Printed from: Interstate Technology & Regulatory Council (ITRC). 2018. *Quality Considerations for Multiple Aspects of Munitions Response Sites.* QCMR-1. Washington, D.C. https://qcmr-1.itrcweb.org/.

3 Evidence-Based Decisions

Evidence-based decisions use of relevant information to provide a transparent, reasonable, and defensible basis for project decisions. Key points to evidence-based decisions include:

- Quality decisions require quality evidence.
- Quality evidence is achieved when information is fit for decision making.
- Evidence fit for decision making is relevant, reliable, based on appropriate data and judgment, and properly reflects uncertainties.
- Evidence may be qualitative or quantitative.
- Using multiple lines of interdependent evidence to test assumptions and strengthen the CSM increases confidence in decisions.
- The CSM and the DQO process are the primary tools for organizing, identifying, and acquiring the appropriate information for evidence-based decisions.
- Defining uncertainty is an integral component of the DQO process.

The goal of quality management is to identify, plan, and acquire the evidence necessary to answer specific questions, and then establish clear quality requirements. Properly assessing those requirements leads to confident and defensible decisions.

3.1 Types of Evidence

Evidence is relevant information collected or acquired to answer specific questions. Quality evidence, particularly multiple lines of interdependent evidence, increases confidence in the decision. The following hierarchy can help prioritize the quality of evidence for confident evidence-based decision making.

- Empirical evidence is observable, testable, repeatable, and falsifiable information collected through direct observation or experimentation.
    - *Qualitative Data*: Qualitative data can be found using the human senses (observation). Expert observers provide the highest quality observational data.
    - *Quantitative Data*: Quantitative data are measurements (such as concentration of munitions or munitions debris, amplitude, mass, or length). The quality of a measurement is defined in terms of measurement uncertainty.
- Historical evidence is evidence of past events that can be verified to a reasonable standard of certainty. Historical evidence includes documents, records, reports, artifacts, or other representations of past

> ***Confirmation Bias***
>
> *Confirmation bias is the tendency to selectively search for or interpret information in a way that confirms one's predeterminations. It is inappropriate to first make the decision and then align the evidence to justify the decision.*

events. Since historical events cannot be observed or repeated in the present, the quality of historical evidence is based on the quality of the source of the information.
        - *Primary sources* are original materials that have not been altered or distorted in any way. A primary source (also called original source or evidence) is an artifact, a document, a recording, or other source of information created at the time (an “eyewitness”) of the event (such as aerial photos, firing order, range map, or an explosive ordnance disposal [EOD] incident report). Primary sources, if relevant, are facts that represent the highest quality historical information and afford the highest confidence in decision making based on qualitative data.
        - *Secondary sources* describe, discuss, interpret, comment upon, analyze, evaluate, summarize, and process primary sources (for example, Archive Search Reports). Secondary sources often lack the firsthand nature of original material. A secondary source is generally one or more steps removed from the event or time period and is written or produced after the fact with the benefit of hindsight. Secondary sources are also documented during the MR-QAPP development on Worksheet #13 (Secondary data uses and limitations).
- Statistical evidence – Statistical evidence forms the basis for making an inference. Inference is the process of making an estimate of a population characteristic based on a sample from that population.
- Anecdotal evidence – Anecdotes, due to biased or small sample size, are frequently not representative of typical experience. Anecdotal evidence can indicate something is possible, but does not establish likelihood of success. This evidence is often used when other evidence is lacking. Often, testimonials used in advertising are anecdotal.
- Analogical evidence – This is a weak form of

evidence that suggests something true about one thing is also true of another due to similarity

3.2 Conceptual Site Model

The CSM represents relevant known and hypothetical site characteristics, conditions, and features developed from evidence collected or acquired throughout the project life cycle. A CSM can be presented in multiple ways including text description, tables, figures, flow diagrams, maps, and pictures. For example, if evidence indicates the MRS was a former artillery range, the CSM would include a high anomaly density area where UXO/DMM may be present in the former target. Known physical features of the MRS, such as geology, and data about the suspected munitions, such as their function and how they were deployed, are used to develop a preliminary depth profile. Additionally, known information about site access and land use are included to form a basis of understanding receptors and potential exposure pathways. A pictorial example of a CSM is presented in Figure 3-1.

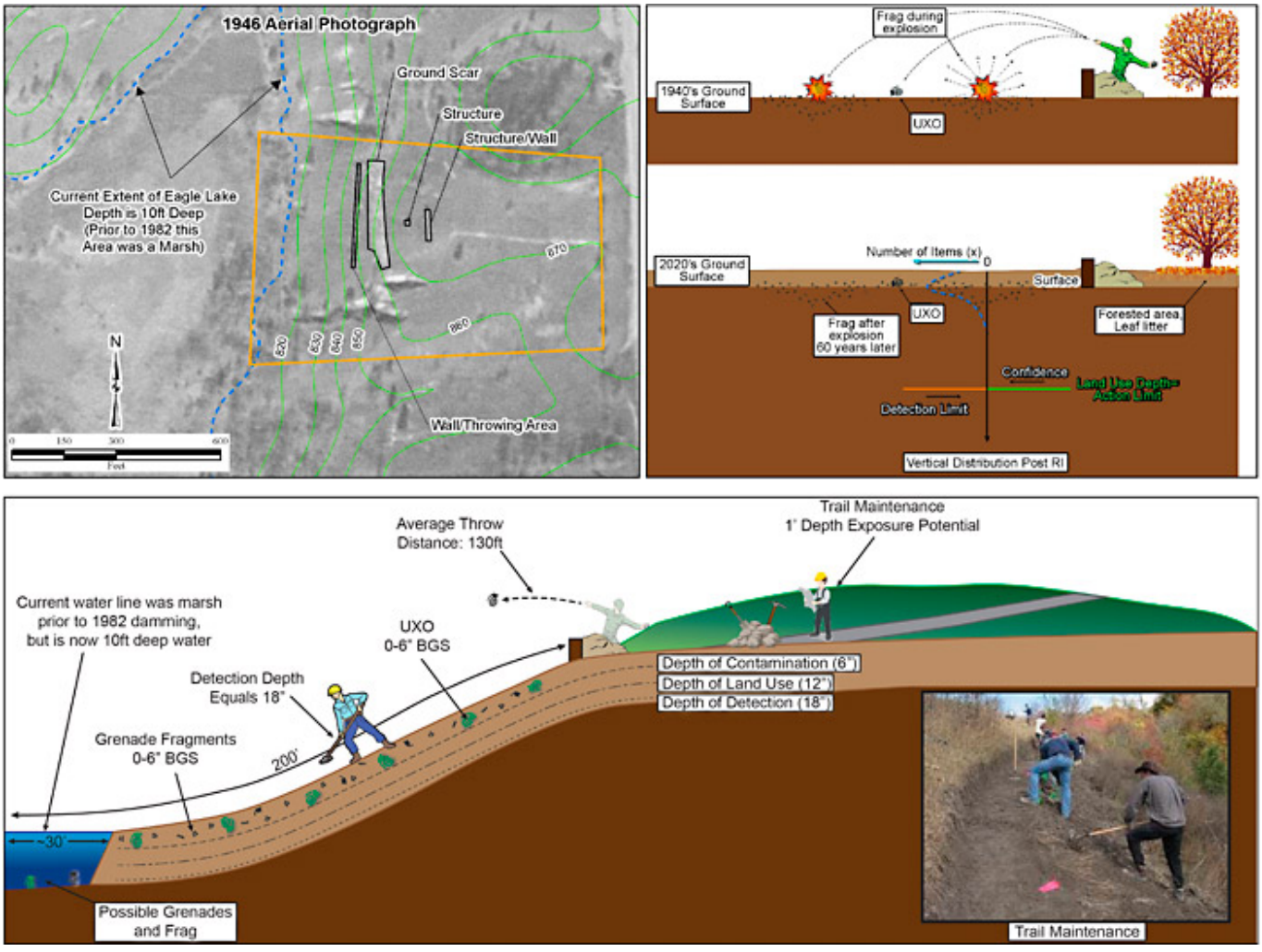


Figure 3-1. Example of a pictorial CSM.

Attributes of a high-quality MR CSM are type of munitions, extent (horizontal and vertical distribution) and concentration of UXO/DMM and MD, as well as past, current, and anticipated future land use, exposure pathways, and receptors. Depending on the phase of the MR, the CSM may be based on historical information or on investigation derived data. The CSM should be updated throughout the process of MR as additional information becomes available and initial assumptions are confirmed or refuted.

The PDT uses the CSM within the Systematic Planning Process to accomplish the following:

1. Identify what is known and unknown about the site.
2. Identify and organize decisions.

3. Identify the sources and quality of information needed to make the decisions.
4. Evaluate the quality of those decisions.
5. Identify data needs or data gaps.

The CSM is documented on MR-QAPP Worksheet #10. Therefore, a well-developed CSM provides a summary of everything known about the site characterization, including site history and findings of previous studies, and identifies data gaps of the site characterization. Data gaps do not, however, necessarily equate to data needs required to characterize the MRS.

The CSM should evolve throughout a project and throughout the project life cycle as new data are collected and as site conditions or receptors change. If changes in site conditions or new data are warranted during phases of an MR project, the PDT should reevaluate the CSM for the MRS to determine whether the project approach should be modified and to plan for any future phases necessary. Further information on these methods and their QC requirements can be found in "Environmental Quality-Conceptual Site Models" Engineer Manual EM-200-1-12 (USACE 2012).

The CSM should be consistent with the current and reasonably anticipated future land use. For Federal Land Management Agencies (FLMAs), the land use and management designations are detailed in federal land management plans. FLMAs provide these plans to the PDT along with any detailed land use information that may affect the DQOs, RAOs, and alternatives to be evaluated. Land Use Control Implementation Plans, which inherently involve acceptance by the landowner, should be negotiated and agreed upon to ensure that remedies involving land use controls will remain protective.

3.3 Systematic Planning: USEPA Data Quality Objective Process

The USEPA Data Quality Objective (DQO) process is the preferred systematic approach for collecting or assessing information or data for evidence-based decision making. The DQO process is used to establish performance and acceptance criteria, which serve as the basis for designing a plan for collecting data of sufficient quality and quantity to support the goals of the study.

The USEPA DQO process is relevant to all aspects of the work performed where data are required to make a decision. The DQO process is most amenable to an environmental measurement but can be used for any decision based on evidence. The process involves seven steps and yields (step 6) the qualitative or quantitative statements (the DQO). A detailed explanation of the DQO process can be found in *Guidance on Systematic Planning Using the Data Quality Objectives Process,* EPA QA/GA-4.

3.3.1 Data Quality Objectives

DQOs are qualitative and quantitative statements derived from the outputs of the first six steps of the DQO process. These statements clarify the study objectives, define the most appropriate type of data to collect, determine the appropriate conditions from which to collect the data, and specify tolerable limits on decision errors which are used as the basis for establishing the quantity and quality of data needed to support the decision (MR-QAPP Worksheet #11)

Although the USEPA DQO process is a seven-step process, DQOs are determined by the responses to the following three questions (Ramsey and Hewitt 2005):

1. **What is the specific question to be answered by data?**
   This step involves identifying the key questions that the study attempts to address, along with alternative actions or outcomes that may result based on the answers to these key questions. Many of these key questions are derived from the CSM. The question should be stated as specifically as possible. Planners should also identify key unknown conditions or unresolved issues that may lead to a solution to the problem.
2. **What is the decision unit (DU)?**
   The DU represents the scale of observation necessary to make the decision. The DU is the specific area, grid, or demarcated extent from which the sample is collected and to which analytical results apply for decision making. There may be one or multiple DUs within an MRS. For example, if the activity is to locate former target areas, the entire MRS may be the DU. Or if the activity is to characterize an area of high anomaly density within the MRS, the high anomaly density area becomes the DU or if the objective is to classify an anomaly as TOI, then the anomaly becomes the DU.
3. **What is the desired decision confidence (level of uncertainty)?**
   When technically feasible, an expression of statistical uncertainty for the decision is desirable because it is considered more objective. But mathematical treatment of uncertainty may not always be technically feasible or

necessary (see Uncertainty). Qualitative expressions of decision confidence through a weight-of-evidence approach may well be sufficient, and in some cases, may be the only option available (USEPA 2001a). Confidence is also related to the underlying quality criteria identified for the project.

3.3.2 Data Quality (Measurement Performance Criteria, Data Quality Indicators, Measurement Quality Objectives, and Quality Control)

To effectively implement DQOs during the project life cycle, planners must define and document qualitative and quantitative requirements and acceptance thresholds and limits for these requirements. These elements are defined during the systematic planning process and documented in the UFP-QAPP as part of that process. This section presents definitions of the elements that comprise DQOs. Figure 3-2 presents the progression of DQO development and the corresponding QAPP Worksheets.

Measurement Performance Criteria – MPC are quantitative measurement performance criteria used to guide the selection of appropriate types of data and measurement. MPCs are developed to ensure collected data satisfy DQOs. MPCs are stated in terms of data quality indicators (DQI) and measurement quality objectives (MQOs). MPCs signify key quality indicators needed for successful decision making at the completion of the project and are documented in MR-QAPP Worksheet #12.

Data Quality Indicators – DQIs are qualitative and quantitative measures of data quality attributes. DQIs include precision, accuracy, representativeness, comparability, completeness, and sensitivity (PARCCS). For purposes of consistency, the following terms are used throughout this document.

- *Precision* – the measure of agreement among repeated measurements of the same property under identical, or substantially similar, conditions.
- *Bias* – systematic or persistent distortion of a measurement process that causes errors in one direction.
- *Accuracy* – a measure of the overall agreement of a measurement to a known value.
- *Representativeness* – the measure of the degree to which data accurately and precisely represent a characteristic of a population, parameter variations at a sampling point, a process condition, or an environmental condition.
- *Comparability* – a qualitative term that expresses the measure of confidence that two or more data sets can contribute to a common analysis.
- *Completeness* – a measure of the amount of valid data obtained from a measurement system.
- *Sensitivity* – the capability of a method or instrument to discriminate between measurement responses representing different levels of the variable of interest.

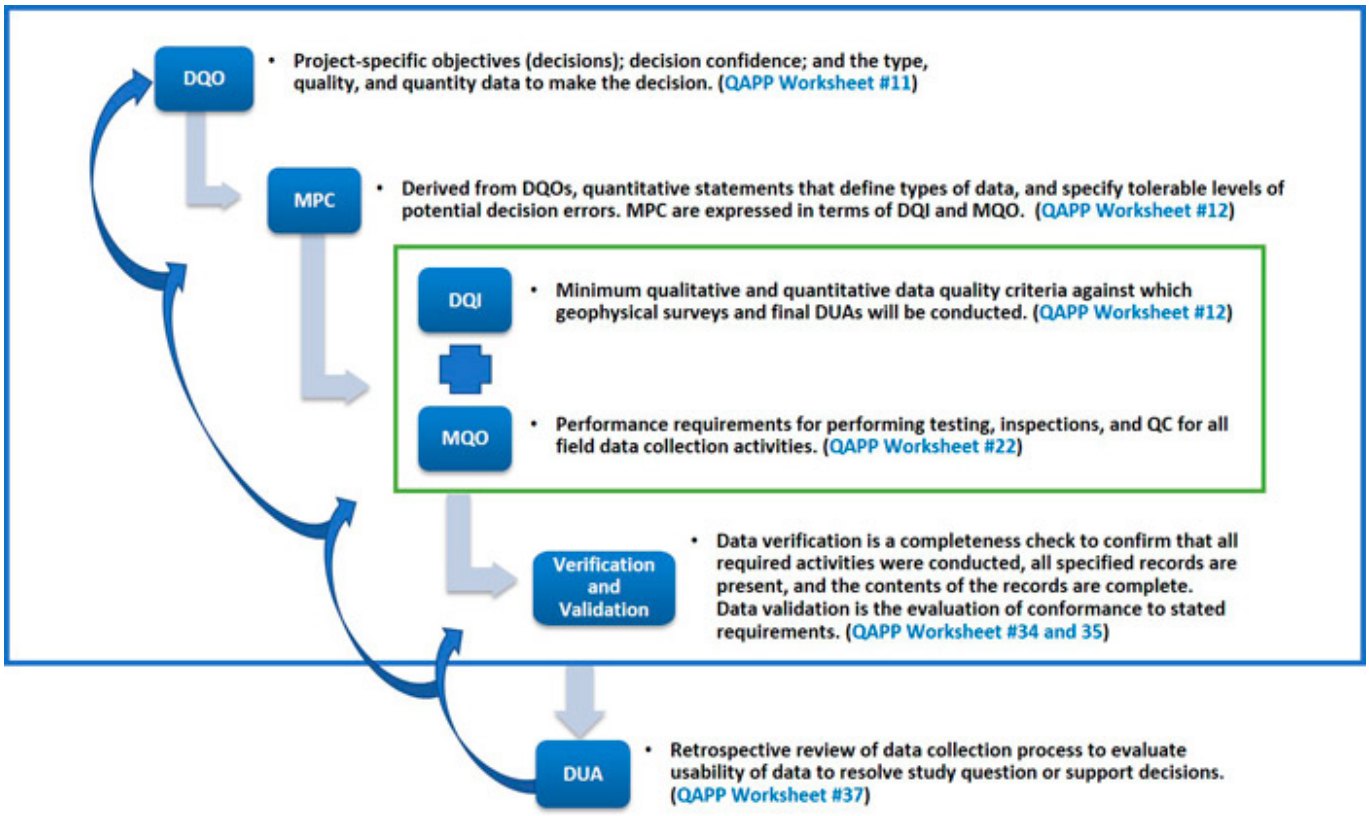


Figure 3-2. Elements of DQOs and the progression of their development.

Measurement Quality Objectives (MQO) – MQOs are the performance requirements for specific DQIs. MQOs specify how good the data must be for confident *decision making*. MQOs are the acceptance thresholds or limits for the collection or analysis of data, based on the MPC, which are derived from the DQOs. MQOs provide QC specifications for measurement processes, including analytical methods, designed to control and document measurement uncertainty in data. MQOs establish QC requirements for all field data collection activities. References to the applicable process (DFW) or SOPs are therefore an integral part of MQOs. Failure of an MQO can be corrected, but the failure response must include a RCA to determine the appropriate corrective action and these concepts are summarized in Appendix A. MQO's are specified on MR-QAPP Worksheet #22.

Verification and Validation – Data verification is a completeness check to confirm that all required activities were conducted, all specified records are present, and the contents of the records are complete. The data validation process evaluates whether data conforms to stated requirements. MR-QAPP Worksheets #34 and #35 define the required documentation and establish the procedures to support these processes.

Data Usability Assessment – The data usability assessment (DUA) is performed by key members of the PDT upon completion of data collection activities for activities within an investigation (the detection survey, the cued survey, and the intrusive investigation) before proceeding to the next phase. The DUA includes a qualitative and quantitative evaluation to determine whether project data are of the right type, quality, and quantity to support the MPCs and DQOs specific to that activity of the investigation. This evaluation also includes a retrospective review of the systematic planning process to evaluate whether underlying assumptions are supported, sources of uncertainty in data have been managed appropriately, data are representative of the population of interest, and the results can be used as intended with an acceptable level of confidence. MR-QAPP Worksheet #37 defines and documents the DUA process.

3.4 Uncertainty

Defining uncertainty is an integral component of the DQO process. The purpose of acquiring and assessing the quality of evidence is to make decisions. The reliability of a decision depends on the quality of the evidence, and the quality of the evidence depends in part on the uncertainties inherent in the evidence. Assessing evidence and deciding, without considering uncertainties, is futile. Likewise, presenting the results of an assessment without giving information about the uncertainties involved makes the results essentially meaningless (Ramsey 2007).

Incorrect estimates of uncertainty or ignoring uncertainties altogether may result in wrong decisions (false positives or false negatives). Wrong decisions are mistakes that may have severe consequences. Information about uncertainty allows decision makers to judge the quality of the evidence and its fitness as a basis for decision making.

Effective procedures are crucial for estimating and reporting uncertainties arising from processes used to acquire evidence. Uncertainties that arise during the assessment should be described and quantified if possible. Best practices for this process vary, depending on the issue and the assessment methods used. In some cases, rigorous statistical methods can be used, especially where the main uncertainties relate to random sampling or measurement error. In other cases, qualitative methods are more appropriate.

3.4.1 Qualitative Uncertainty

Estimates of uncertainty in qualitative evidence are based on the source of the evidence. In the hierarchy of sources, empirical qualitative data has less uncertainty than historical information, and primary historical sources have less uncertainty than secondary information. The purpose of the hierarchy is to provide a basis for assessing uncertainty and consequently the quality of decisions supported by qualitative information.

For example, decisions such as establishing the boundaries of an MRS are often based on historical information alone. However, the consequences of incorrectly locating the MRS are severe. If this decision is wrong, the focus and the results of subsequent investigations may be hopelessly biased and lead to a wrong decision. Because the consequences of being wrong are severe, this decision should be based on the highest quality (lowest uncertainty) historical information available (such as aerial photos and range records). Moreover, any decision based solely on historical information should be considered preliminary until confirmed with empirical data. Because of the uncertainty associated with historical information, empirical data is required to validate the information for confident decisions.

Additional details on the historical records review process and data sources are provided in the ITRC document Munitions Response Historical Records Review (UXO-2) (ITRC 2003).

3.4.2 Quantitative Uncertainty (Measurement Error)

Quantitative uncertainty is the error in a measurement due to random and systematic effects on the measurement process and is quantified as precision and bias, respectively. The combined estimate of precision and bias errors in the measurement is the total error/uncertainty in the measurement. Total measurement error is the most important single parameter that describes the quality of measurements, because uncertainty fundamentally affects the decisions that are based upon the measured result. Too much uncertainty or error in the result increases the chance of a wrong decision.

3.4.2.1 Random Error

Random error is a component of uncertainty which, during many measurements of the same characteristic, remains constant or varies in an unpredictable way (ISO 3534-1: 3.9). Random errors are present in all measurements. Unlike systematic errors, random errors are unpredictable, which makes them difficult to detect. Random errors, however, often have a Gaussian (normal) distribution and in such cases, statistical methods may be used to analyze the data and estimate error. There are several measures of random error, including but not limited to:

- confidence intervals
- standard error
- coefficient of variance
- P values

Random error can be subdivided into sampling errors and analytical errors. The estimate of random sampling and analytical error is made by replicate sampling and analysis. Traditionally, these errors have been quantified as sampling and analytical precision.

Random error occurs for a variety of reasons, including:

- too few samples or transects
- environmental conditions such as temperature and humidity, which can be controlled in AGC work by, for example, taking background reading every two hours
- equipment sensitivity, such as when an instrument is not be sensitive enough to respond to or indicate a change in some quantity or effect, or the observer may not be able to discern the change

- equipment noise (electrical noise or static)
- imprecise requirements

Precise measurement requirements are required to avoid ambiguities in interpretation. It is necessary to state the measurement quantity (length, mass, concentration, amplitude), but also important to specify the decision unit to which the measurement applies (see DU).

Methods to reduce random error (imprecision) include:

- Collect more samples.
- Make multiple measurements.
- Use the best measuring equipment available.
- Establish precision requirements for measurements.

3.4.2.2 Systematic Error (Measurement Bias)

Systematic error is a component of uncertainty which, during many measurements for the same characteristic, remains constant or varies in a predictable way (ISO 11074-2).

Systematic error deviates by a fixed amount from the true value of measurement. Systematic errors are difficult to detect but, as opposed to random errors, easier to correct. Like random error, systematic error can be subdivided into both analytical/measurement and sampling error. These sources of errors are usually quantified as measurement bias (difference between measured value and true value) and sampling bias, although the quantification of sampling bias is difficult if not impossible.

The wrong instrument, a poorly calibrated instrument, and operator error can all produce analytical bias. Analytical/measurement bias can be estimated by measuring the difference between the measurement and true value on well-matched certified reference materials (such as ISO standards), or by taking it directly from the validation of the analytical/measurement method.

Eliminating bias in sampling is critical. Sampling bias results from selecting nonrandom or nonrepresentative samples. Estimating and detecting sampling bias is difficult if not impossible. Sampling bias is controlled by adhering to quality systems, using standard practices and procedures, using appropriate and properly calibrated equipment, and employing skilled operators.

Total quantitative uncertainty is the total estimated error from random (precision) and systematic (bias) effects throughout the entire measurement process. Therefore estimating, calculating, and accounting for uncertainty in a measurement can only be accomplished when the entire measurement process and sources of error within the process are completely understood. Once the entire process is understood and uncertainty requirements have been established, QC can be designed and implemented to evaluate and estimate uncertainty throughout the entire process.

The overall objective in estimating uncertainty is to obtain a reliable estimate of the overall uncertainty of measurement. This objective does not require that all the individual sources of uncertainty to be quantified, only that the combined effect be assessed. If, however, the overall level of uncertainty is found to be unacceptable (the measurements are not fit for purpose) then the uncertainty must be reduced (Ramsey 2007).

3.4.3 Uncertainty in Scientific Models

Models are used to project or predict outcomes based on assumed inputs. A scientific model is accepted as valid only after it is tested against data from the real world and the evidence supports its results. Visual Sampling Plan (VSP) is a model used to develop MRS sampling plans using certain assumptions such as the size and shape of target areas, background density, number of potential TOIs per acre and others. AGC models features extracted from subsurface anomalies to predict TOI. Both models have been tested extensively and considered valid.

Model results, however, include uncertainty. To reduce uncertainty of decisions based on models, site-specific data is gathered to see whether the real-world data agree with the predictions. For VSP, after site-specific data is acquired, it is used as VSP inputs to confirm the sampling plan (transect design) develop based on assumptions was sufficient for the project. For AGC predictions, all TOIs are excavated and evaluated to ensure accurate predictions and several non-TOI are dug and assessed to evaluate the accuracy of those predictions. If predictions do not agree with the actual data, sampling plans may be revised, or corrective actions implemented to determine cause of the failure.

Evidence-based decisions require a systematic and rational approach to researching, collecting, and assessing the best available evidence to make logical empirically supported decisions. (USEPA 2001d; USACE 2009, Section 3.3; DON ***in progress***). The best, most reliable evidence is information or data that is fit for decision making. The fitness of evidence can only be judged by having reliable estimates of uncertainty.

3.4.4 Uncertainty and Data Gaps

The overall objective in estimating uncertainty is to obtain a reliable estimate of the overall uncertainty of measurement. This objective does not require that all the individual sources of uncertainty to be quantified, only that the combined effect be assessed. If, however, the overall level of uncertainty is found to be unacceptable (the measurements are not fit for purpose) then the uncertainty must be reduced (Ramsey 2007).

Strategies for managing uncertainties include assessing the quality of available information (see evidence hierarchy and uncertainty, evaluating error (conformance to QC requirements), identifying data gaps, assessing the impact the overall uncertainty has on the pertinent decisions (what if the wrong choice is made?) and if necessary, deciding how to resolve the uncertainty (minimize error) or fill data gaps.

Missing or insufficient information increases uncertainty. Reconsidering and interpreting historical information or preexisting evidence is one way to fill information gaps. Collecting new data, however, may be the only option to fill data gaps. For example, a data gap could be unavailable or incomplete historical aerial photography during periods of suspected munitions use onsite. Analysis of aerial photography during that time frame could support the presence or absence of ground surface features related to munitions use in a specific location such as an artillery or aerial bombing target.

The PDT should also use the CSM to identify data gaps. As the first step of the DQO process, the CSM can identify data gaps and focus the problem statement for the investigation. Regardless, once the PDT decides to assess or reexamine existing data or collect new data, the team implements the DQO process to determine type and quality of data required to fill the data gaps.

Back to Top ■ Next Section