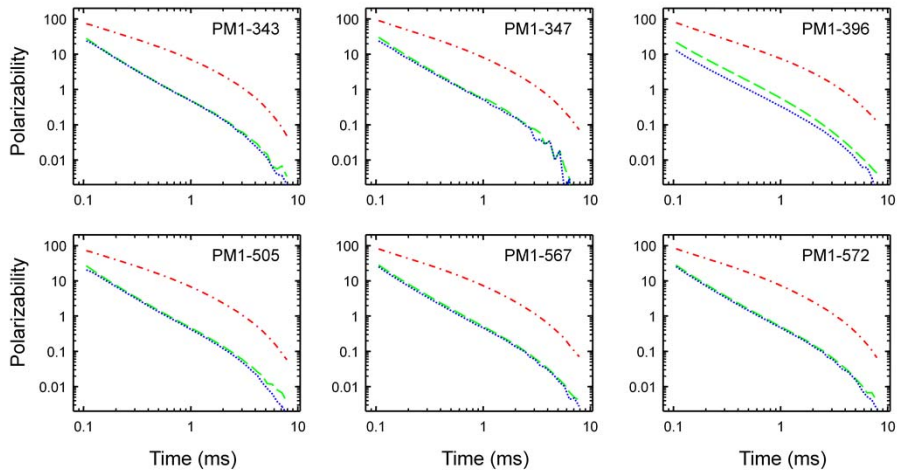

The first step is to examine all the anomalies that fall near the munitions we know about and see how they match munitions. By match in this case we mean the entire polarizability curves match, not just that they are close in this 2-D space.

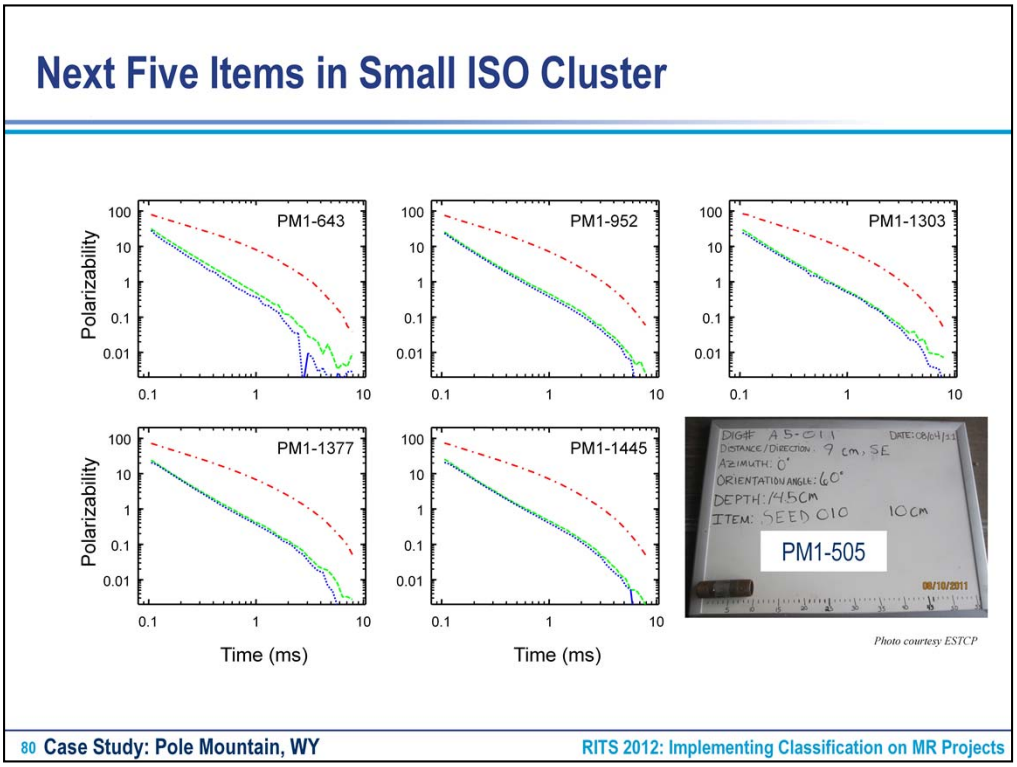



We'll use the cluster that corresponds to the pipe nipples called Industry Standard Objects as an example. There was an ISO in the IVS so we know where this object is located on this 2-D plot. We'll look at all the points around that location and compare the full curves. It turns out there are 11 anomalies whose polarizability curves match that of the ISO from the IVS.

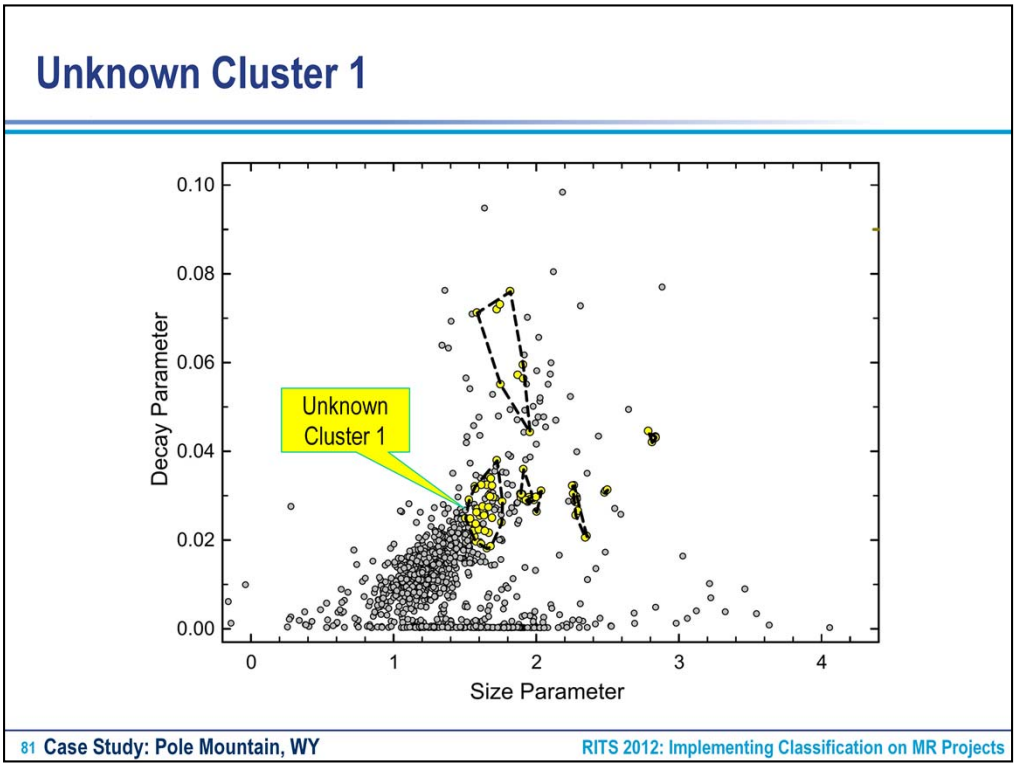
The essential point here is that we use this simplified 2-D plot to make it easy to look at the data but we only make decisions based on the full curves.

# First Six Items in Small ISO Cluster

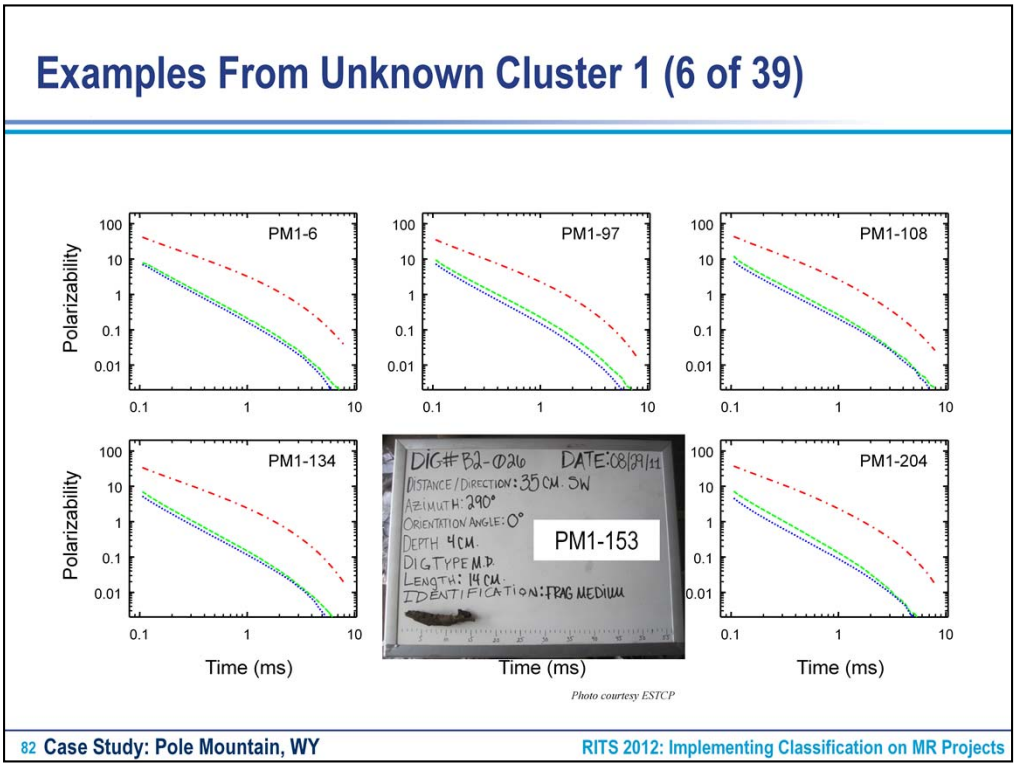




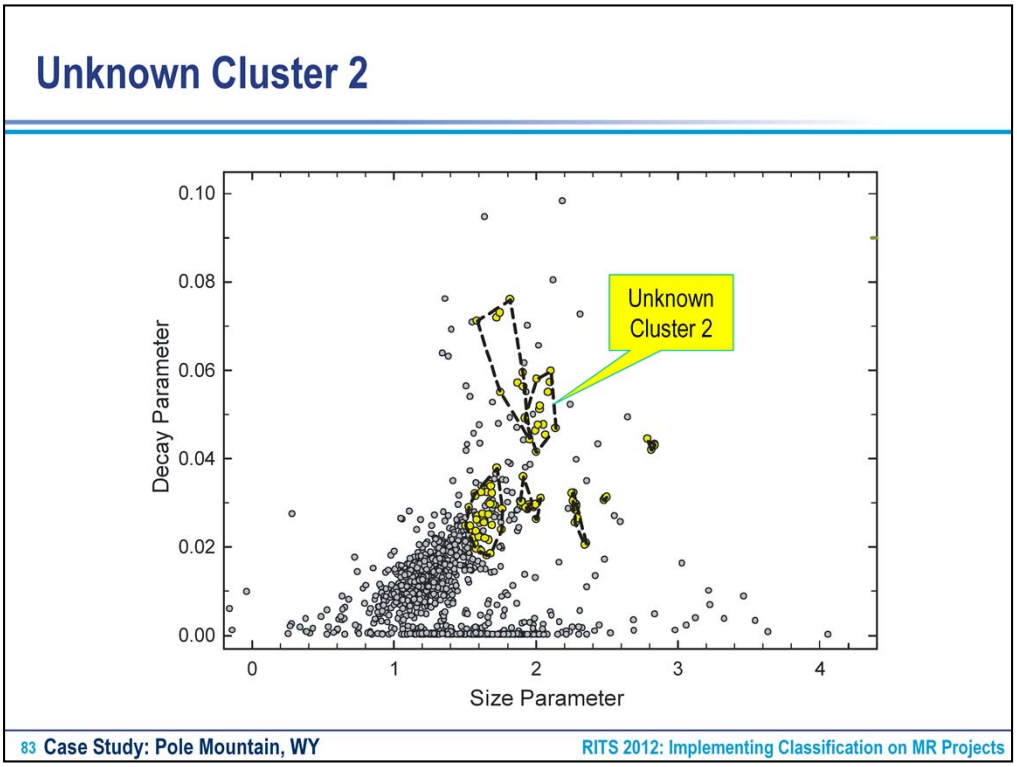

Here are the remaining five anomalies in the ISO "cluster." When we dig one up to check, we are not surprised to see that it is a small ISO. We can follow the same process for all six clusters of anomalies that match munitions we already know. All of those items go on the red part of the ranked anomaly list.



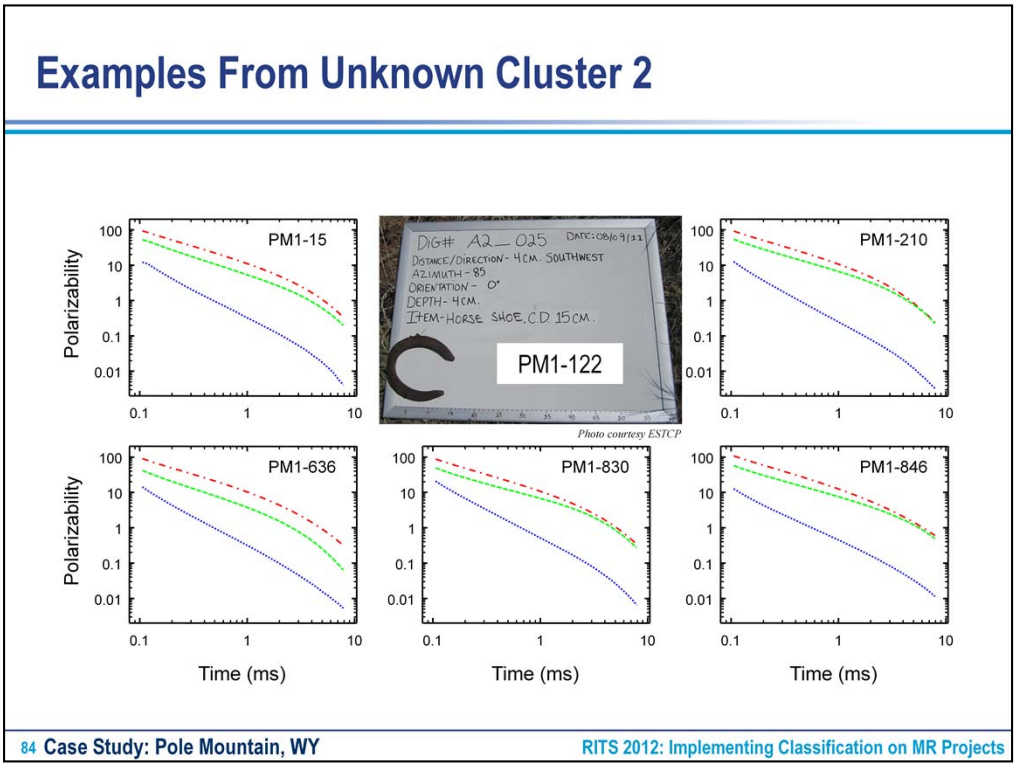

Now we start looking for clusters (groups of anomalies with matching polarizability curves) that do not correspond to munitions we expect. There is a cluster of 39 anomalies that shows up as shown on this slide.




Here are the polarizability curves for the first six of the 39 anomalies in this cluster. They all look roughly cylindrical (the two smaller curves don't match exactly) but have smaller amplitude than the pipe nipples (which is why they were to the left on the plot). When we dig one up to check, it is a small piece of frag that is roughly symmetric. The other 38 of these anomalies can go on the list as high-confidence clutter.

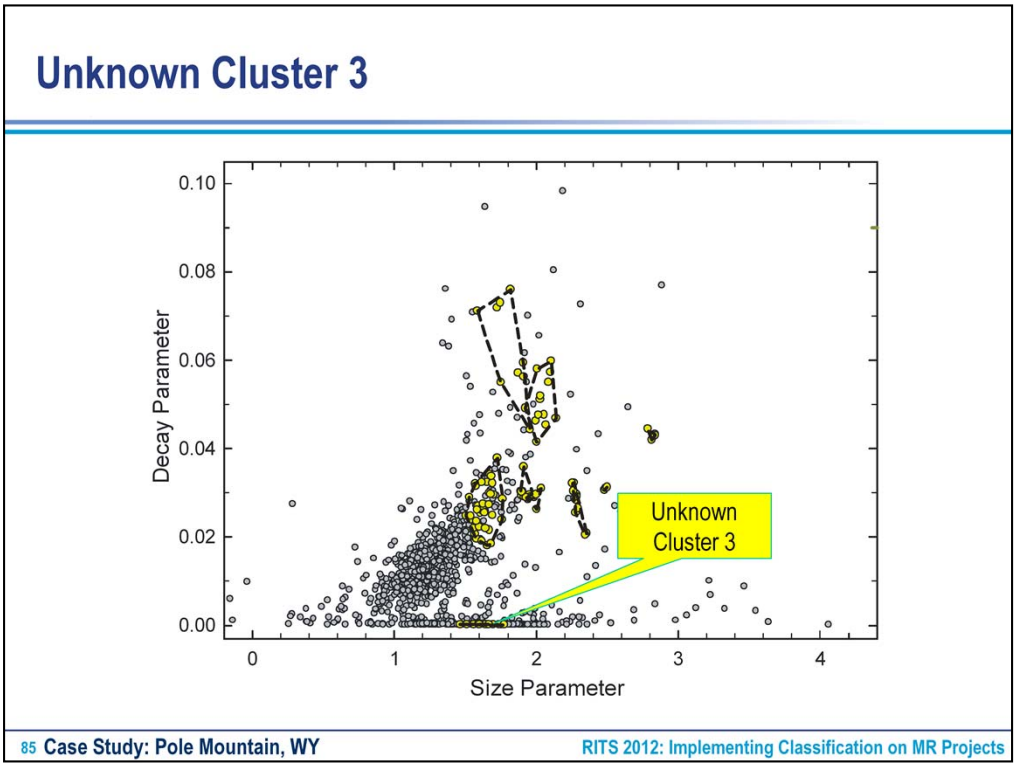



Here is another “cluster” of anomalies that is in the part of the curve where munitions are found.

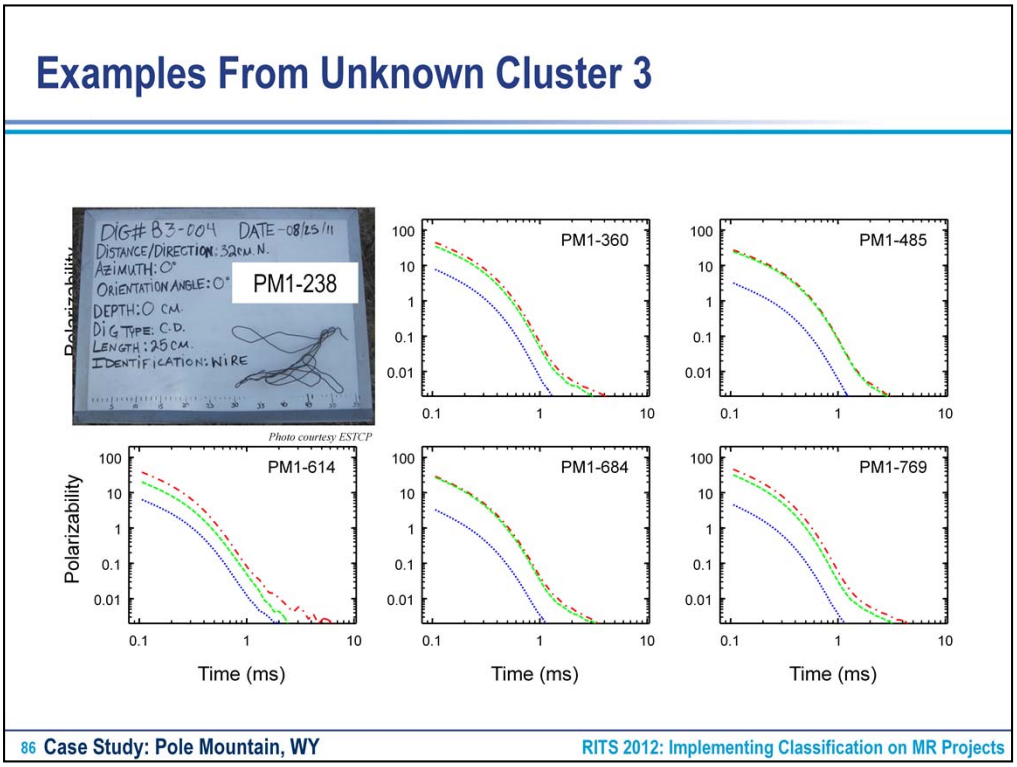



When we look at the first six of these, we see that they are plate-like, two large and one small response, and not symmetric. When we dig one up we see that it is our old friend the horseshoe. We can put the rest of these anomalies on the high-confidence clutter part of the list.

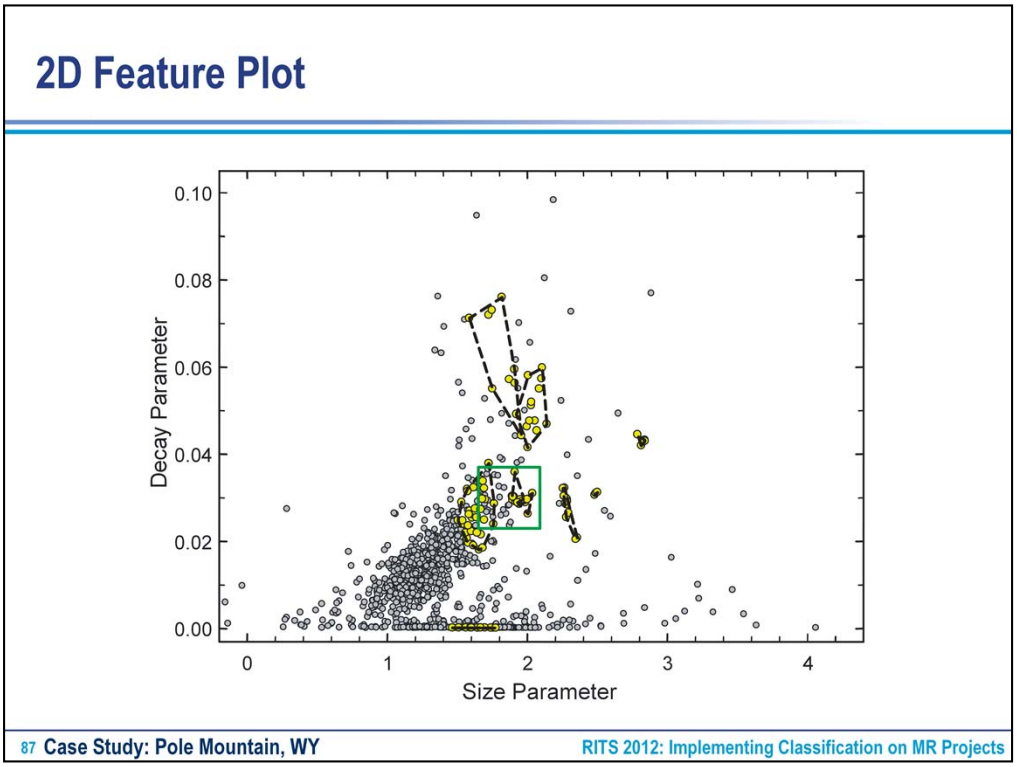
If there ever is a call for a horseshoe detector we've just proven we have a good one.



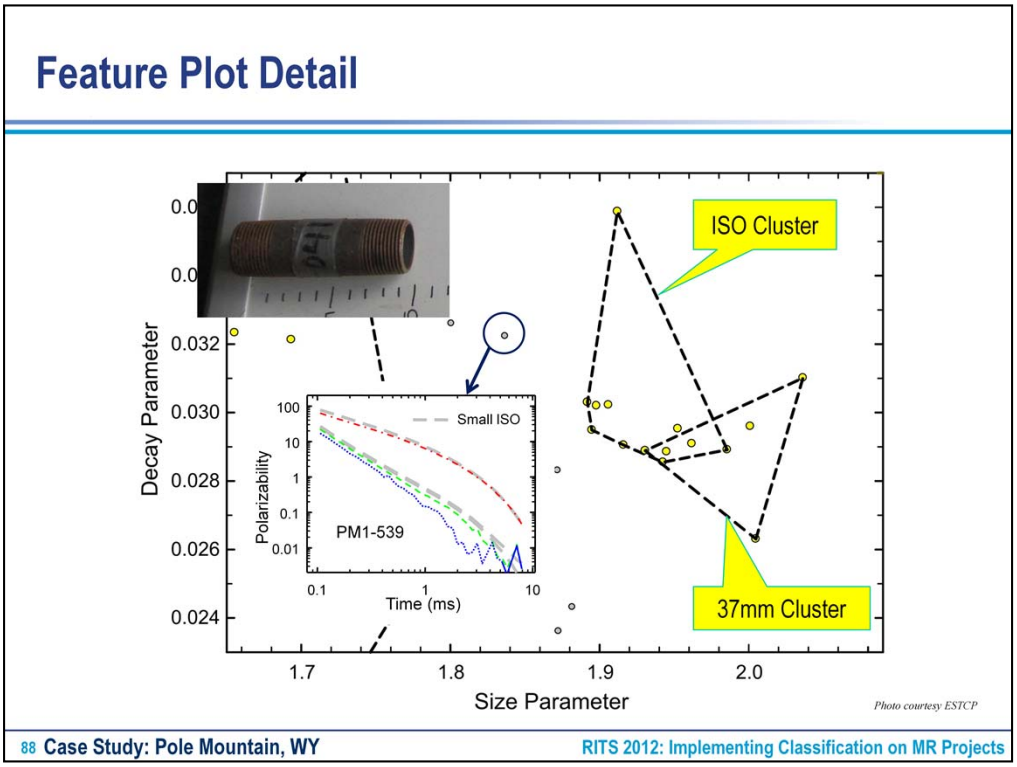

One more example. Here we've marked a large cluster far from where we expect munitions – the decay is very fast meaning the items have thin walls.



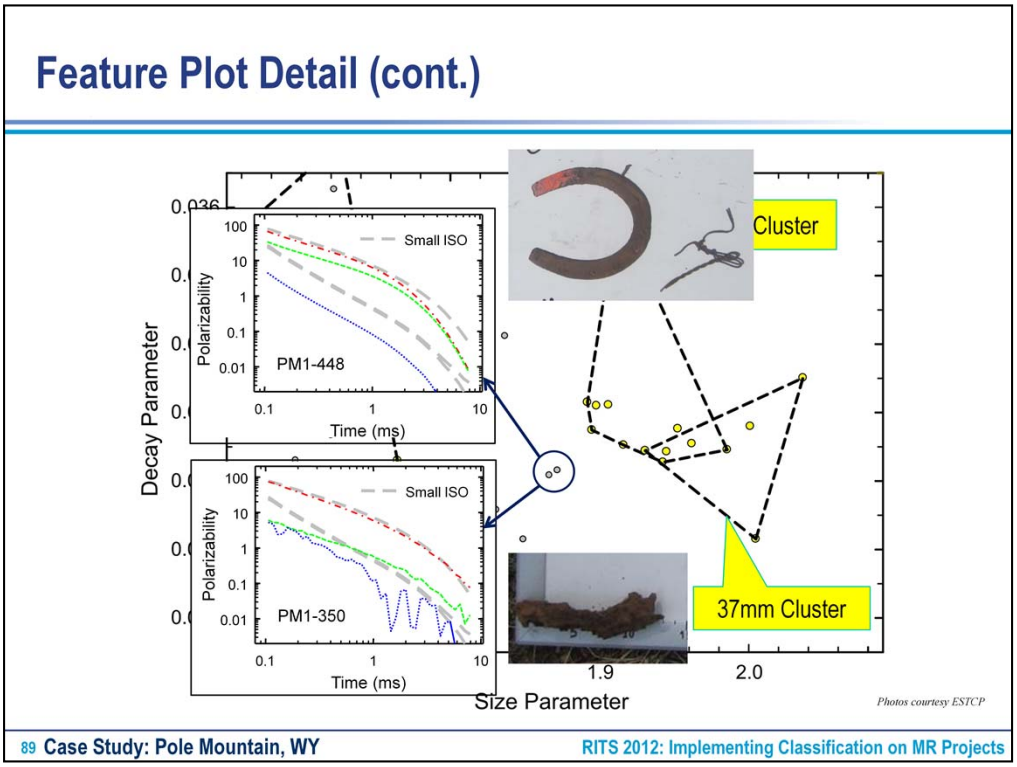

Here are the first six items in the cluster. The signatures are plate-like (two large responses and one small) and the polarizabilities decay very quickly. When we dig one up, we see that it is a flattened coil of wire. Another large group of anomalies for the green part of the list.




Our work is done. We've trained the classifier (and our own eyes) and are ready to construct a dig list. There are a couple of dozen anomalies between the clusters we have identified, we're going to be conservative and mark them all to be dug.




This is a blow-up of the part of the feature plot between the clusters. We marked these all for digging to be conservative. The point circled has the polarizability curves shown in the inset. In grey, we have plotted the small ISO results from the IVS. Even though the third curve is a little ragged, this is a good match so we are not surprised when the item turns out to be a pipe nipple.




Now let's look at the two anomalies just down and over from that pipe nipple. They are very close in the 2-D plot but when we look at the full curves they look nothing alike. One more reminder that the advanced sensors provide the information to make this work. Anomaly 448 looks very plate-like and turns out to be a horseshoe. Anomaly 350 is cylindrical but not completely symmetric. It turns out to be a bent up fragment. We dug both of these to be conservative but if we had examined each one in detail we could easily have left 448 in the ground and maybe 350.

## Ranked Anomaly List

| Rank | Decision Metric | Comment |
|---|---|---|
| -9999 | -9999 | Can't extract reliable features |
| 1 | .97 | |
| 2 | .96 | High confidence munition |
| 3 | … | |
| … | … | Can't make a decision |
| … | … | Can't make a decision |
| … | … | |
| … | … | |
| … | … | High confidence non-munition |
| … | .03 | |
| … | .03 | |
| … | .02 | |
| N | .01 | |



This is the ranked anomaly list we were directed to provide.

## Ranked Anomaly List – Final

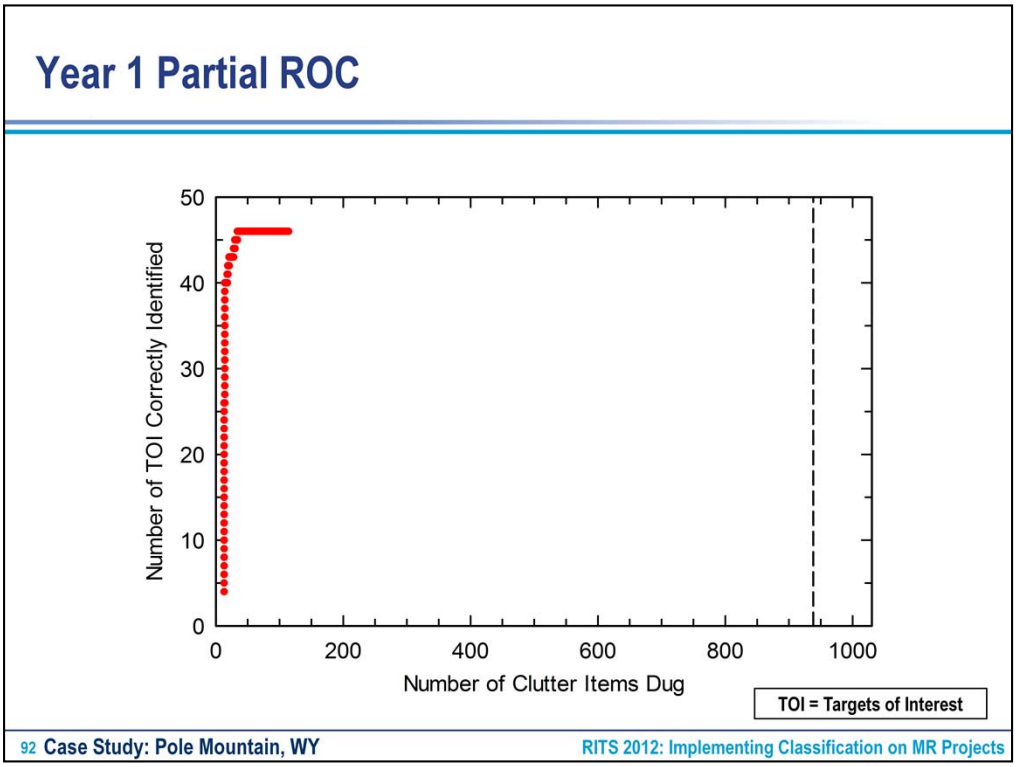
| Rank | Decision Metric | Comment |
|---|---|---|
| -9999 | -9999 | Can't extract reliable features |
| 1 | .97 | |
| 2 | .96 | High confidence munition |
| 3 | … | |
| … | … | |
| … | … | |
| … | … | |
| … | … | |
| … | … | |
| … | … | High confidence non-munition |
| … | .03 | |
| … | .03 | |
| … | .02 | |
| N | .01 | |

Threshold

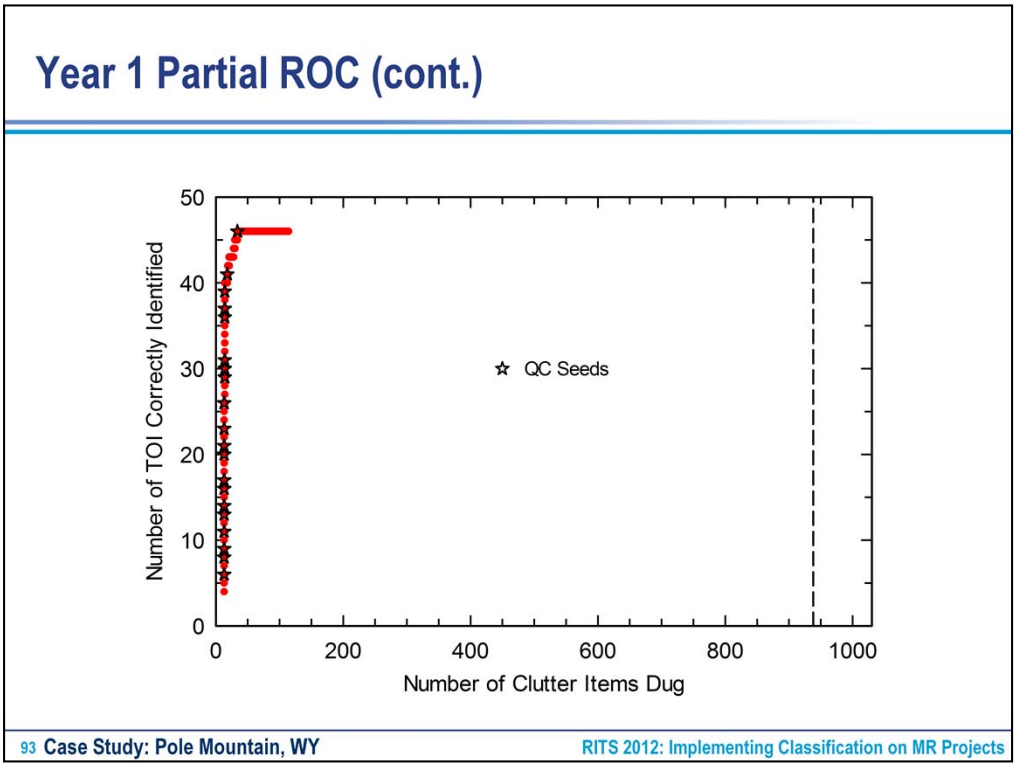
91 Case Study: Pole Mountain, WY

RITS 2012: Implementing Classification on MR Projects

We made a decision on everything so this is what we actually submitted.




We dug down through the red part of the anomaly list and this is the partial ROC we get. We labeled the y-axis as Targets of Interest identified because we used some pipe nipples as QC seeds and it is not correct to call them munitions. We said we were going to be conservative with our list; notice that the last 75 anomalies we dug were clutter.



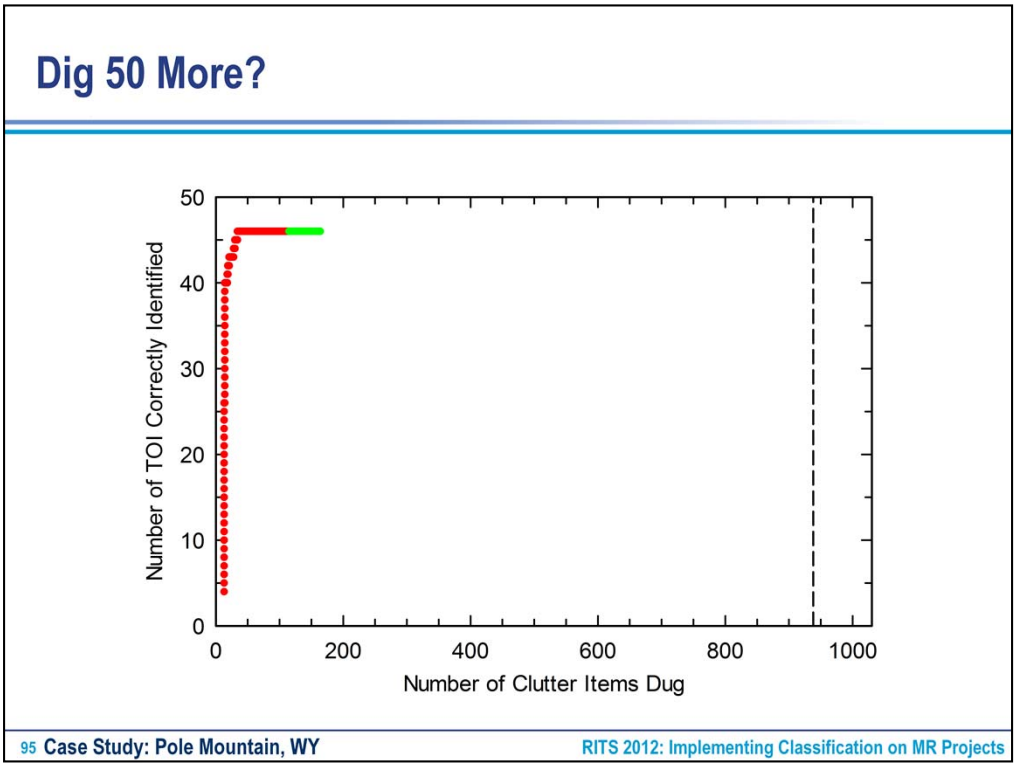

As a check, you, the site team want to make sure we correctly classified the QC seeds. We did, their positions on the partial ROC curve are indicated by stars.

## Year 1 Analysis

- Is the analysis acceptable?
- Do we accept the stop-dig point?

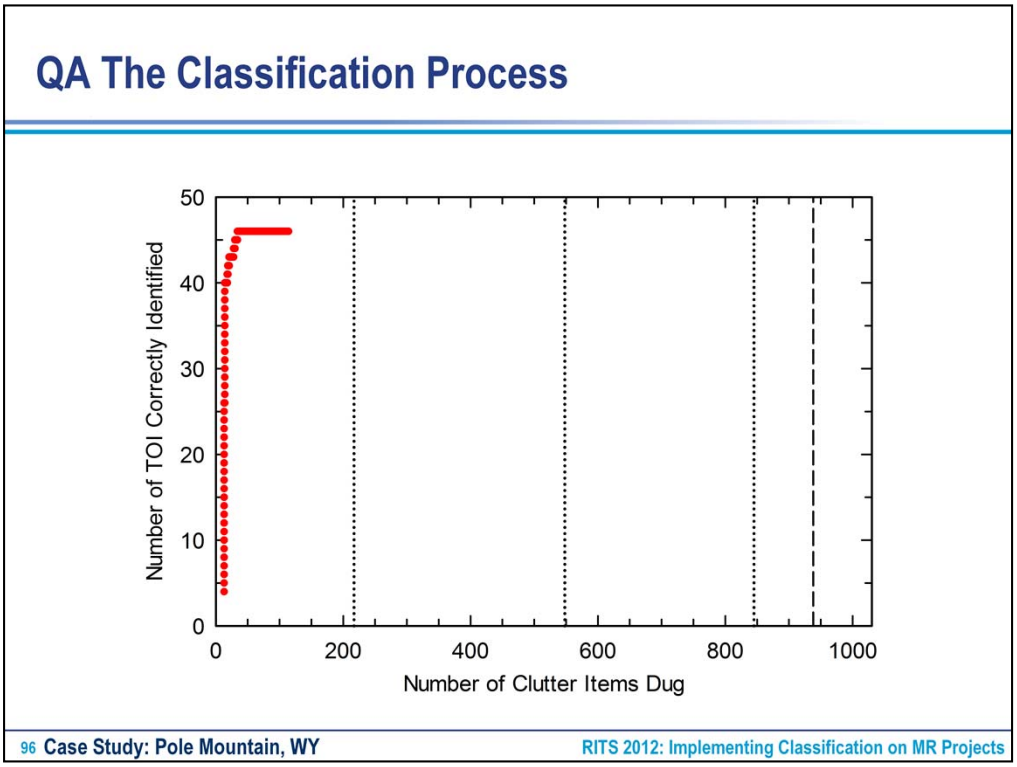


It is now your turn to decide if we can stop digging.

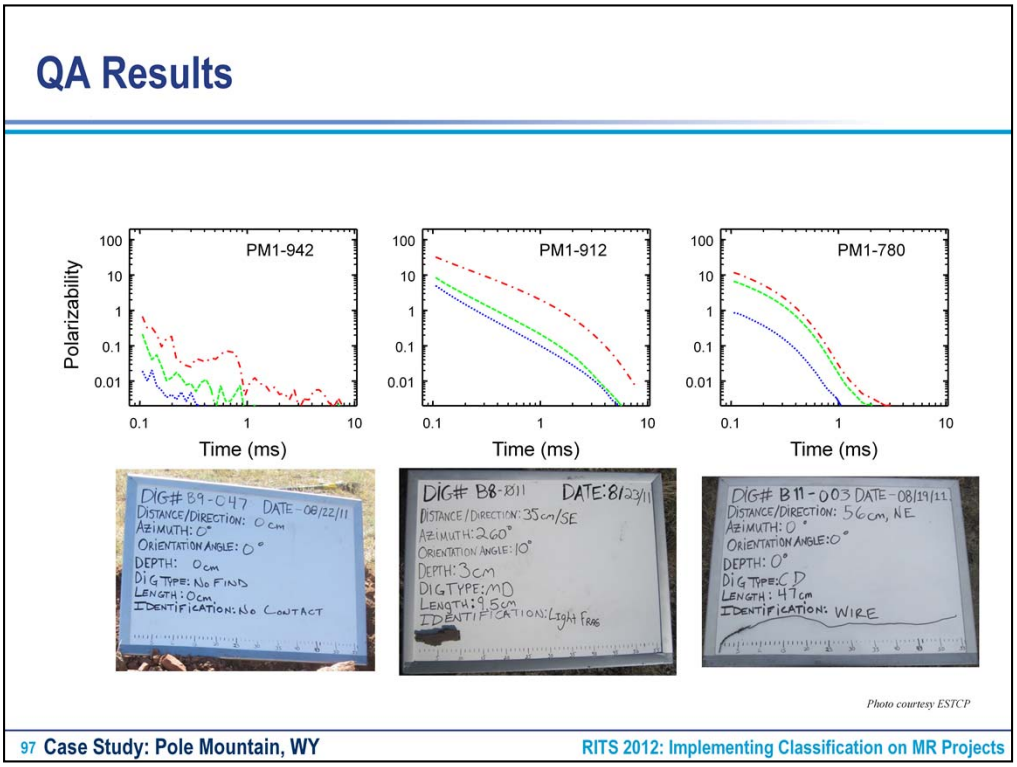



Sometimes people say that they need a little more convincing. One audience asked to dig the next 50 items on the anomaly list. We did that and they were all clutter. We colored those points green because we were confident they were not munitions from our classifier.

At this point, most groups say we can stop digging. They still want to have some sort of QA to make sure we didn't miss anything. One suggestion we often hear is to dig some random points further down the anomaly list to check for UXO. This is not a very good idea for the following reason. UXO are rare at this site (as they are at most sites). We have already dug 47 TOI so at worst there are one or two left. You would have to be exceedingly lucky to find one more UXO by digging tweny more holes. All that would do is give you a false sense of security.

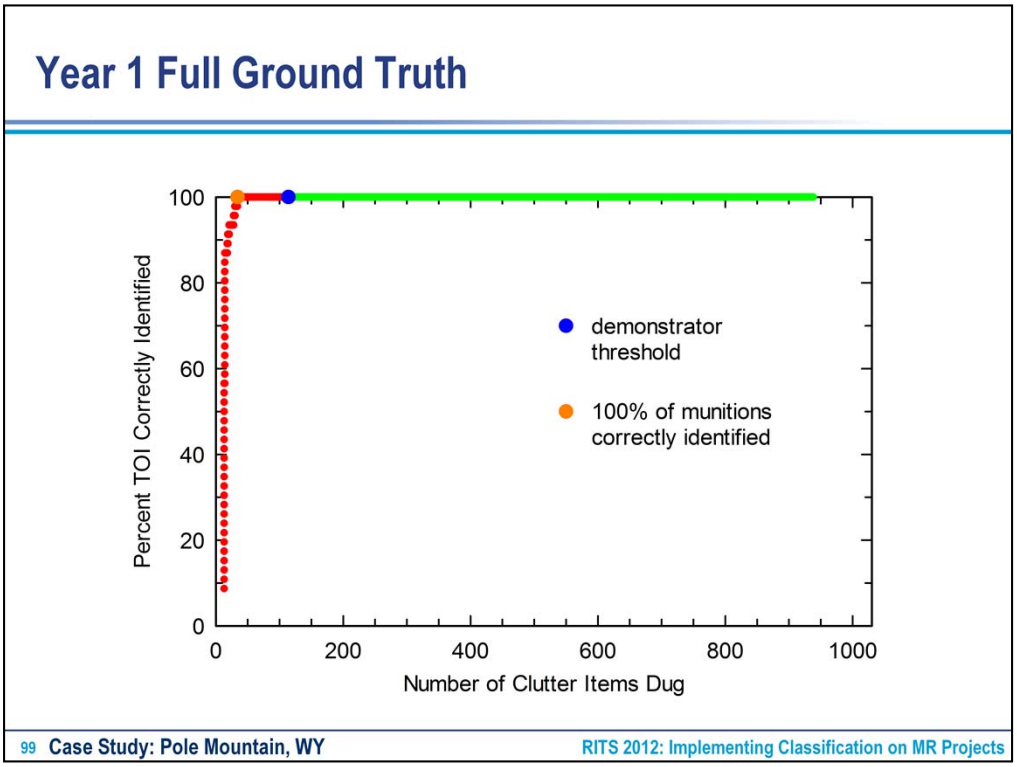



We think a better approach is to QA the whole classification process. Let's dig some randomly chosen anomalies further down the list not expecting to find UXO but making sure that the classification process worked. For illustration I have chosen three more anomalies at random for QA digging.



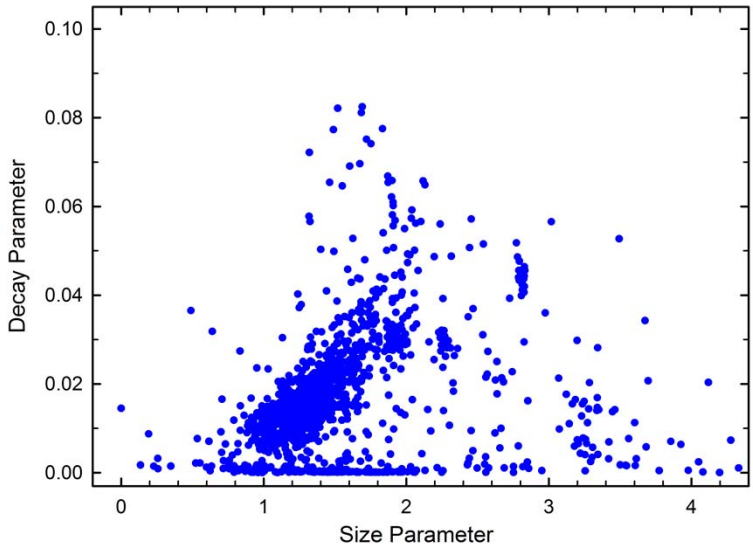

The first two sets of polarizability curves matches what we recovered. Anomaly 942 is some small ragged response, just above the noise. The UXO techs were not able to find anything in that hole. Anomaly 912 looks like a small piece of frag and it is. Anomaly 780 looks just like the flattened coils of wire we saw earlier but the photo shows a straight piece of wire. When we query the dig team, they tell us they straightened out the wire so we could see how long it was. We are relieved and tell them not to try to help us any more.

In the real world, we would never know how we did. This is a demonstration so we can see whether our decision was correct.




We did fine. As promised, the original dig threshold was chosen conservatively. Even with that, we were able to avoid 800 digs.

## Year 2 Feature Plot




100 Case Study: Pole Mountain, WY

RITS 2012: Implementing Classification on MR Projects

We've collected data from the remainder of the site. The feature plot looks very much like year 1 so we can feel confident using the classifier and thresholds from year 1.

## Overview of Presentation

- Introduction
- Classification Basics
- Workflow and Classification Planning
- Case Study: Pole Mountain, WY
- ► Summary

## Where Do We Stand Today?

√ Commercially-available sensor

√ Available analysis tools

√ Trained contractor base

√ Trained Government geophysicists

~ Regulator and Stakeholder acceptance

? Government business practices

? Government PM acceptance

102 Summary RITS 2012: Implementing Classification on MR Projects

The MetalMapper is commercially available. The Corps of Engineers has purchased 4 of these that can be made available as GFE.

The analysis software is freely available through Geosoft Oasis montage. Other software will continue to transition.

OSD supports the transition to classification as a means to make munitions response more efficient and accomplish more within available resources.

We recognize that there are challenges in contracting and regulatory acceptance. We are working within DoD to tackle the contracting challenges. On the regulatory issues, we are working with groups such as ECOS, ASTSWMO, ITRC and NARPM to educate regulators and gain acceptance.

Ongoing demonstrations involve both the production contractor work force and geophysicists from the Army COE. ESTCP will continue to offer training.

Environmental Council of the States (ECOS), Association of State and Territorial Solid Waste Management Officials (ASTSWMO), Interstate Technology & Regulatory Council (ITRC), and National Association of Remedial Project Managers (NARPM).

## Additional Training and Resources

- SERDP/ESTCP website contains webcasts, tutorials, and demonstration reports on the classification technology
- ESTCP Technology Transfer Workshop (June 6-7, 2012 Golden, Colorado)
  - Results of site demonstrations, outdoor technology demonstrations, tutorials, and software
- Advanced Munitions Response Site Management Workshop
  - Discusses the RI/FS and classification technology
- Navy Scope of Work Templates being modified to include classification technology

## Classification at Your Site?

- ESTCP will be hosting several demonstration projects in the next couple of years
  - RPMs should contact your MR Workgroup member to discuss it as a potential demonstration site
  - Contractors can submit live site proposals to the SERDP-ESTCP Website
- The technology is still relatively new, so an experienced contractor who has had sufficient training and field work with the classification process is recommended
- Consider performing a treatability study to determine site specific areas where the technology is suitable and to determine costs
- Third party QA will require a contractor with a thorough understanding of the entire classification process

104 Summary RITS 2012: Implementing Classification on MR Projects

What should an RPM do if she thinks her site is a good candidate for classification?

# Classification Cost Estimating

- **For RPMs use only, the "RI/FS Cost Estimating Tool" can be modified to include an additional geophysical survey, a reduction in the number of digs, etc. to generate an estimate for performing classification on the site**

INSTALLATION NAME: _____
DATE ESTIMATE DEVELOPED: _____

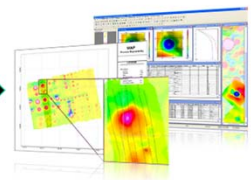
| | MRS AREA 01 Site Conditions/Scope | MRS AREA 02 Site Conditions/Scope | MRS AREA 03 Site Conditions/Scope |
|---|---|---|---|
| **General Site Information** | | | |
| Site Name/Number | | | |
| Munitions Response Site (acres) | | | |
| % of Site to Investigate during RI | | | |
| Site/Range Type | Select from List | Select from List | Select from List |
| Munitions Type | Select from List | Select from List | Select from List |
| Site Topography | Select from List | Select from List | Select from List |
| Site Vegetation | Select from List | Select from List | Select from List |
| Degree Regulatory/Public Concern | Moderate | Moderate | Moderate |
| Surface Anomalies (lbs/acre) | | | |
| Subsurface Anomalies/Acre | 0 | 0 | 0 |
| Total Subsurface Anomalies within Investigation Area | 0 | 0 | 0 |

Site Info | Summary | Calc_Area 1 | Calc_Area 2 | Calc_Area 3 | Cost Range | Cost Backup | ProfLabor | Definitions | References

## Where Are We Going?

- Smaller sensors for use in challenging terrain and vegetation
- One-pass classification – eliminate the need for stationary cued data
- Transition from developers to production geophysics companies
- Harder to classify munitions such as 20mm

$$\mathbf{H}(\mathbf{r}) = \sum_{i=1}^{N_v} \overline{\overline{\psi}}_i(\mathbf{r}) \cdot \boldsymbol{b}_i,$$



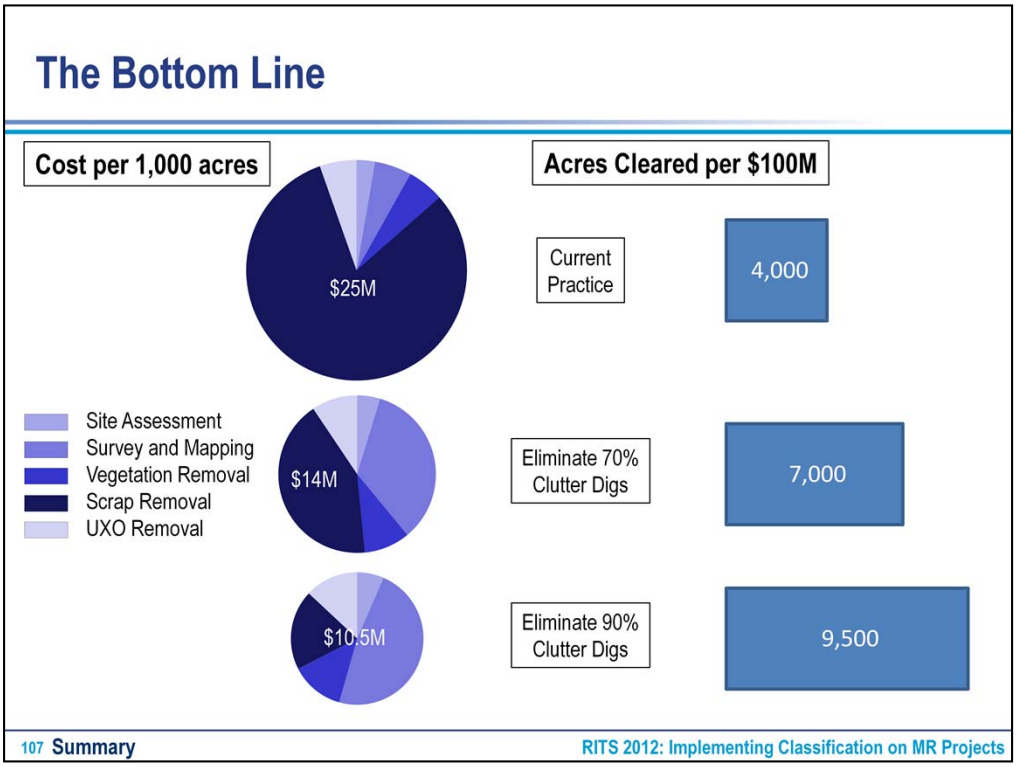
106 Summary RITS 2012: Implementing Classification on MR Projects

Most of the demonstrations to date have involved vehicle-towed sensors that operate in a cued mode – that is, have to stop over each potential target to collect data and make a decision.

Future demonstrations are planned to more extensively test the man-portable versions of these sensors.

A system intended to do one-pass classification from data collected while the platform is in motion is currently under development and planned for demonstration in 2012.

Finally, as the demonstrations proceed, the role of the technology developers is expected to continue to decrease and the role of production geophysics companies to increase. We view this hands-on experience as the most valuable type of training.




This slide looks at the impact that classification can have at the level of the MMRP program. We look at the cost to clear every 1000 acres and the amount of land that could be cleared for every $100M spent.

At the top we have current practice. Clearing 1000 acres with current practice costs about $25M, with the fraction of the cost dedicated to each task indicated by the pie chart. The bulk of the funds are spent for scrap removal (in the purple). Every $100M can clear on average 4000 acres of land.

If you can use classification to reduce the number of clutter items that are dug by 70%, which has been routinely achieved with advanced sensors, the distribution of how the money is spent changes. The fraction spent on survey and mapping increases and the fraction spent on digging clutter decreases and the overall cost to clear 1000 acres decreases to about $14M. Overall, the amount of land that can be cleared for every $100M spent increases from 4000 acres with current practice to 7000 acres.

If you can eliminate 90% of the clutter digs using classification, the impact is even more dramatic. We think this will be possible on some sites. Again, the fraction of the money spent on survey and mapping increases and the fraction spent on digging scrap decreases. The overall cost to clear 1000 acres decreases to just over $10M and the amount of land cleared of real hazard increases to 9500 acres for every $100M.

One important thing this analysis illustrates is that the transition to classification will require a transition in the contractor work force. More geophysicists will be needed and fewer UXO techs.

## Take-Home Messages

- Accelerate pace of cleanup at a constant funding level
  - Completion dates forecast decades out
  - Reduce risk more rapidly
- Better understood and transparent process
  - Higher quality data collected
  - All decisions documented and reviewable
  - Minimize operator effect on quality
  - Ability to adapt to new information
- Managing residual risk
  - Removals are not perfect under current practice
  - Some residual risk will always remain and must be managed
  - Clearing more land sooner is better