Sampling design part 1 – application

Contaminated Land Guidelines

Draft for consultation

NSW EPA

www.epa.nsw.gov.au

Cover: Contaminated soil being sampled. Photo: Bottlebrush Media/EPA.

Published by:

NSW Environment Protection Authority
4 Parramatta Square, 12 Darcy Street
Parramatta NSW 2150
Locked Bag 5022, Parramatta 2124
Phone: +61 2 9995 5000 (switchboard)
Phone: 131 555 (NSW only – environment information and publications requests)
Fax: +61 2 9995 5999
TTY users: phone 133 677, then ask for 131 555
Speak and listen users: phone 1300 555 727, then ask for 131 555
Email: info@epa.nsw.gov.au
Website: www.epa.nsw.gov.au

Report pollution and environmental incidents
Environment Line: 131 555 (NSW only) or info@epa.nsw.gov.au
See also www.epa.nsw.gov.au

ISBN 978 1 922447 09 8
EPA 2020P2437
August 2020

Contents

Figures

Tables

1. Introduction

The NSW Environment Protection Authority (EPA) has prepared these guidelines to assist contaminated-land consultants, site auditors, regulators, planning authorities, landholders, developers, and members of the public who have an interest in the outcomes of the assessment and management of contaminated land. They will help consultants to design sampling for contaminated sites, with regard to where samples are collected, how many samples are collected, and how the data is compared to relevant criteria: they are intended to help users obtain data that is appropriately representative for the purposes of the sampling and the media being sampled, and to carry out the subsequent analysis and interpretation of the collected data.

As when following any guidance, users should justify the approaches they use, and demonstrate that they are appropriate and fit for purpose.

The guidelines are in two parts. The first part (this document) describes the application of sampling design; the second part provides guidance on interpretation of the results.

1.1. Background

In assessing site contamination, a major objective is to measure the level of contamination by collecting representative environmental samples for characterisation and chemical analysis. The type of sampling carried out, and the methods used to analyse and interpret the resulting data, significantly influence the validity of the assessment.

This document provides specific recommendations and procedures for the consultants and reviewers of site investigations. However, it is not all-encompassing. For methods it does not describe, or for more complex problems, refer to other relevant guidelines and information sources – this document points to many – or consult an environmental statistician.

These guidelines should be used at the beginning of the site investigation, when the preliminary conceptual site model (CSM) has been developed and data gaps for site characterisation have been identified. The next steps are to identify the processes that could have resulted in contamination, the potential contaminants of concern (PCoCs) and the target media for the investigation.

1.2. Scope of these guidelines

Section 2

Introduction to systematic planning, including CSMs and data quality objectives (DQOs). Additional information on DQOs is provided in Appendix A, and a hypothetical worked example is given in Appendix B.

Section 3

General considerations regarding environmental sampling and statistical aspects in site contamination assessment.

Section 4

Objectives of sampling programs, including a discussion of the correspondence of characterisation and validation, modes of contamination and sampling objectives.

Section 5

Main sampling strategies and considerations for environmental media, including soil and fill, stockpiles, groundwater, surface water and air, along with information for determining background conditions in various media.

Section 6

Methods for detecting significantly elevated concentrations of contamination (that is, hotspots). Appendix C provides methods for determining sampling grids for hotspot detection, including the recommended grid sizes for characterisation using a systematic sampling pattern.

Section 7

The number of samples required, including existing guidance and statistical tests for determining the number of samples, using the combined risk value (CRV) method and the maximum probable error (MPE) method. Appendix D summarises NSW guidance regarding sampling design and Appendix E and Appendix F give procedures and worked examples using the CRV and MPE methods.

1.3. Legal framework, policy and relationship to other guidelines

These guidelines have been made under the *Contaminated Land Management Act 1997* (CLM Act). They should be read in conjunction with the CLM Act, the Contaminated Land Management Regulation 2013 (CLM Regulation), and any guidelines made or approved by the EPA under the CLM Act.

The guidelines complement other guidelines made by the EPA, and several national guidance documents that have been approved by the EPA. Those guideline documents are listed in the reference section and are specifically referenced in the text, where appropriate.

1.4. Environmental media

These guidelines address the sampling of soil and solid media, as these are the most common targets in the assessment of site contamination. They also provide information about other media, including groundwater, surface water, sediments and air. Most of the statistical procedures described in these guidelines can be applied to all these media. General advice is provided regarding sampling for emerging contaminants, along with specific references.

This document does not specifically address biota sampling and ecotoxicity testing. For these specialty areas, see the following references:

- Australian and New Zealand Governments and Australian State and Territory Governments (ANZG) (2018) *Australian and New Zealand Guidelines for Fresh and Marine Water Quality*, Water Quality Australia, Canberra ACT. Available at: www.waterquality.gov.au/anz-guidelines.
- Department of Environment and Science (DES) (2018), *Monitoring and Sampling Manual: Environmental Protection (Water) Policy 2009*, Queensland Department of Environment and Science, Brisbane.
- Department of Environment and Conservation (DEC) (2004*), Australian River Assessment System (AUSRIVAS) Sampling and Processing Manual 2004*, NSW Department of Environment and Conservation, Sydney.

2. Systematic planning

A systematic planning process should be used to define the objectives of all site assessment and remediation programs, and to develop sampling, analysis and quality plans (SAQPs) for the collection and evaluation of representative data to achieve those objectives. The *National Environment Protection (Assessment of Site Contamination) Measure* (NEPM 2013, B2) recommends the use of conceptual site models (CSMs) and data quality objectives (DQOs) for systematic planning.

2.1. Conceptual site models

Conceptual site models provide a spatial and temporal overview of the contamination at sites and their surrounds, highlighting the contaminant sources and potential receptors, and the exposure pathways between the sources and receptors. Robust CSMs should include the known and potential:

- sources of contamination and contaminants of concern, including the mechanism(s) of contamination, for example 'top down' spills, placement of fill, sub-surface release, etc
- affected media, such as soil, sediment, groundwater, surface water, soil vapour and air quality (indoor air and ambient air), both on- and offsite
- human and ecological receptors, both on- and offsite
- complete exposure pathways, both on- and offsite.

CSMs should logically explain the existing information, evidence, and data from the area under study, and be predictive. Where CSMs have poor predictive capabilities, the supporting information and evidence should be reviewed, and the CSM appropriately revised and updated, stating any data gaps.

After identifying the potential contaminants of concern (PCoCs), consideration should be given to their physico-chemical properties, such as solubility in water, volatility, miscibility and interactions with environmental media. This is especially important when considering uncommon or emerging contaminants.

When conducting a preliminary site investigation (PSI), the available environmental information and site history information should be synthesised into a CSM. At every subsequent stage of site assessment, the CSM should be refined, with the information and data from each investigation stage. Each refined CSM should be used to inform subsequent decisions on the condition of the site or area under study.

A CSM should identify uncertainties and data gaps in relation to the contamination and the potential exposure pathways. Any theories or assumptions underlying CSMs should be clearly identified to ensure adequate transparency. CSMs should address:

- how representative the available data are likely to be
- the potential sources of variability and uncertainty
- how important the identified gaps are to the objectives and reliability of the site assessment.

CSMs can take various forms, including text, tables, graphics, and flow diagrams. They can also take the form of site-specific plans and figures, including cross-sections.

While statistical analysis can provide a quantitative basis for decision making, the assessment of site contamination relies on a multiple lines of evidence/weight of evidence approach, and other critical lines of evidence, including site histories, field samples, and geological and hydrogeological data and information. This approach allows scientifically defensible decision making with robust CSMs assisting in this process.

2.2. Decision makers

Various decision-makers can have an interest in the outcomes of site contamination assessment projects and investigation results, including:

- the consultant team
- clients, landowners, property developers

- accredited site auditors
- planners and other technical specialists
- regulators, including the local government, EPA and other state government bodies
- other relevant stakeholders, such as adjacent landholders, the local community, and non-government organisations.

In the planning and reporting of site contamination assessment projects, consultants should recognise that other decision makers are, at times, not technical specialists. Therefore, the methods used in the collection and analysis of site contamination assessment sampling data should be clearly documented and discussed.

In all instances, clear and appropriate explanation and justification of the implemented sampling program should be provided, including the benefits of the approach selected, along with all assumptions, limitations and remaining data gaps. Importantly, where the sampling program deviates from made or approved guidance, this should be clearly articulated, including the rationale and justification for any such deviations.

2.3. Modes of contamination

In the assessment of site contamination, understanding the mode of contamination affecting the site is very important. The duration of the spill or leak, volume of contaminant lost, contaminant type and nature, the sub-surface material and whether preferential pathways are present, will all affect the distribution of the contamination.

Example modes of contamination include:

- filling or emplacement of materials (from on- or offsite) from areas with unknown contamination issues: historical industrial waste from combustion furnaces or waste products, fill sourced from agricultural lands, building and demolition wastes or abandoned production materials
- heterogeneous filling: if fill has been sourced from a number of different unknown sources, the site may vary in its spatial distribution of matrices and contaminant levels in ways that are not predictable
- top-down contamination: a leak or spill of a substance occurring on the surface of the site and infiltrating down through the sub-surface, from sources such as above-ground tanks, drums, direct application of liquid wastes and spent liquors, transfer systems or vehicles
- subsurface leaks: contaminant losses from sub-surface infrastructure such as underground petroleum storage systems (UPSSs), trade-waste systems, septic tanks, sumps, pits, transfer lines or pipelines
- in-situ contamination: similar to sub-surface leaks, but relating to contamination already located within the sub-surface. Examples include a leachate plume emanating from a landfill or contaminated soil, or phase-separated hydrocarbon (PSH) in the vadose zone, both of which can contaminate groundwater.

The modes of contamination should be considered in the investigation objectives and discussed in the CSM.

2.4. Data quality objectives

The DQOs process is used to develop performance and acceptance criteria (or data quality objectives) that clarify study objectives, define the appropriate type of data, and specify tolerable levels of potential decision errors. These criteria are used as the basis for establishing the quality and quantity of data needed to support decisions. EPA policy is that DQOs must be adopted for all assessment and remediation programs, and that the process must be conducted before any investigative works begin (EPA 2017; NEPM 2013, B2).

Developed as part of the environmental data life-cycle process by the United States Environmental Protection Agency (USEPA), the seven-step DQOs process is a method for systematic planning that includes options for the type of problem to be addressed, based on the intended use of the data to be collected. The two primary types of intended use are classified as **decision making** and **estimation**.

The DQOs process is further described in Appendix A. Refer to USEPA (2000, G-4HW; 2006, G-4) for details of the process for collecting environmental data. Appendix B gives a worked example of a hypothetical investigation-level decision problem.

3. Environmental sampling considerations

Not all of a population can be measured, and a collection of measurements or observations is made as a sample of the population. In the assessment of site contamination, populations commonly include such things as the soil at a site or in a decision area, all of the fill in a stockpile, all of the gas in the soil, or all of the groundwater beneath the site. The characteristics determined from the sample are then used to provide information regarding the population.

However, the sampling of environmental media presents unique challenges for measurement due to matrix interferences, large-scale spatial variation, small-scale variations from matrix heterogeneity, and the generally small number of measurements made relative to the media being assessed. Because of this, both multiple lines of evidence and weight of evidence approaches must be used in the assessment of site contamination, to synthesise the physical and numerical information that characterises a site and its surrounds. The CSM and the associated data-gap analysis are the key tools for this synthesis. In following this process, it is imperative that the reporting includes the full physical and numerical dataset and that all methodologies are documented and explained, including all assumptions and any associated limitations.

3.1. Types of samples

In the assessment of site contamination, a sample is usually a physical object: it can be a jar of soil, a cannister of soil gas, a bottle of water, an individual specimen of biota, etc., that can be chemically analysed at a laboratory for the PCoCs. The term ‘sample’ can also refer to visual and olfactory observations, descriptions and field logging, which can be field-screened and then subject to other non-laboratory assessments and tests. These are known as **field** samples.

Any sample that is sent to a laboratory to be analysed is known as an **analytical** sample. An analytical sample is a field sample, but a field sample may not necessarily be an analytical sample[1]. In statistics, ‘sample’ is also used to mean **n**, the number of samples or individual measurements.

Statistical analysis and inference with prescribed error rates is done mainly with analytical samples. Under the multiple lines of evidence/weight of evidence approach, field samples are critical to inform the CSM and assist in defining the sources and pathways. The number and type of analytical samples is determined by the CSM and statistical determinations, and the iterative nature of the process requires assigning an appropriate, but variable, weighting to the available evidence from both types of samples.

[1] When field samples are collected, some may not be analysed immediately. Those samples must be analysed within the laboratory holding times for extraction and analysis for the various contaminants, or new samples may have to be collected.

4. Objectives of sampling programs

Clear definition of sampling objectives is essential to developing a sampling strategy, as this influences the sample types, the sampling pattern adopted, and the number of samples taken. In general, information is being sought as to the type, location, extent and severity of the contamination, which often requires comparison to relevant threshold values. In instances like this, the objective of the sampling is to enable decision-making.

The specific objectives of any sampling program will need to be defined on a case by case basis, depending on the project level objectives, the CSM, the media to be assessed and the stage of the project. The NEPM (2013) describes that:

> *The purpose of site assessment is to determine whether site contamination poses an actual or potential risk to human health and the environment, either on or off the site, of sufficient magnitude to warrant remediation appropriate to the current or proposed land use.*

The NEPM (2013) also notes that adequate site characterisation is the foundation for appropriate assessment of health and environmental risks associated with site contamination.

4.1. The process of assessing site contamination

The assessment of site contamination generally includes sequential stages of assessment and management, shown in Table 1, along with types of environmental sampling conducted at each stage. See Section 5.2 for more details of 'probabilistic' and 'judgmental' sampling.

Table 1 Site contamination assessment investigation stages and associated sampling

Investigation stage	Type of sampling
Preliminary site investigation (PSI)	Occasionally sampling is performed but it is generally limited to judgmental sampling of soil, fill, and/or surface water.
Detailed site investigation (DSI)	Both judgmental and probabilistic sampling are performed, commonly of soil, fill and groundwater, but sometimes also of soil gas, indoor air, ambient air, surface water and sediments.
Implementation of the remedial action plan (RAP)	Includes sampling for compliance monitoring, which is generally judgmental, and waste classification, which is probabilistic. Also includes investigations of unexpected finds uncovered during the physical works, which can include probabilistic and judgmental sampling.
Validation investigation	Conducted using probabilistic sampling for broad areas and judgmental sampling for validating beneath former structures or within excavations, tank pits, trenches, etc.
Ongoing monitoring (if required)	Targeted to specific locations, such as sentinel groundwater wells or air monitoring in basements, as the extent and magnitude of contamination has been identified in a previous stage.

Specialist studies may also be required as part of the site contamination assessment process, for instance to provide data for human health or ecological risk assessments, assessment of the broader environment adjacent to and/or down-gradient from the site, and as part of the remedial design. While the assessment is usually represented as sequential steps, often the steps consist of multiple, overlapping investigations. For example, soil sampling can lead to further delineation of the extent of the contamination and potentially also groundwater sampling or soil gas sampling, which can lead to further soil sampling to close subsequently identified data gaps.

4.2. Characterisation and validation

Guidance has traditionally made a clear distinction between **characterisation** and **validation**. While this may be appropriate in some circumstances, it should be recognised that there is no practical distinction between a final characterisation sample and a final, post-remediation, validation sample, when they are both taken as the final sample which concludes that a sample location is below any specific criterion or action level. Accordingly, the required quality of both representativeness and usability for final characterisation samples and final validation samples should be identical.

It should also be noted that where assessment and remediation projects occur over extended periods of time, areas that were characterised as suitable for the proposed uses must be maintained throughout as being suitable. If subsequent uses occur which have the potential to cause contamination, such as stockpiling of potentially contaminated material, uncontrolled dumping of wastes, or ongoing industrial use, then further characterisation or validation will be required. Similarly, if there are significant information gaps in the site history between characterisation and proposed changes in site use, then further characterisation or validation will be required.

4.3. Sampling objectives

Project objectives are broad: for example, to determine if a site is suitable for a specified land use. Sampling objectives, however, need to be very specific and concise, and set out the media to be sampled, the PCoCs and the principal study question (including the possible outcomes resulting from the study question).

As sampling objectives are necessarily situation-specific, it is not possible to be prescriptive about objectives and sampling designs, although the typical objectives of a sampling design for a site contamination assessment are to:

- characterise the nature and extent of contamination at a site
- characterise soil, fill, stockpiles or waste materials for waste classification
- assess whether contamination levels exceed a criterion or action level
- determine the background condition of a specified media
- determine if contaminant concentrations significantly exceed background
- determine whether certain characteristics of two populations differ by some amount
- estimate certain population parameters such as mean, variation or the 95th (or greater) percentile
- identify the location of hotspots of a specified size, or provide evidence that they do not exist within specified confidence limits
- delineate groundwater or surface water plumes
- identify if a preferential pathway exists
- determine if offsite impacts have occurred to any media
- determine if identified contaminants pose a human-health or ecological risk.

The sampling objectives should be defined as part of the DQOs process, and clearly documented. Analysis and interpretation of the resulting data should be conducted in the context of the defined sampling objectives. More detail on the DQOs process is provided in Appendix A, and a worked example is discussed in Appendix B.

5. Sampling design

Following an overview of the categories of sampling design, broad sampling strategies are discussed which may be applied to all media in Section 5.2, with media-specific information in Sections 5.3 to 5.8.

5.1. Probabilistic and judgmental sampling design

There are two main categories of sampling design; **probabilistic** sampling and **judgmental** sampling.

A **probabilistic** sampling design is one that uses random selection (that is, the different units in the population under study have an equal probability of being selected). This type of design, properly applied, results in unbiased and independent data. The advantages of probabilistic sampling designs are that they:

- enable statistical inferences to be made
- provide the ability to calculate uncertainty associated with estimates
- provide reproducible results within uncertainty limits
- produce decision error criteria that are incorporated into the interpretation and presented in results (usually as confidence statements).

However, for an optimal design using probabilistic sampling, an accurate CSM is required, including a clear definition of the population to be sampled.

Judgmental sampling, also called targeted sampling, involves making a specific decision about where and/or when to collect the samples. In the assessment of site contamination, it relies on good site histories and/or site features being clear and distinct. While it can be an efficient method for determining the areas of worst-case impacts, or is useful where the site history is inadequate or the features of concern are obscured or not discernible, if undertaken using an underdeveloped CSM it provides poor quality data for site characterisation and should not be solely relied upon. Judgmental sampling relies on a high level of experience and expertise to both choose the sampling locations and to subsequently interpret the resulting data.

Statistical determinations relating the sample data to the population parameter, such as estimating confidence intervals or conducting hypothesis tests, are only valid if the sample data is unbiased and independent. Consequently, data collected using judgmental designs are not suited for use in statistical determinations.

Where judgmental samples **are** used for statistical determinations, if they are targeted to areas of contamination (fill material, stained and odorous soils, impacted groundwater, etc.), the resulting data will probably be biased upwards, increasing the likelihood of a Type II decision error, i.e. the site will be determined to be more contaminated than it really is. If you do use biased data for a statistical determination, you should clearly document this, and identify and discuss the ramifications and limitations. It is recommended that the results from judgmental sampling and probability sampling are treated as two different populations.

5.2. Sampling strategies

The sampling strategies that are generally employed in the assessment of site contamination include:

- judgmental (i.e. targeted) sampling
- systematic sampling
- random sampling
- stratified sampling.

Determining site contamination involves two main tasks: first, delineating the spatial properties of the environmental medium, stratum or decision area of concern, and second, characterising the physical properties and chemical concentrations of the PCoCs for that medium, stratum or decision area. Note that any such characterisation may also have a time-varying component, particularly for waters and air.

A combination of strategies is often used – targeted sampling for known features and/or specific media, and systematic sampling to provide adequate site coverage and data for statistical inference.

For consultants wishing to determine the number of samples required for site characterisation or site validation as a function of variance in the dataset and confidence levels, refer to Section 7.

5.2.1. Judgmental sampling

Judgmental sampling is also called targeted sampling.

Judgmental sampling points are selected based on the investigator's knowledge of the probable distribution of contaminants at the site, with known or suspected areas of contamination being specifically targeted, based on the CSM. It is an efficient sampling method that makes use of site history and field observations, but it has the disadvantage of being statistically biased. The quality of the resultant data depends in part on the experience and judgment of the consultant, and the available site history information and observable site features.

Judgmental sampling also tends to result in uneven distributions of sampling locations, in which case additional sampling locations are required to provide site coverage. Judgmental sampling should not be used as the only method for site characterisation, unless detailed documentation of the history and site information with high integrity exist and can be provided to support the decision.

Judgmental sampling is recommended to validate the remediation of solid media and the removal of infrastructure such as UPSSs (underground petroleum storage systems). The number of judgmental samples taken is determined in part by the number and size of potential identified sources, and the number and area of observable features, such as staining, odours, wastes, extent of fill material, etc. Sample locations may also be targeted along potential migration routes of surface drainage or permeable materials.

Judgmental sample results must be reported in separate results tables from other results.

As contamination concentrations can vary greatly over short distances, single judgmental samples may not provide a complete understanding of the potential contaminant range. Where areas of contamination are identified, further 'step-out' and depth sampling is often required, to determine if the likely maximum contaminant levels have been identified. Further sampling should seek to identify if the concentrations are increasing or decreasing away from the already identified contamination, noting that sharply defined boundaries rarely exist.

For environmental media other than soil, the use of judgmental sampling designs is often required, as it is often not possible to sample these media at random locations and times. For groundwater, soil vapour and ground gas studies, consultants should seek to target all of the potential sources, receptors and pathways. Random sampling locations, while being more defensible statistically, essentially serve no purpose unless a large (and generally cost-prohibitive) number of samples is taken. See DEC (2007) and for further guidance.

For surface waters, sampling often targets specific locations, such as upstream and downstream of a site located on a watercourse. Unless the watercourse or body is particularly large, randomisation of sample locations is impractical. Generally, there is a high degree of natural mixing and homogeneity in surface waters, but stratification can occur in water bodies. Sampling design should ensure appropriate controls, to minimise unrepresentative samples. Refer to ANZG (2018) for more information on sampling surface waters and sediments.

Figure 1 below shows an example of a judgmental sampling pattern, with sampling point locations concentrated around a potentially contaminating object.

Figure 1 Judgmental sampling pattern example

Source: Dyllen Redman/NSW EPA

5.2.2. Systematic sampling

Systematic sampling is a probabilistic strategy that involves selecting points at regular intervals. It is statistically unbiased as long as the coordinates of the first sampling point are determined by uniform random allocation.

Alternatively, systematic random sampling can be used whereby a convenient site feature or boundary can be selected to establish a grid. The sample-location coordinates within each cell are then selected using uniform random allocation between zero and the grid cell size in each dimension. This design can often be more practical to apply at operational sites or where significant infrastructure exists, where selected locations are blocked. Further coordinates can be generated using the same randomisation process.

Whichever approach is used, the randomisation should be as discussed in .

In the assessment of site contamination, systematic sampling is usually done over a grid, although transects may be appropriate when lineal features are being assessed, such as the validation of former pipeline trenches. Gilbert (1987) notes that uniform coverage in many cases yields more accurate critical parameters of a contaminant distribution, such as the mean. The NEPM (2013, B2) states that “systematic and grid sampling is used to search for hotspots and to infer means, percentiles or other parameters”. Systematic sampling does not generate clusters of sampling points but rather ensures an even coverage of the site or decision area, which makes this approach ideal for the characterisation of sites or decision areas.

Grids can be square, rectangular, or triangular; however, square grids are generally used. For regular square grids, the required grid size is set out, and samples are taken at the same location from each cell, ideally the centre of the cell, or as determined by the ‘systematic random’ approach. Square grids have

the advantage of being simple to establish. They have also given adequate results (BSI 10175:2013) in studies that have evaluated the relative efficiencies of various systematic sampling patterns for hotspots of different shapes. For rectangular or triangular grids, consult Gilbert (1987).

The mesh size (the distance between grid cells in both the x and y direction) is related to the size of a hotspot and the required probability of detecting a hotspot of specified size. Where elongated hotspots are expected (possibly due to land slope), differential x and y mesh sizes will help to detect them. Appendix C gives a method for calculating grid size. Section 5.2.5. provides recommended number of samples for systematic sampling.

Figure 2 below shows an example of a systematic sampling pattern, with sampling point locations placed at regular intervals in a square grid across the investigation area.

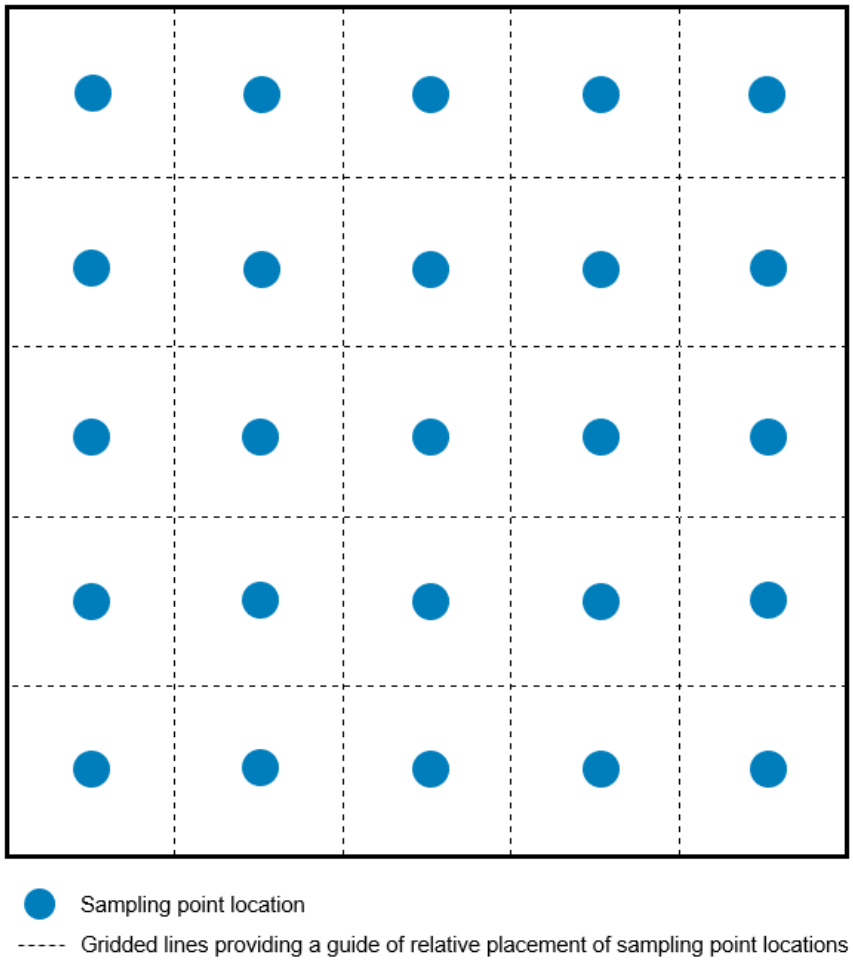

Figure 2 Systematic sampling pattern example

Source: Dyllen Redman/NSW EPA

5.2.3. Random sampling

The NEPM (2013, B2) states that simple random sampling “is most useful when the area of interest is relatively homogenous, and no major pattern or hotspots are expected”. Examples may include specific decision areas for which no information is available, as part of a PSI. Where used in the assessment of site contamination, the limitations of random sampling should be considered and appropriately documented.

With random sampling, sampling points are selected randomly but not arbitrarily[2]. A legitimate uniform pseudo-random number generator, for example a computer program, should be used to determine sampling point coordinates. The randomisation process ensures unbiased data, as any location within the sampling area has an equal chance of being selected as a sampling point.

While random sampling is statistically unbiased, sampling points can cluster together by chance. This makes them deficient for detecting hotspots and for giving an overall picture of the spatial distribution of the contamination. In practice, random sampling has limited use in the assessment of site contamination, unless combined with systematic grid sampling (discussed in Section 5.2.2).

Figure 3 below shows an example of a random sampling pattern, with sampling point locations placed at random locations across the investigation area.

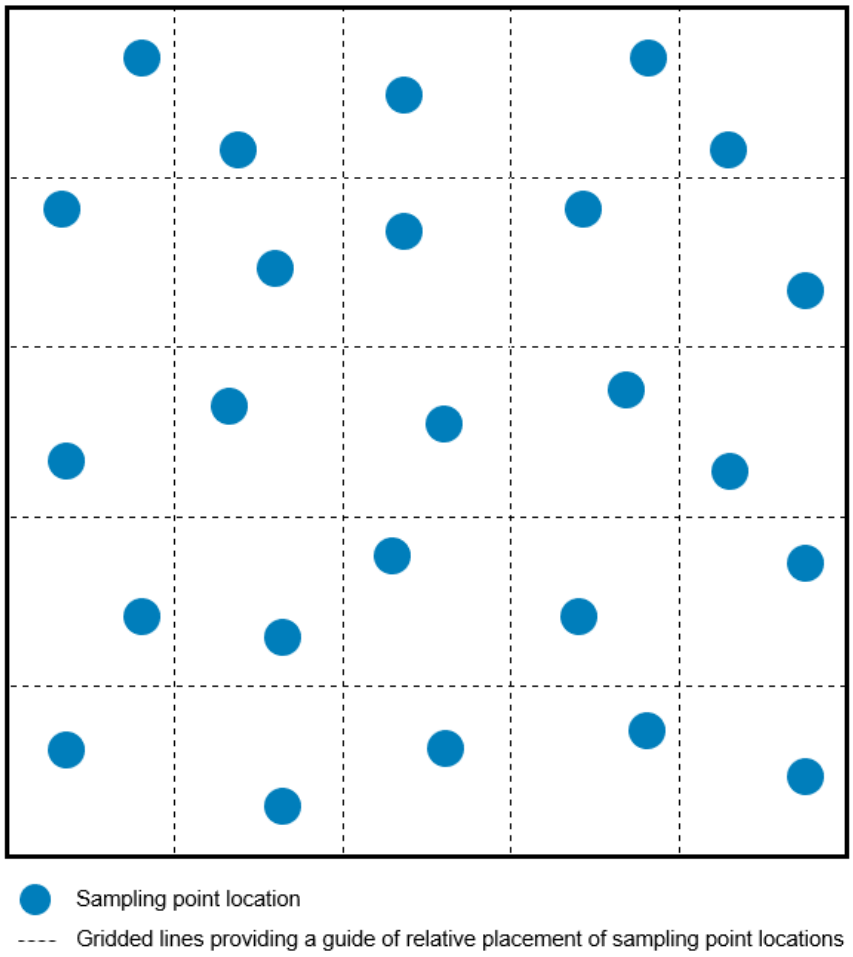

Figure 3 Random sampling pattern example

Source: Dyllen Redman/NSW EPA

5.2.4. Stratified sampling

A stratified sampling pattern may be appropriate for the investigations of large sites with different uses and features, and complex contaminant distributions. Under this approach, the site is divided into various non-overlapping sub-areas, according to geological and geographical features, the nature of the contamination, former usage of the site, or other relevant factors. Each sub-area can then be treated as an individual decision area, and different sampling patterns and sampling densities applied.

[2] Arbitrary samples are also considered judgmental samples, as it is not possible to rule out unconscious bias.

The NEPM (2013, B2) describes the following advantages of implementing a stratified sampling pattern:

- potential for achieving greater precision in estimates of the mean and variance where the measurement of interest is strongly correlated with the variable used to define the strata
- calculation of reliable estimates for subgroups of special interest.

A stratified sampling strategy requires reliable prior knowledge of the site. The sampling strategy used in the individual strata can vary; for example, one stratum might require a targeted sampling strategy while a neighbouring one needs a systematic strategy.

In some cases, stratified sampling patterns may require more complex statistical analysis, as discussed in Gilbert (1987) and USEPA (2006 G–9S).

Figure 4 below shows an example of a stratified sampling pattern, with separate investigation areas and sampling point locations due to different characteristics of the site. Investigation area for group 1 contains an underground storage tank, group 2 is sampling fill material from an unknown source, and group 3 is natural soil with no known contaminating activity on this area. Different sampling strategies and densities have been applied to each area and will be analysed separately.

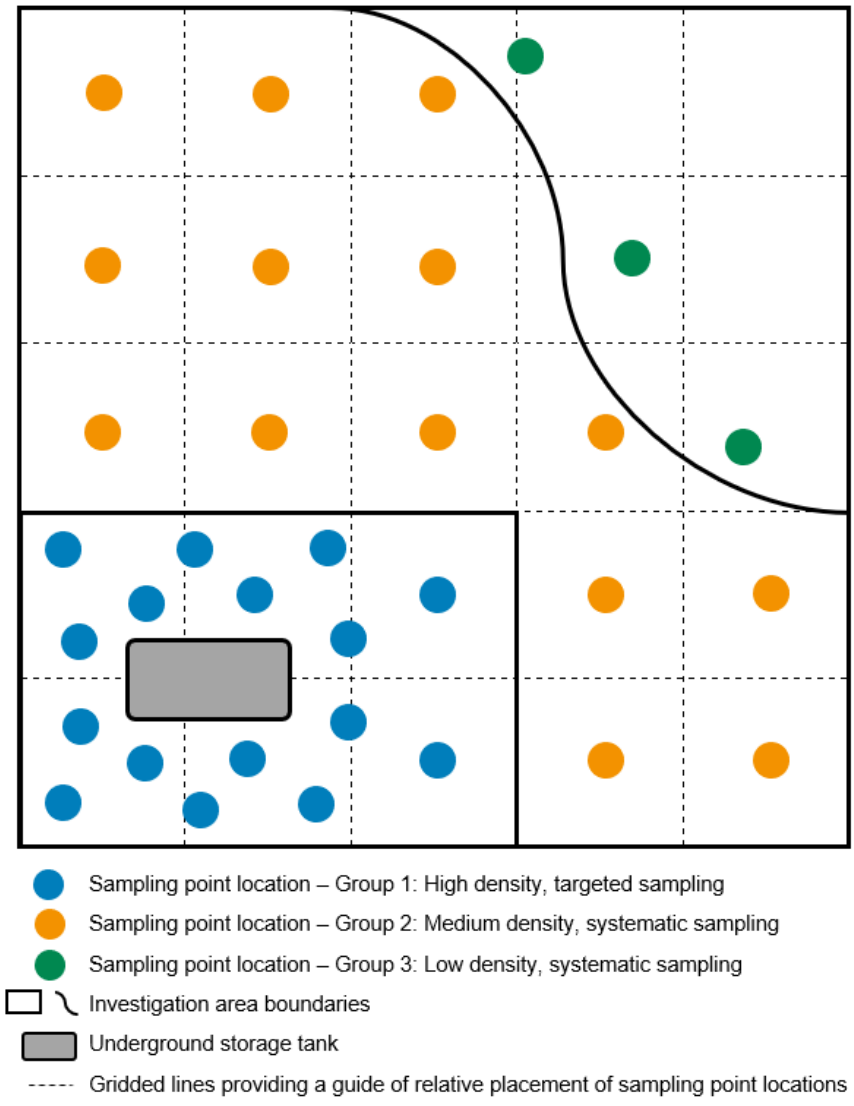

Figure 4 Stratified sampling pattern example

Source: Dyllen Redman/NSW EPA

5.2.5. Recommended number of samples for systematic sampling

Site histories can be incomplete and the locations of buried contaminant sources, such as drums, livestock dips or animal carcasses, might not appear on historical records. In addition, fill of unknown origin can be a source of contamination which can only be identified by intrusive investigations supported by the collection of analytical samples that are analysed by a laboratory.

To account for the lack of or uncertainties in site history information, the EPA recommends using a regular square-grid systematic pattern with a grid size based on the proposed land use (see Table 2). This approach is intended for investigating areas with consistent features. If there are any areas of concern, such as dead vegetation, structures, or evidence of disturbed ground, then the area should be separated and subject to stratified sampling as discussed in Section 5.2.4. If the site history indicates potentially contaminating activities have been carried out at the site, then targeted sampling should be performed in the vicinity of potentially contaminating activities, as well as systematic sampling performed on a grid, as discussed below.

Table 2 Systematic sampling grid size by proposed land use

Proposed land use	Grid size
Residential with garden/accessible soil (home grown produce <10% fruit and vegetable intake, (no poultry), also includes children's day care centres, preschools and primary schools	12 m
Residential with garden/accessible soil (home grown produce <10% fruit and vegetable intake, (no poultry), also includes children's day care centres, preschools and primary schools	12 m
Residential with minimal opportunities for soil access includes dwellings with fully and permanently paved yard space such as high-rise buildings and flats	14 m
Public open space such as parks, playgrounds, playing fields (e.g. ovals), secondary schools and footpaths. It does not include undeveloped public open space (such as urban bushland and reserves) which should be subject to a site-specific assessment where appropriate	16 m
Commercial / industrial such as shops, offices, factories and industrial sites	18 m

Table 2 proposes the maximum grid spacing for four different land use scenarios, and therefore the minimum recommended number of sampling points. Where the proposed land use is unknown, apply a 12-metre grid.

British Standard (2013) suggests a greater sampling density (therefore smaller grid sizes including less than 12 m) should be considered where:

- heterogeneous contamination is indicated, for example on a former gasworks site
- contaminant concentrations identified during an earlier investigation are close to the critical levels of interest, recognizing the uncertainties of measurement in the concentration values
- a high level of confidence is required for the outcome of a risk assessment
- delineation is required along the edges of known areas of contamination
- the averaging area is small.

If there is enough pre-existing information on likely contaminant distribution, the grid size may be increased by up to 2 m for the selected land use scenario. This decision must not be made without appropriate justification, which must be documented in the sampling plan and subsequent report. This allows for a total range of 12–20 m in grid spacing for systematic sampling patterns.

For areas of ecological significance, including wetlands and their adjacent areas, a comprehensive site history must first be prepared, then a conceptual site model developed using desktop information. If field work is required, it should only be undertaken using existing roads for access and hand tools, so as to not impact the ecological values. Consultants should carefully weigh the outcomes of possible remediation against potential damage to the land's ecosystem.

Regardless of the grid spacing selected, areas of 500 m^2 or less need at least five (5) sampling points.

The minimum number of sampling points required for a systematic sampling program has been calculated for a range of site areas in Table 3. The number of samples required must be sufficient to satisfy the following acceptance criteria:

- the arithmetic average concentration of the contaminant(s) must be less than an acceptable limit, at a 95% or higher confidence level
- a site must be free of hot spots larger than a critical size, at a 95% or higher confidence level.

The formula for calculating the 95% upper confidence level of the arithmetic mean can be found in appendices J, K, and L in Part 2 of these guidelines. The formula for calculating the critical size of the hotspot can be found in Appendix C.

Where the number of samples required is not an integer (a whole number), the number of samples must be rounded up.

The formula for calculating the number of sampling locations for specific site areas is included below.

Table 3 Number of sampling locations based on grid size

Area size (m^2)	12 m	13 m	14 m	15 m	16 m	17 m	18 m	19 m	20 m
500	5	5	5	5	5	5	5	5	5
1,000	7	6	6	5	5	5	5	5	5
2,000	14	12	11	9	8	7	7	6	5
3,000	21	18	16	14	12	11	10	9	8
4,000	28	24	21	18	16	14	13	12	10
5,000	35	30	26	23	20	18	16	14	13
6,000	42	36	31	27	24	21	19	17	15
7,000	49	42	36	32	28	25	22	20	18
8,000	56	48	41	36	32	28	25	23	20
9,000	63	54	46	40	36	32	28	25	23
10,000	70	60	52	45	40	35	31	28	25
15,000	105	89	77	67	59	52	47	42	38
20,000	139	119	103	89	79	70	62	56	50

The number of sampling locations based on grid size can be represented by the formula $n = A/G^2$, where:

- n is the number of sampling locations, to be rounded up to the next whole integer
- A is the area of the site or decision area in metres squared
- G is the grid size in metres, i.e. the distance between nodes of grid.

For more information, see:

- Appendix C for the number of samples needed to detect hotspots of given sizes
- Section 7 for how to calculate the number of samples required, based on variance in the dataset and confidence levels
- Appendix E for determining the number of samples by the CRV method
- Appendix F for determining the number of samples by the MPE method.

5.3. Soil and fill material

'Soil' describes the naturally occurring or residual soil that forms due to weathering or geomorphological process. While generally homogenous in the absence of anthropogenic contaminants, Hamon et al. (2004) note that trace elements have naturally high variability. Depending on the associated parent material, metals can be highly variable: for example, ultra-mafic rocks can lead to naturally elevated chromium (Cr) and nickel (Ni). Soil is the primary medium of concern in the assessment of site contamination and has traditionally been the focus for the assessment and management of contaminated land.

Emplaced fill or 'made ground' refers to excavated earthen materials or wastes that have been placed at a site by artificial means, often for the purpose of building up or levelling the surface of a site. Depending on the site, its location and when the site was filled, the material may consist of reworked soils from onsite cut-and-fill activities; overburden material received from offsite locations; industrial wastes such as furnace wastes, ash, slags and tailings; construction and demolition wastes; biosolids; and other industrial wastes and residues.

While some monolithic deposits can be highly homogenous, fill is often highly heterogeneous. The potential for fill to be present at a site should be identified during the PSI and detailed in the CSM as a potential contaminant source.

Each fill layer must be sampled discretely and separately from underlying natural materials. This is because both fill and natural soils should be sampled as part of site characterisation, with care taken to collect discrete samples from the specific target stratum, and not sample across strata. Identified fill material should be appropriately described and logged during site investigations, and data analysis should be conducted by material type, rather than different soils and fills being analysed as one material, which can result in interpretation errors.

5.3.1. Depth of sampling

The sampling depth and interval is dependent on the CSM and mode of contamination. The NEPM (2013, B2) states that "at the surface, samples at 0–100 mm or 0–150 mm should be taken unless there is evidence of a thin surficial layer of contamination". Examples of such situations include broadacre agricultural sites, where some analytes tend to accumulate in a thin surficial layer, and areas that have received surface applications, such as termiticide sprays beneath slabs. Samples should be collected from both the emplaced fill and natural soils, at intervals of generally no more than 500 mm, and at locations where distinct differences in permeability or other observable features occur.

The following should be considered when deciding on the sampling depth interval:

- the likely fate and transport of the PCoCs
- whether permeable layers are present within fill and natural soils
- the mode of contamination
- visual/olfactory indicators of contamination (in these cases, the use of field screening tools such as photoionisation detectors (PIDs) can help identify the depth of sampling).

The sample should be collected beneath the point where fill meets the underlying natural soil.

If validating infrastructure, samples should be at a depth that is likely to intercept any potential leaks (i.e. if underground tanks have been removed, samples should be collected from the lower half of the excavation wall).

5.4. Stockpiles

It is preferable to characterise soil and fill *in situ*, but at times site- or project-specific constraints require the material to be stockpiled before it is sampled. The excavation and stockpiling of material can result in the mixing and dilution of contaminated materials with uncontaminated materials, so the excavation and placement of the material should be supervised to ensure different types of soil and fill materials are kept segregated.

Where stockpiles are pre-existing, site investigations should ensure they are fully examined and that the entire stockpile is sampled, rather than just the surface.

When characterising a stockpile, first take the CSM into account: it should include all known information about the stockpiled materials, the size of the stockpile, and the proposed end use of the material. The sampling strategy should also consider that a stockpile is three dimensional, and requires systematic sampling, using three-dimensional grids for characterisation.

As with all investigative stages in the assessment of site contamination, a multiple lines of evidence/weight of evidence approach should be used when characterising stockpiled material, taking into account appropriately described and logged field samples, along with the resulting sampling data from the analytical samples.

Table 4 Minimum number of samples recommended for initial assessment of stockpiles

Stockpile volume (m^3)	No. of samples
<75	3
75–100	4
100–125	5
125–150	6
150–175	7
175–200	8

Table modified from NEPM (2013, B2)

The sampling frequency recommended in the NEPM relates to materials of homogenous soils, suspected of contamination. Greater sampling densities are required for stockpiles that contain heterogenous material or have large ranges in contaminant concentrations.

Where there is a large range in contaminant concentration, then either the maximum concentration should be assumed for disposal purposes or additional samples collected and analysed, and the situation re-evaluated (NEPM 2013, B2). Different sampling rates may be appropriate for soil quantities greater than 200 m^3. Statistical methods to apply in this situation are discussed in Section 7.

The NEPM (2013, B2) should be consulted for further information on the sampling of stockpiled materials, and EPA (2014b) and the Resource Recovery Orders and Exemptions should be consulted for specific information relating to sampling material for waste disposal and resource recovery; these are available on the NSW EPA website.

5.5. Use of composite samples

Composite sampling of soils involves the mixing of several discrete samples or sub-samples, collected a maximum of 20 m apart, to form one composite sample for analysis. The circumstances where compositing samples can be used are very limited, and are described in DEC (2005a)

In principle, the concentration of the composite sample represents the average of the sub-samples. However, composite sampling has three major drawbacks:

- it cannot be used for the assessment of pH, volatile or semi-volatile contaminants including TRH, BTEXN, OCPs, OPPs and low molecular weight PAHs
- it should never be used for clay soils
- a simplistic analysis of composited samples can result in a sub-sample that contains a high concentration of contaminant can remain undetected due to the dilution effect of the compositing process, potentially resulting in unrepresentative data and associated decision errors.

Further information about composite samples can be found in the NEPM (2013, B2) and DEC (2005a).

5.6. Groundwater

Potential groundwater contamination must be considered when designing sampling programs at contaminated sites. These design requirements are impacted by the type and nature of the site's groundwater system, which can be complex and have multiple interacting aquifers. Aquifer-specific issues that must be considered include whether the groundwater is perched, unconfined, semi-confined or confined.

The appropriate method for the assessment of groundwater is determined first by undertaking a PSI (preliminary site investigation). This incorporates a desktop hydrogeological assessment and the development of a site-specific CSM, which must include groundwater. To inform the CSM, use existing published geological reports and hydrogeological information for the surrounding area, including that from the NSW Department of Planning, Industry and Environment. The geological information is used to determine the number and location of groundwater wells as well as screen intervals and well depths.

Groundwater wells are generally installed in locations that will maximise the likelihood of intercepting and defining the extent of groundwater contamination. This includes targeting contamination sources and known plumes, and then locating wells hydraulically up-gradient, down-gradient and cross-gradient (lateral) to the areas of concern. If a potential for offsite migration exists, groundwater monitoring wells should be installed as close to the down-gradient site boundary as practical. The location of the nearest down-gradient groundwater receptors should be considered as well.

A minimum of three wells is required for simple groundwater systems, to define the groundwater plane (water table or potentiometric surface). The water levels in the three wells are triangulated to find the groundwater flow direction (assuming a single aquifer is being assessed). Where multiple aquifers exist, more wells are required to determine the slope and direction of groundwater flow in each aquifer.

The screen in the wells need to target the appropriate aquifer/water-bearing zone or the zone of interest within that aquifer. The physical and chemical characteristics of the contaminant have the potential to affect their distribution in groundwater, e.g. light non-aqueous phase liquids (LNAPLs) such as oils are less dense than water and tend to accumulate at the top of the water table, whereas dense non-aqueous phase liquids (DNAPLs), such as some solvents, are more dense than water and tend to sink and accumulate at the bottom of the water-bearing zone. Multiple wells or wells with multiple screens/nested wells may be required to characterise the vertical groundwater profile and contaminant distribution.

Various sampling methods are available for the collection of groundwater samples. It is important that site-specific conditions and the contaminants of concern are considered when selecting an appropriate method.

The potential for natural variability across a site should also be considered if inter well comparisons are likely to made. For example, are the different pH levels between an upgradient well due to potential contamination or is natural variation the cause? Or does the aquifer vary in permeability across the site? If it is likely that natural variation may be present at the site, intra well comparisons maybe more suitable: these include concentration changes over time.

Specific guidance on groundwater sampling design and sample collection can be found in the NEPM (2013, B2) and DEC (2007). Methods for the statistical analysis of groundwater data, including intra well and inter well comparisons, can be found in ITRC (2013) and USEPA (2009).

5.7. Surface water

The sampling design for a surface-water program should take into account the CSM, the purpose/objective of the program, the chemical characteristic of the contaminants and the pathways and receptors.

The ANZG (2018) website must be consulted prior to the design of any surface-water monitoring program.

5.8. Sediment

Sampling of sediments should be undertaken with reference to ANZG (2018) and Simpson & Batley (2016).

5.9. Vapour

For vapour investigations, multiple lines of evidence should be used. The CSM needs to include:

- the design and condition of buildings, including the presence of elevators and ventilation systems
- preferential pathways – both constructed pathways (such as building sumps, drains, services and permeable backfill) and natural pathways (such as tree roots, differential soil permeabilities, fractured bedrock, etc.)
- environmental factors such as diurnal fluctuations, short-term and seasonal fluctuations in weather conditions, and variations in soil moisture and temperature
- confounding sources of contamination that may contribute to the VOCs measured at the sites, for both indoor and ambient air.

5.9.1. Soil vapour

Soil vapour sampling is the preferred approach in most situations where vapour from a sub-surface source is likely to exist. The number of samples recommended spatially for a vapour investigation depends on site-specific conditions. Access constraints (for example building construction and occupation) can significantly impact sample locations for soil vapour assessments, and different types of samples such as indoor air may be needed to obtain a weight of evidence approach.

Soil vapour sampling should be aimed at targeting the highest concentrations, either known or expected, at the site, and the location of current or future receptors (that is, inhabited buildings or the location of a proposed building). As a minimum, the NEPM (2013, B2) recommends that samples are collected above the maximum source concentration near or under a building and at each corner or along each side of the building. Where there is a known point source of vapour contamination, at least one vapour sample should be taken as close as possible to an area of highest concentration. Additional samples should be collected between the source and the potential receptors.

Refer to EPA (2020a) for further information on vapour intrusion and soil vapours (trace ground gases).

Various sampling methods are available for the collection of soil gas samples, including active and passive methods. It is important that site-specific conditions and the contaminants of concern be considered when selecting an appropriate method.

Sample frequency

At least one round of sampling should be taken in weather conditions (e.g. temperature, pressure, and soil moisture) that are likely to result in the highest vapour concentrations. Repeat sampling should be undertaken where site conditions may change – for example, where there is a fluctuating source, varying meteorological conditions, varying building use or conditions, or where remedial work is undertaken. ITRC (2007) identifies that there is no need to repeat sampling if soil gas values are a factor of 5–10 times below the risk-based screening levels, unless there is a major change in conditions (such as an elevated water table) that would significantly change vapour concentrations.

5.9.2. Indoor and ambient air

Indoor air sampling is the most direct method of sampling VOC exposure where the CSM has identified that vapour intrusion is a potential pathway. Where concentrations of contaminants of concern in indoor air attributed to a source of contamination exceeds relevant criteria, then the appropriate parties should be notified and the need for mitigation measures should be assessed.

The number of samples recommended for representative indoor air sampling depends on the size of the indoor area and the building's internal divisions (which may limit air movement). Indoor air samples

should be obtained from the crawl space and/or basement if present, and the living area at the height where occupants sit or sleep. Overall, sample locations should be targeted to inhabited buildings, with samples collected from a location similar to that of the actual breathing zone.

Sources of VOC's inside of buildings should be considered prior to sampling in assessing whether indoor air sampling is an appropriate line of evidence in the context of a weight of evidence approach.

5.9.3. Ground gases

Ground gases associated with operating or closed landfills or buried putrescible wastes are generated as a result of the biological, chemical and physical decomposition of spilled or dumped wastes. The assessment of ground gases is a complicated area of investigation and is beyond the scope of these sampling design guidelines: see the specific information in EPA (2020a) and its references.

5.10. Determining background concentrations

In the assessment of site contamination, an understanding of the background concentrations of an analyte at a site is important in understanding how much contamination may be present, particularly when assessing metals and metalloids. Metals and metalloids are naturally occurring elements: their natural 'background' concentrations in soils are highly variable, and depend on the parent rocks from which the soils originate and the processes occurring during soil formation (Gray & Murphy 1999). Arguably, as a consequence of the industrial revolution, natural background concentrations no longer exist, at least in surficial soils, due to anthropogenic sources and the global transportation of contaminants. The NEPM states that "the term ambient background concentration (ABC) … is used rather than background concentration" (2013, B5b).

The NEPM (2013, B1) assumes that ecosystems are adapted to the ABCs of metals in soils, and that it is only the addition of contaminants above this background concentration that has an adverse effect on the environment. It notes that:

> *The ABC of a contaminant is the soil concentration in a specified locality that is the sum of the naturally occurring background level and the contaminant levels that have been introduced from diffuse or non-point sources by general anthropogenic activity not attributed to industrial, commercial, or agricultural activities, for example, motor vehicle emissions.*

This definition can be extended to other media.

Determining the ambient concentrations for any medium relies on identifying sites or decision areas that have not been affected by the same or similar contaminating activities as the subject site or decision area (or, if that is not possible, not affected to the same magnitude).

The following should be considered when determining the ABCs of various media.

5.10.1. Soils

- Take care to ensure that the background areas consist of the same soil types as the site or decision area.
- Collect and compare samples with soils/sediments from the same soil horizon layer.

5.10.2. Groundwater

- Construct background groundwater monitoring wells in the same manner as the subject site wells, targeting the same aquifer.
- Consider potential sources of contamination up-gradient of the well.
- Assess preceding rainfall and standing water levels.
- Collect and record physico-chemical parameters at both the decision site and an unaffected site.
- In highly fractured or karstic geological environments, seek specialist hydrogeological support if required.

5.10.3. Surface water and sediments

- For fresh water, ensure sample locations are upstream of the source of contamination.
- Consider the impacts of tidal flow, stratification, other sources of contaminants and the potential re-suspension of contaminants that have sorbed to sediments.
- Assess the weather before and at the time of sampling, and any potential impacts of this.
- Collect and record physico-chemical parameters at all sampling locations.
- Sampling of surface waters and sediments should be undertaken with reference to ANZG (2018) and to Simpson and Batley (eds) (2016).

6. Hotspot detection

Hotspots are defined as localised areas characterised by significantly higher contaminant concentrations relative to other areas of a site or decision area.

Systematic sampling aimed at detecting hotspots of specified shapes and sizes is required to characterise or validate sites or decision areas. The sampling grids are placed at regularly spaced intervals, as discussed in Section 5.2. The grid size and pattern required for site characterisation depends on what is already known about the site and described in the CSM, as well as the shape and size of the target hotspots.

In regard to determining the grid size, the NEPM (2013, B2) says that:

> *Determining grid size/sampling density from mathematical formulae (for example, Appendix D of Standard AS 4482.1–2005) is not an acceptable approach without consideration of likely contaminant distribution and acceptable hotspot size.*

The number of sampling locations required for site characterisation is based on the following principles:

- the number of samples derived from the systematic sampling is adequate to indicate the true value of other critical parameters of a contaminant distribution, such as the average concentration and variability
- the spacing between sampling locations should be determined according to the conceptual model, the phase of the investigation, acceptable levels of uncertainty and the requirements of the risk assessment. (BSI, 2013).

The land to be sampled may be intended for subdivision. If so, the minimum hotspot size for investigation should be no larger than the size of the proposed or likely land parcels. While lot sizes depend on location and development type, an average lot size of between 400 m^2 and 500 m^2 is a reasonable assumption.

This concept is, in part, derived from the NEPM (2013, B2), which says:

> *If a site is to be subdivided, the size of the subdivided lots should be taken into account when determining the sampling density. While predictions may be made on a 'macro' scale, residents or owners may seek information about their own particular area of land and the risks associated with this land, especially if the potential contamination on the original site was uneven in distribution and type.*

Hotspots rarely have sharply defined boundaries. Contamination, strata and fill types, are often present as heterogeneous pockets across sites due to site features and past uses. Accordingly, the use of systematic sampling grids generally aims to allow appropriate location and mapping of the materials at the site, to provide representative data, and to determine the 'true value' of other critical parameters of a contaminant distribution, more so than necessarily finding distinctly definable hotspots.

Methods for determining the required grid size for circular and elliptical hotspots are shown in Appendix C.

The recommended sampling points required for site characterisation provided in Appendix C should not be considered as fixed: for irregularly shaped sites, more sampling points may be needed to detect hotspots of the calculated minimum size. The grid size is the key metric: as the number of sampling points required is in part based on the geometry of the site or decision area, the actual number of sampling points required is dependent on applying the specified grid size to the actual site or decision area. This is illustrated in Figure 5.

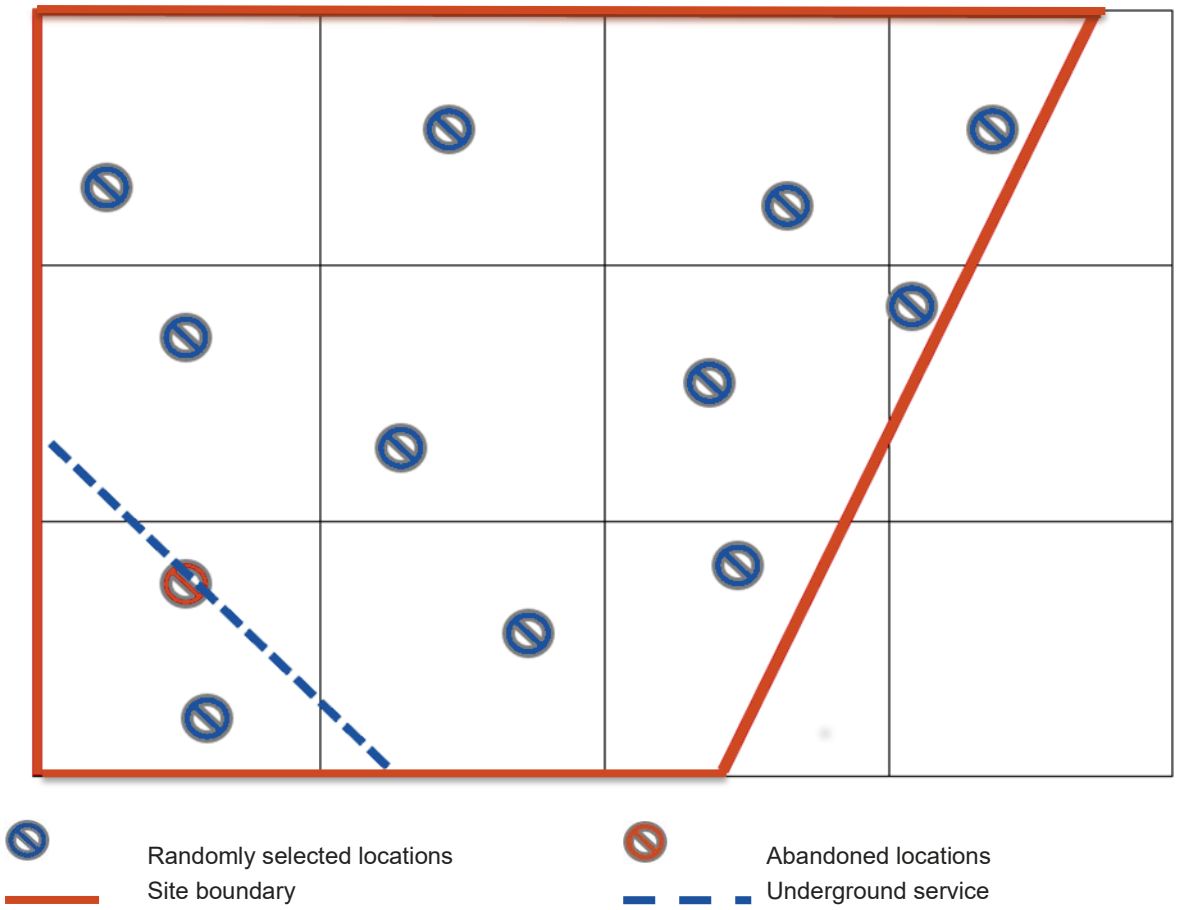

Figure 5 Placement of sampling grid and randomly selected sampling locations

Source: Marc Salmon/Easterly Point Environmental Pty Ltd

When deciding whether to use a square grid pattern, consider the site characteristics and specific investigation objectives. Ferguson (1992) suggests that systematic random designs are less efficient than aligned square grid designs for detecting hotspots of a specified size (although this finding depends on target shape and orientation, and is somewhat contradicted by later findings (BSI 2013)).

With systematic random designs, the randomness of the sampling points can be maintained by simply generating a new random coordinate set, if the design location is obscured. This is shown in Figure 5.

7. Number of samples required

This section provides methods for calculating the number of samples required to be representative of a population, by considering factors such as variance in the dataset and confidence levels. Consultants wishing to determine the number of samples required for site characterisation or site validation as a function of the site area should refer to Section 5.2.

The aim of environmental sampling is to collect sampling data that is representative of the population being sampled. **Representativeness** is a measure of the degree to which the sampling data accurately and precisely represents the characteristics of the population being studied. The probability of achieving representative data is, in part, controlled by the number of samples.

The number of samples required is defined by a number of interacting factors, including:

- the purpose of the sampling
- the sampling strategy selected
- the inherent variability of the target population
- the minimum effect size that needs to be determined, and
- the certainty required, including both the specified confidence level and the statistical power.

Consultants should consider all these interrelated factors when determining the number of samples required to achieve an investigation's objectives.

The effect of some of these factors is illustrated in Figure 6, which shows the sample size needed for a one-sample t-test at a 95% confidence level and at various statistical powers (α = 0.05, β = 0.05, 0.1 and 0.2) (from USEPA 2002, G–5S). The number of samples required increases significantly as the effect size (defined below) becomes a smaller fraction of the estimated value, and also as the required confidence level and power increase.

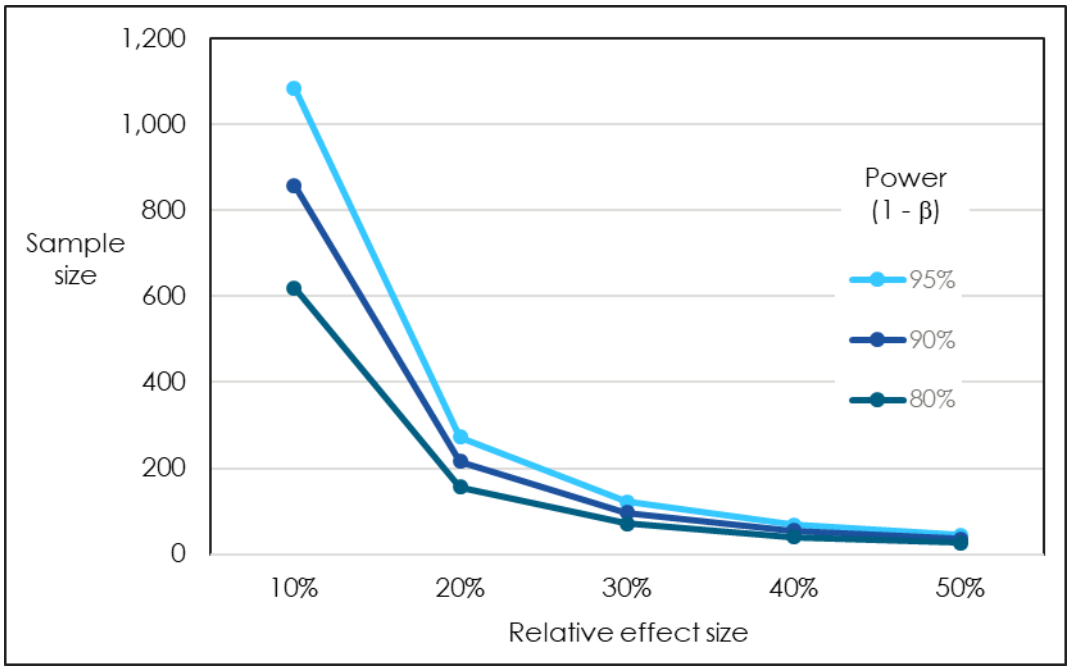

Figure 6 Sample size at 95% confidence level, based on effect size as fraction of estimated value and power required

Source: Marc Salmon/Easterly Point Environmental Pty Ltd

Two statistical methods are provided for determining the number of samples (n) required: the CRV (combined risk value) method shown in Appendix E and the MPE (maximum probable error) method shown in Appendix F.

The CRV method is in part determined by the **effect size**, which is defined as the magnitude of the difference between the populations or groups being studied. In the assessment of site contamination, the effect size typically measures the difference between, for example the 95% UCL$\bar{x}$, and the criterion or action level. Many of the procedures used to determine n will provide a small n for large effect sizes, or a

large n for small effect sizes. This includes the CRV method, which may provide an unrealistically small (<0.1) or large (>1,000) value for n based on the effect size. Accordingly, this method may be unsuited for determining if a site or decision area has been adequately characterised, or meets a specified criterion, unless other methods are used to confirm n. Both the MPE method and hotspot detection approach should also be used, where appropriate, as part of the multiple lines of evidence/weight of evidence approach.

Within the DQOs (data quality objectives) process, the phenomenon of the sample size increasing as the effect size becomes smaller is addressed through the use of the **grey region**. This is where the results are 'too close to call' (USEPA 2006, G-4), and the consequences of making a decision error are considered to be relatively minor. USEPA (2000, G-4HW) describes the grey region as "the range of possible parameter values near the action level where the cost of determining that the alternative condition is true outweighs the expected consequences of a decision error".

In hypothesis testing approaches, the width of the grey region is called the **minimum detectable difference** Δ (the uppercase Greek letter delta). It is determined by the parameter values for which the α and β probabilities are set and is the region for which decision errors are considered tolerable. In general, the narrower the grey region, the greater the number of samples needed to determine whether to accept or reject the null hypothesis, H_0.

As H_0 is that the site is contaminated in the assessment of site contamination, the grey region represents the probability of a Type II or false acceptance decision error, and values within this region have a higher probability of being falsely accepted. When a UCL$\bar{x}$ is used for hypothesis testing, the probability of making a Type I or false rejection decision error is controlled; however, this approach does not control against making a Type II or false acceptance decision error.

If the decision rests on showing that the UCL$\bar{x}$ is less than the criterion, the number of samples required will depend on how close the arithmetic mean is to the criterion. The narrower the gap between the mean and the criterion, the greater the number of samples that will be required to statistically demonstrate that the UCL$\bar{x}$ is less than the criterion.

7.1. Existing guidance

Appendix D summarises EPA-made and -approved guidance on sample design, and other relevant guidelines.

When using the CRV and MPE methods to assist in determining the number of samples required to achieve the project objectives, take into account the sample design guidance for the specified land use, media and/or contaminant types and incorporate it into the sampling design where relevant.

Neither the methods described below nor those referenced in Appendix D are to be considered minimum requirements. Rather, the method to be used needs to be chosen according to the situation-specific requirements of the investigation, and to be fully explained and documented, including any assumptions and limitations.

7.2. Combined risk value method

The number of samples needed to show that the average concentration of a contaminant is below a defined criterion or action level can be determined using the CRV method. The CRV method can be used for a variety of media samples.

The determination is based on the principle of hypothesis testing, with the alpha (α) value for a Type I error, or false rejection of the hypothesis, and the beta (β) value for a Type II error, or false acceptance of the hypothesis, being used to determine the CRV. As the methodology is based on parametric methods, it assumes nearly-normal distribution and independent and unbiased sampling data.

The CRV method is used in hypotheses testing of the arithmetic mean to determine if n is sufficient. It is based on the specified values of α and β, and the effect size resulting from the difference between $\bar{x}$ and the specified criterion or action level. If the null hypothesis is not rejected, then the only potential decision error is false acceptance (β), and the CRV method can be used to determine if the error rate has been satisfied (USEPA 2006, G-4). If n as determined by the CRV method is less than the number of

analytical samples, then a Type II error may have been made. In such a case, the only way to maintain the selected probability is to increase the number of analytical samples.

Appendix E shows how to use the CRV method to determine the number of samples required and gives a worked example.

7.3. Maximum probable error

When the objective of the sampling includes the estimation of the population arithmetic mean at a specified confidence level, the MPE method as described in Provost (1984) and Gilbert (1987) can be used. This method uses the margin of error (MoE), the standard deviation (s), and the t critical value, at a 95% confidence level or higher.

As MPE is based on parametric methods, it assumes nearly-normal distribution and independent and unbiased sampling data. The MPE equation ultimately approaches n = n, that is, all other parameters cancel out; MPE therefore cannot be used to retrospectively demonstrate sufficient sampling. Rather, it provides a guide to an appropriate number of samples based on the variability of the data (standard deviation, s), and the required precision of the data (margin of error, MoE). Once the standard deviation of the sample dataset is known, the desired MPE can be selected, and the number of samples required to achieve that MPE can be determined.

The MPE method can be used for all media, for areas and stockpiles. However, it is insensitive to the area or volume of interest, and should be used in conjunction with other methods to confirm that a sufficient number of analytical samples has been collected and analysed.

Appendix F shows how to use the MPE method to determine of the number of samples required and gives a worked example.

8. Abbreviations and glossary

8.1. Acronyms

ABC	Ambient background concentration
ANZG	Australian and New Zealand water quality guidelines
CECs	Contaminants of emerging concern
CLT	Central limit theorem
CLM	Contaminated land management
CRV	Combined risk value
CSM	Conceptual site model
CV	Coefficient of variation
DNAPLs	Dense non-aqueous phase liquids
DQIs	Data quality indicators
DQOs	Data quality objectives
DSI	Detailed site investigation
DUs	Decision units
EPA	Environment Protection Authority
HIL	Health-based investigation level
HSL	Health screening level
ISM	Incremental sampling methods
LNAPLs	Light non-aqueous phase liquids
LOR	Limits of reporting
Metals	Arsenic (As), cadmium (Cd), chromium (Cr), copper (Cu), lead (Pb), mercury (Hg), nickel (Ni) and zinc (Zn)
MoE	Margin of error
MPE	Maximum probable error
MQOs	Measurement quality objectives
NEPM	National Environmental Protection Measure
NHST	Null-hypothesis significance testing
NOW	New South Wales Office of Water
OEH	New South Wales Office of Environment and Heritage
PAHs	Polycyclic aromatic hydrocarbons
PFAS	Per- and poly-fluorinated alkyl substances
PFOS	Perfluorooctane sulfonate
PFOA	Perfluorooctanoic acid
PFHxS	Perfluorohexane sulfonate
PSH	Phase-separated hydrocarbon
PSI	Preliminary site investigation
PCoCs	Potential contaminants of concern

PID	Photoionisation detector
PTFE	Polytetrafluoroethylene
QAPP	Quality assurance project plan
QA/QC	Quality assurance/quality control
Q-Q	Quantile–quantile
RAP	Remediation action plan
RSD	Relative standard deviation
SAQP	Sampling and analysis quality plan
SOPs	Standard operating procedures
STP	Sewage treatment plant
SWL	Standing water level
TOFA	Total organic fluorine assay
TOPA	Total oxidisable precursor assay
TRHs	Total recoverable hydrocarbons, including volatile C6–C10 fractions and semi- and non-volatile C11–C40 fractions
UCLs	Upper confidence limits
$UCL_{\bar{x}}$	Upper confidence limits of means
UPSS	Underground petroleum storage system
USEPA	United States Environmental Protection Agency
UST	Underground storage tank
VOCs	Volatile organic compounds

8.2. Statistical notations

$1 - \alpha$	Confidence level
α	Type I error rate (see Glossary)
β	Type II error rate (see Glossary)
c	Criterion/action level
df	Degrees of freedom
exp	Exponential function
H_A	Alternative hypothesis
H_0	Null hypothesis
n	Number of samples or measurements in a sample (see sample definition)
θ	Scale parameter of the gamma distribution
σ	The population standard deviation, which is generally not known
σ^2	The population variance, which is generally not known
p-value	Probability value
Δ	Uppercase Greek letter delta, denoting the width of the grey region associated with hypothesis testing
s	The sample standard deviation, which is determined from the measurements taken
s^2	The sample variance, which is determined from the measurements taken

δ_0	Difference (delta) of zero
$t\alpha$	Critical value
t_0	Test statistic
μ	The population mean, which is generally not known
$UCL\bar{x}$	Upper confidence limit of arithmetic mean
$\bar{x}$	The sample mean, which is determined from the measurements taken
x_i	The i^{th} measurement in the dataset

8.3. Glossary

α risk

The probability, expressed as a decimal, of making a 'type I error' when the hypothesis is tested statistically. A type I error wrongly rejects a null hypothesis when in fact the null hypothesis is true. In this document, the null hypothesis always assumes that the site is 'contaminated' and thus the α risk refers to the probability of a site being validated 'uncontaminated' when in fact it is 'contaminated'.

β risk

The probability, expressed as a decimal, of making a 'type II error' when a hypothesis is tested statistically. A type II error wrongly accepts a null hypothesis when in fact the null hypothesis is false. In this document, the null hypothesis always assumes that the site is 'contaminated' and thus the β risk refers to the probability that a site is concluded 'contaminated' when in fact the site is 'uncontaminated'.

Acceptable limit

A threshold concentration value below which the level of contamination is regarded as acceptable. An acceptable limit can either be adopted from the appropriate guidelines, or it can be derived on a site-specific basis using risk assessment. Where site remediation is involved, acceptable limits are often referred to as 'clean-up standards' or 'remediation standards'.

Acceptance criteria

A statistical statement specifying how a contaminant distribution will be compared with an acceptable limit (see above definition) to determine whether a site should be evaluated as 'contaminated' or 'uncontaminated'. The concentrations of a contaminant can vary over orders of magnitude in a sampling area. All site assessments must state the appropriate acceptance criteria, as well as the appropriate acceptable limits.

Ambient air

External air environment, not including the air environment inside buildings or structures.

Arithmetic mean

The arithmetic mean is commonly referred to as the average and is used to describe the centre of the data distribution. It is obtained by summing all the values and dividing the result by the number of values.

Central tendency

The central or typical value for a probability distribution and may be considered the average value in a set of data. It is generally described by the mode, median, or, more commonly, the mean, and describes where a sample distribution is centred.

Chi-squared distribution

A type of cumulative probability distribution that varies depending on the degrees of freedom (df). It is used to test relationships between categorical variables in the same population.

Coefficient of variation (CV)

CV is the measurement of the relative homogeneity of a distribution. Low CV values, e.g. 0.5 or less, indicate fairly homogeneous contaminant distribution, while CVs with values over I–1.2 imply that the concentration distribution of a contaminant is heterogeneous and probably highly skewed to the right.

Composite sample

The bulking and thorough mixing of soil samples collected from more than one sampling location to form a single soil sample for chemical analyses.

Conceptual site model (CSM)

Provides a three-dimensional overview of the contamination at sites and their surrounds, highlighting the sources, receptors and exposure pathways between the sources and receptors.

Confidence level

The probability, expressed as a percentage, that a statistical statement is correct. Confidence level is the opposite expression of 'risk' (see definitions of α and β risks). For the purpose of this document in which a risk that needs to be regulated, the confidence level is always equal to I - α.

Contaminated

For the purpose of this document and depending on the context, 'contaminated' can have slightly different meanings. If a site or a sampling area is evaluated as 'contaminated', it means that the site or the sampling area as a whole has not met the acceptance criteria (see definition of acceptance criteria). 'Contaminated' can also be used to describe a localised area or soil that has contaminant concentrations exceeding an acceptable limit (see definition of acceptable limit). Note: depending on what the acceptance criteria are, an entire site could be considered 'uncontaminated' even though a certain percentage of the site is expected to be 'contaminated'.

Data quality objectives (DQOs)

A systematic planning process used to define the type, quantity and quality of data needed to support decisions relating to the environmental condition of a site or a specific decision area.

Decision area

A specific area or medium within a site, or offsite, about which data is being gathered so a decision can be made. For example, a decision can include part of a site, soil, a stockpile, soil gas, groundwater, surface waters or sediments.

Estimate

An estimate is a value that is inferred for a population based on data collected from a sample of units from that population. For example, the measured data from a sampling event used to calculate the sample mean ($\bar{x}$) is then used to estimate the population mean (μ).

Estimation

A technique that systematically adjusts the sample data to determine an estimated value for the population.

Geometric mean

This is similar to the **arithmetic mean** (described above), in that it is also a measure of the central tendency of the distribution of a population or sample. It is sensible to calculate geometric means only on populations or samples that contain positive values Y.

Grab samples

Samples collected from different locations that will not be composited but analysed individually.

Hotspot

A localised area where the level of contamination within that area is noticeably greater than that in surrounding areas. Note that a hotspot is only **relatively** high in contamination.

Inter well

Comparison between two groundwater monitoring wells that are separated spatially.

Intra well

Comparison of measurements over time at one groundwater monitoring well.

Maximum

The maximum observed value in a data. Important, as it generally provides a conservative estimate of the potential exposure risks. It is generally assumed that if the maximum is below the action level, then the site should be suitable for the associated land use.

Median

The middle value of the distribution. Half the data values are less than the median and half are greater.

Minimum size effect

The acceptable magnitude of the difference between the populations or groups being studied.

Mode

The value that occurs most frequently. It is determined by counting the number of times each value occurs.

Modules

A series of discrete DQOs outputs, based on logical categories, that address selected components of a site investigation. Modules can be selected for contaminant types, media, decision areas, or a workable combination of these.

Neyman–Pearson method

A method of statistical inference used to determine if a null hypothesis (H_0) should be rejected in favour of an alternative hypothesis (H_A), at a specified level of confidence.

Outlier

A data point that sits outside the expected range of the data. An outlier can have either a high or low value. Unless there is a demonstratable reason for rejecting them (such as coding error, sample contamination or equipment failure), outliers need to be retained within sample datasets.

Parameters

Numerical measures of the characteristic of interest in the population being sampled. Typical parameters are the population mean (μ), variance (σ^2) and standard deviation (σ). Parameter values are usually unknown.

Percentiles and quartiles

As their names suggest, these are descriptive values used to equally split a dataset into 100 parts. A percentile is the value that a given percentage of observations in a dataset is equal to or less than, e.g. 80% of observations in a dataset are at or below the 80th percentile, while 20% are above.

Quartiles are commonly used to break the dataset up into four equal parts, providing an indication of the distribution and variance of the data.

First quartile – the 0^{th} percentile up to (and including) the 25^{th} percentile

Second quartile – from the 25^{th} percentile up to (and including) the 50^{th} percentile

Third quartile – from the 50^{th} percentile up to (and including) the 75^{th} percentile

Fourth quartile – from the 75^{th} percentile up to (and including) the 100^{th} percentile

Population

Any large collection of objects, things or individuals with some characteristics in common, that is being studied and for which information is sought. The population under consideration must be clearly and succinctly defined to allow effective sampling design and subsequent reporting.

The population can be further defined as the **target population** and the **sampled population**, and ideally these should be the same. The target population is the set of all units that comprise the items of interest, that is the population about which a decision is required, and the sampled population is that part of the target population that is accessible and available for sampling. If the two diverge significantly, the target population should be redefined.

Probabilistic sampling

Probabilistic sampling occurs when each member of the population has a given probability (greater than zero and less than one) of being included in the sample. If the probability is the same for all population members then, and only then, will the sample be unbiased. Because inclusion in the sample is based on probability, subsequent samples won't necessarily include the same members.

Range

The range of a dataset measures the spread between the highest and lowest values in the dataset. Other measures (such as the standard deviation and the interquartile range) are required to provide an understanding of the distribution of the data.

Residual soil

The soil at a site that is not contaminated by industrial, commercial, or agricultural activities, consistent with the term 'ambient background concentration' (ABC) from the NEPM. Residual soils can include natural soils, reworked natural soils and historically imported material. Residual soils may have naturally occurring background levels of contaminants, contaminants that have been introduced from diffuse or non-point sources by general anthropogenic activity, and only low levels of contaminants attributed to industrial, commercial, or agricultural activities.

Sample

'Sample' has a number of meanings in the assessment of site contamination, including:

- as more broadly used in statistics, a representative group drawn from a population for description or measurement

- a physical amount of a material (soil, water, air, etc.) or an aliquot, taken for testing or chemical analysis
- a sampling point or sample location, being the location in plan at which a sample is collected, including description (e.g. geological logs) and field screening (e.g. PID, XRF, etc.).

Sample size

The number of samples or sampling points selected in a sampling program.

Sampling, analysis and quality plan (SAQP)

Incorporates the CSM and the DQO outputs, to provide the context and justification of the selected sampling and analysis. The methods, procedures and quality control (QC) samples associated with the DQIs, including the frequency and MQOs, along with any associated contingencies, are also documented. The SAQP ensures that the data collected is representative and provides a robust basis for site assessment (NEPM 2013).

Sampling pattern

The locational pattern of sampling points within a sampling area.

Sampling point

The location at which a soil sample is collected.

Site characterisation

The assessment of the nature, level and extent of contamination. A typical site characterisation involves a preliminary site investigation (PSI), followed by a detailed site investigation (DSI), where warranted.

Site validation

The process of showing that a site is successfully remediated.

Standard deviation

Calculated by taking the square root of the variance (described below). It provides an indication of a population or sample data's typical deviation from its mean.

Statistic

Any summary number that describes the sample, such as an average or percentage. For example, the mean of a sample is described as $\bar{\mathbf{x}}$ (x-bar) and the standard deviation as **s**. When describing the population from which the sample is drawn, a summary number is called a **parameter**.

Statistical power

The probability of correctly determining a positive result (e.g. a change or difference in the population) based on sample data.

Sub-sample

A sample that will be bulked together with other sub-samples to form a composite for chemical analyses.

Systematic planning

A planning process based on a scientific method, and which leads the project to unfold logically. Systematic planning includes established management and scientific elements. In the assessment of site

contamination, it includes the application of the **DQOs** process and development of both a **CSM** and an **SAQP**.

Variable

A characteristic, number or quantity that is the subject of the inquiry. In the assessment of site contamination, it is usually continuous numerical variables that are being assessed, for example the concentration of a contaminant in soil, soil gas or water. Discrete or discontinuous variables are at times considered, such as the number of fish in a waterbody. These are both quantitative variables in that they are derived by measurements.

Qualitative or categorical variables include ordinal or ranked variables and nominal variables. Ordinal variables are observations that take a value that can logically be ordered or ranked, such as first, second, third, whereas nominal observations take a value that cannot be organised in a logical sequence, such as presence or absence. Categorical variables are not commonly used in the assessment of site contamination and are not considered further.

Variance

The average squared distance of population or sample data points from the associated mean.

Weight of evidence/lines of evidence

‘Weight of evidence’ describes the process of collecting, analysing and evaluating a combination of different qualitative, semi-quantitative or quantitative lines of evidence to make an overall assessment of contamination.

Applying a weight-of-evidence process incorporates judgements about the quality, quantity, relevance and congruence of the data contained in the different lines of evidence (ANZG 2018).

9. References

Australian and New Zealand Environment and Conservation Council (ANZECC) and Agriculture and Resource Management Council of Australia and New Zealand (ARMCANZ) 2000, *Australian and New Zealand Guidelines for Fresh and Marine Water Quality*, paper no. 4, ANZECC and ARMCANZ, Canberra.

British Standards Institution (BSI) 2013, *Investigation of Potentially Contaminated Sites: Code of Practice*, BS 10175:2011+A1:2013, BSI Standards Limited.

Clements L, Palaia T & Davis J 2009, *Characterisation of Sites Impacted by Petroleum Hydrocarbons: National Guideline Document*, CRC Care Technical Report no. 11, CRC for Contamination Assessment and Remediation of the Environment, Adelaide.

CRC Care 2013*, Petroleum Hydrocarbon Vapour Intrusion Assessment: Australian Guidance*, CRC CARE Technical Report no. 23, CRC for Contamination Assessment and Remediation of the Environment, Adelaide.

Crumbling DM 2002, In search of representativeness: evolving the environmental data quality model, *Quality Assurance*, vol.9, pp.179–90, http://cluin.org/download/char/dataquality/dcrumbling.pdf.

Davis GB, Wright J & Patterson BM 2009, *Field Assessment of Vapours,* CRC CARE Technical Report no. 13, CRC for Contamination Assessment and Remediation of the Environment, Adelaide.

Department of Agriculture and Water Resources 2018*, Australian & New Zealand Guidelines for Fresh and Marine Water Quality*, Department of Agriculture and Water Resources, Canberra, www.waterquality.gov.au/anz-guidelines.

Department of Environment and Conservation (DEC) 2004, New South Wales (NSW) *Australian river assessment system (AUSRIVAS) sampling and processing manual*.

Department of Environment and Conservation (DEC) 2005a, *Contaminated sites: Guidelines for assessing former orchards and market gardens*, DEC 2005/195, DECCW NSW, Sydney.

Department of Environment and Conservation (DEC) 2005b, *Information for the assessment of former gasworks sites*, DEC 2005/237, DECCW NSW, Sydney.

Department of Environment and Conservation (DEC) 2007, *Contaminated sites: Guidelines for the assessment and management of groundwater contamination*, DEC 2007/144, DEC NSW, Sydney.

Department of Environment, Climate Change and Water (DECCW) 2009, *Guidelines for implementing the Protection of the Environment Operations (Underground Petroleum Storage Systems) Regulation 2008*, DECCW 2009/653, DECCW NSW, Sydney.

Department of Environment (DoE) Queensland 1998, *Draft guideline for the assessment & management of contaminated land in Queensland*, DoE, Brisbane.

Department of Environment and Science (DES) Queensland, 2018, Monitoring and sampling manual: Environmental protection (water) policy 2009 [sic], DES, Brisbane.

Department of Health and Ageing and EnHealth Council 2012, *Environmental Health Risk Assessment: Guidelines for Assessing Human Health Risks from Environmental Hazards*, Department of Health and Ageing, Canberra.

Devore J & Farnum N 2005, *Applied Statistics for Engineers and Scientists*, 2nd Edition, Brooks/Cole, Cengage Learning, Belmont CA.

Environment Protection Authority (EPA) 1995, *Contaminated sites: Guidelines for the vertical mixing of soil on former broad-acre agricultural land*, EPA 2003/28, NSW EPA, Sydney.

Environment Protection Authority (EPA) 1997, *Contaminated sites: guidelines for assessing banana plantation sites*, EPA 97/37, NSW EPA, Sydney.

Environment Protection Authority (EPA) 2014b, *Resource Recovery Order under Part 9, Clause 93 of the Protection of the Environment Operations (Waste) Regulation 2014: The excavated natural material order 2014.*

Environment Protection Authority (EPA) 2014c, *Best practice note: Landfarming*, EPA 2014/0323, NSW EPA, Sydney.

Environment Protection Authority (EPA) 2015a, *Guidelines on the duty to report contamination under the Contaminated Land Management Act 1997*, EPA 2015/0164, NSW EPA, Sydney.

Environment Protection Authority (EPA) 2015b, *Technical note: Light non-aqueous phase liquid assessment and remediation*, EPA 2015/0553, NSW EPA, Sydney.

Environment Protection Authority (EPA) 2017, *Contaminated land management: Guidelines for the NSW site auditor scheme (3rd edition)*, EPA 2017/P0269, NSW EPA, Sydney.

Environment Protection Authority (EPA) 2018, *Guidelines on resource recovery Orders and Exemptions: For the land application of waste materials as fill*, EPA 2017/P0392, NSW EPA, Sydney.

Environment Protection Authority (EPA) South Australia 2005, *Composite soil sampling in site contamination assessment and management*, Government of South Australia, Adelaide.

Environment Protection Authority (EPA) Victoria 2009, *Industrial waste resource guidelines: Soil sampling*, IWRG702.

Ferguson CC 1992, The statistical basis for spatial sampling of contaminated land*, Ground Engineering*, vol. 25, no. 6, pp. 34–38.

Gilbert RO 1987, *Statistical methods for environmental pollution monitoring*, John Wiley & Sons Inc., Brisbane.

Gray JM & Murphy BW 1999, *Parent material and soils: A guide to the influence of parent material on soil distribution in eastern Australia*, Technical Report No. 45, NSW Department of Land and Water Conservation, Sydney.

Hamon RE, Mclaughlin MJ, Gilkes RJ, Rate AW, Zarcinas B, Robertson A, Cozens G, Radford N & Bettenay L 2004, Geochemical indices allow estimation of heavy metal background concentrations in soils, *Global Biogeochemical Cycles*, vol. 18, GB1014.

Harr ME 1987, *Reliability-Based Design in Engineering*, McGraw-Hill, New York.

HEPA 2018, *PFAS National Environmental Management Plan*, Heads of EPAs Australia and New Zealand. (Note: this document is updated frequently. For the latest guidance, see http://www.environment.gov.au/protection/chemicals-management/pfas.)

Interstate Technology and Regulatory Council (ITRC) 2007, *Vapor intrusion pathway: A practical guideline*, VI-1, ITRC Vapor Intrusion Team, Washington DC, USA, www.itrcweb.org/Documents/VI-1.pdf.

Interstate Technology and Regulatory Council (ITRC) 2012, *Incremental sampling methodology*, (ISM-1), ITRC, Washington DC, USA, www.itrcweb.org/Guidance/ListDocuments?topicID=11&subTopicID=16.

Interstate Technology and Regulatory Council (ITRC) 2013, *Groundwater statistics and monitoring compliance: Statistical tools for the project life cycle*, GSMC-1, ITRC, Washington DC, USA, http://www.itrcweb.org/gsmc-1/.

Interstate Technology and Regulatory Council (ITRC) 2017, *Naming conventions and physical and chemical properties per- and polyfluoroalkyl substances (PFAS)*, ITRC, Washington DC.

Lock WH 1996, Composite sampling, in *National Environmental Health Forum Monographs, Soil Series No. 3*, South Australian Health Commission, Adelaide.

McDougall KW & Macoun TW 1996, *Guidelines for the Assessment and Clean Up of Cattle Tick Dip Sites for Residential Purposes*, NSW Agricultural in conjunction with CMPS&F Environmental, Wollongbar NSW.

Naidu R, Jit J, Kennedy B & Arias V 2016, Emerging contaminant uncertainties and policy: The chicken or the egg conundrum, *Chemosphere*, vol. 154, pp. 385–390.

National Environment Protection Council (NEPC) 2013, *National Environment Protection (Assessment of Site Contamination) Amendment Measure 2013 (No. 1)*, Schedule A and Schedules B(1)–B(9), National Environment Protection Council, Canberra.

National Environment Protection Council (NEPC) 2013, *National Environment Protection (Assessment of Site Contamination) Amendment Measure 2013 (No. 1)*, Schedule A and Schedules B(1)–B(9), National Environment Protection Council, Canberra.

Nickerson RS 2000, Null hypothesis significance testing: A review of an old and continuing controversy, *Psychological Methods*, vol. 5, no. 2, pp. 241–301.

New Jersey Department of Environmental Protection (NJDEP) 2005, *Vapor intrusion guidance*.

Northern Territory Environment Protection Authority (EPA) 2013, *Guidelines on conceptual site models*, NT EPA, Darwin.

Perezgonzalez JD 2015, Fisher, Neyman–Pearson or NHST? A tutorial for teaching data testing, *Frontiers in Psychology*, vol. 6, article 223, p. 1.

Provost LP 1984, Statistical Methods in Environmental Sampling, in Schweitzer GE and Santolucito JA (eds), *Environmental Sampling for Hazardous Wastes*, American Chemical Society, Washington DC.

Reinhart A 2015, *Statistics Done Wrong: The Woefully Complete Guide*, No Starch Press, San Francisco CA.

South Australian Health Commission (SAHC) 1995, *Guidelines for the composite sampling of soils*, SAHC, Adelaide.

Simpson S and Batley G (eds) 2016, *Sediment Quality Assessment: A Practical Guide*, CSIRO Publishing, Melbourne.

US Environmental Protection Agency (USEPA) 1996, *Soil Screening Guidance: User's Guide* (2nd edition), Attachment B, Soil Screening DQOs for Surface Soils and Subsurface Soils, EPA/540/R-96/018, USEPA, Washington DC.

US Environmental Protection Agency (USEPA) 2000, *Data Quality Objectives Process for Hazardous Waste Site Investigations (QA/G-4HW)*, EPA/600/R-00/007, USEPA, Washington DC.

US Environmental Protection Agency (USEPA) 2001, *Guidance on Data Quality Indicators (QA/G-5i)*, USEPA, Washington DC.

US Environmental Protection Agency (USEPA) 2002, *Guidance on Environmental Data Verification and Data Validation (QA/G-8)*, EPA/240/R-02/004, USEPA, Washington DC.

US Environmental Protection Agency (USEPA) 2002, *Guidance on Choosing a Sampling Design for Environmental Data Collection for Use in Developing a Quality Assurance Project Plan (QA/G-5S)*, EPA/240/R-02/005, USEPA, Washington DC.

US Environmental Protection Agency (USEPA) 2002, *Guidance for Quality Assurance Project Plans (QA/G-5)*, EPA/240/R-02/009, USEPA, Washington DC.

US Environmental Protection Agency (USEPA) 2006, *Data Quality Assessment: Statistical Methods for Practitioners (QA/G-9S)*, EPA/240/B-06/003, USEPA, Washington DC.

US Environmental Protection Agency (USEPA) 2006, *Guidance on Systematic Planning Using the Data Quality Objectives Process (QA/G-4)*, EPA/240/B-06/001, Appendix: Derivation of sample size formula for testing mean of normal distribution versus an action level, USEPA, Washington DC.

US Environmental Protection Agency (USEPA) 2006, *Data Quality Assessment: A Reviewer's Guide (QA/G-9R)*, EPA/240/B-06/002, USEPA, Washington DC.

US Environmental Protection Agency (USEPA) 2007, *Guidance for Preparing Standard Operating Procedures (SOPs) (QA/G-6),* EPA/600/B-07/001, USEPA, Washington DC.

US Environmental Protection Agency (USEPA) 2009, *Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities: Unified Guidance*, EPA 530/R-09-007, USEPA, Washington DC.

US Environmental Protection Agency (USEPA) 2014, *Test Methods for Evaluating Solid Waste: Physical/Chemical Methods Compendium (SW-846)*, USEPA, Washington DC.

US Environmental Protection Agency (USEPA) 2015a, *ProUCL Version 5.1.002: Technical Guide: Statistical Software for Environmental Applications for Data Sets with and without Nondetect Observations*, EPA/600/R-07/041, USEPA, Washington DC.

US Environmental Protection Agency (USEPA) 2015b, *ProUCL Version 5.1.002: User Guide: Statistical Software for Environmental Applications for Data Sets with and without Nondetect Observations, EPA/600/R-07/041*, USEPA, Washington DC.

Wendelberger J & Campbell K 1994, *Non-Detect Data in Environmental Investigations*, American Statistical Association, Toronto, Canada.

Wilson S, Card G & Haines S 2009, *Ground Gas Handbook*, Whittles Publishing, Dunbeath, UK.

Appendix A: DQOs and the environmental data life-cycle process

Environmental data life-cycle process

The data quality objectives (DQOs) are not a standalone process, but rather are an integral part of the USEPA's project-level quality system for the collection of environmental data. This system includes various components which, taken together, form an environmental data life-cycle process for environmental assessments. This highlights that, as used by the United States Environmental Protection Agency (USEPA), the DQOs process is one component of a larger, project-level quality system aimed at producing 'defensible products and decisions'.

Implicitly, partial or incomplete application of any of the individual components will result in data that is unlikely to achieve all the desired outcomes, that is, defensible products and decisions. The components of the USEPA's project-level quality system (USEPA 2002) are shown in Figure 7 and summarised below.

- **Systematic planning** – to identify the expected outcome of the project, the technical goals, the cost and schedule, and the acceptance criteria for the final result before a project begins. The DQOs process includes developing or refining the conceptual site model (CSM).
- **Sampling design** – fundamental to data collection for scientifically based decision making, which seeks to ensure that the data collection program collects appropriate and defensible data that accurately represents the problem being investigated
- A **quality assurance project plan** (QAPP) – to document performance criteria and the project-specific plan for obtaining the type, quality, and quantity of data needed for a specific use. It is analogous to a sampling, analysis and quality plan (SAQP) in the Australian context. In addition to systematic planning and sampling design, inputs to the QAPP include:
    - the data quality indicators (DQIs) planning, to address the principal data quality attributes and the associated measurement quality objectives (MQOs)
    - any standard operating procedures (SOPs) to document the procedures necessary to carry out routine or repetitive administrative and technical activities.
- **Conducting the study or investigation** – the implementation of the study or investigation, based on the preceding inputs. This can include technical assessments (project audits), such as reviews to document the degree to which the procedures and processes specified in the QAPP are being implemented.
- **Data verification and validation** – to determine if data has been collected in accordance with the QAPP with respect to compliance, correctness, consistency and completeness, and to evaluate the technical usability of the data with respect to the planned objectives or intention of the project.
- **Data analysis and interpretation** – to provide a scientific and statistical assessment to determine whether the data is of the right type, quality and quantity to achieve the objectives of a project.

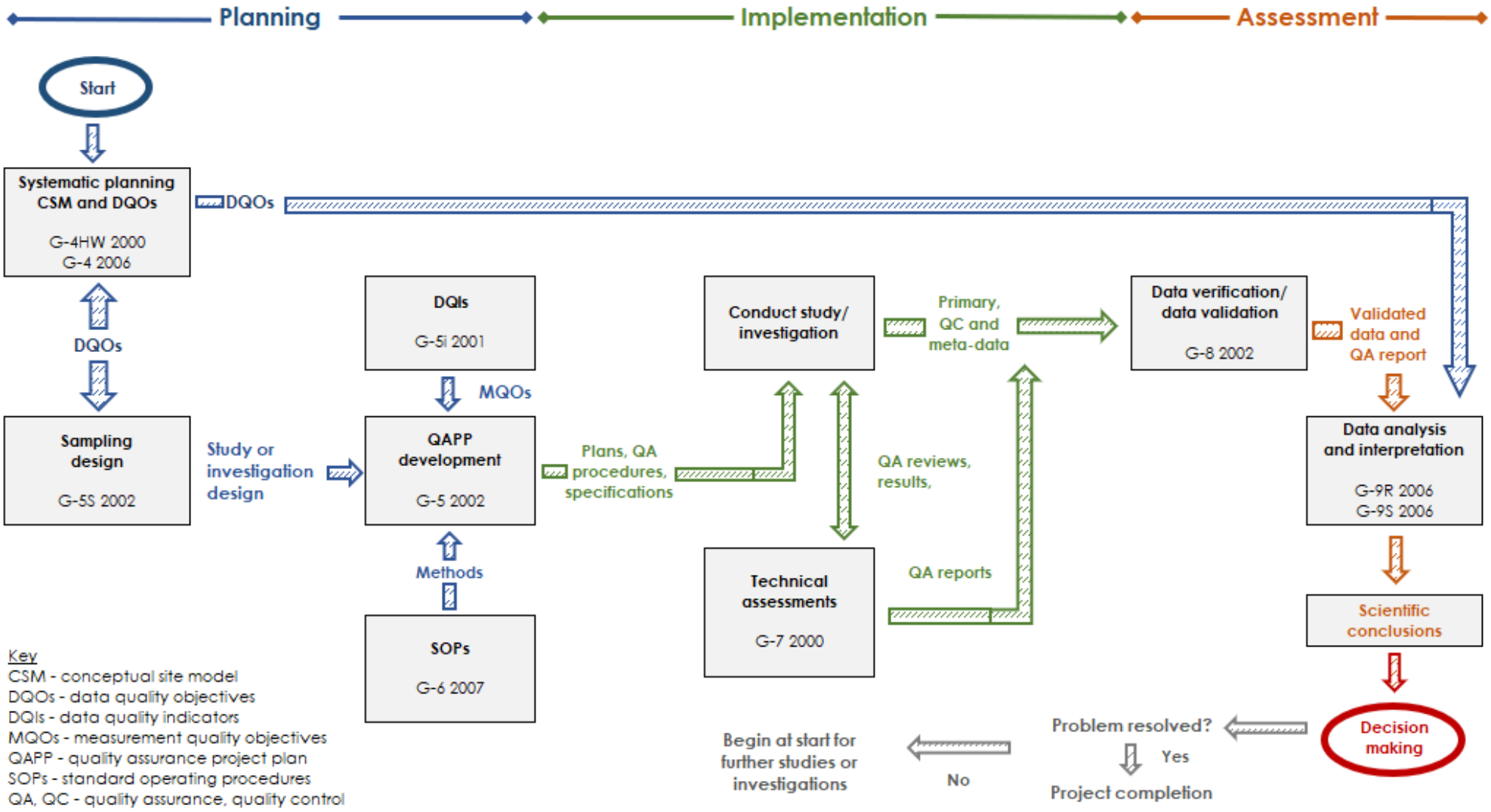

Figure 7 Environmental data life-cycle process and relevant USEPA guidance

Figure modified from USEPA G-4HW 2000 and www.epa.gov/quality/agency-wide-quality-system-documents#preview.

This is an iterative process, and prior steps may need to be revisited based on the outcomes of later steps.

Clarification of ‘quality’

Within the DQOs process, is the term “quality” is used in the context of the quality of the estimates derived from the data, or to refer to a desired level of quality, such as the statistical precision of the data.

USEPA (2000) points out that the DQOs process represents an evolution from valid concerns about the quality of data, to concerns about the quality of the decisions that will be made from the data. It notes that

> *[d]ata quality, as a concept, is meaningful only when it relates to the intended use of the data. Data quality does not exist without some frame of reference; one must know the context in which the data will be used in order to establish a yardstick for judging whether or not the data set is adequate.*

For decision problems, decision quality measures include such things as the true value of the parameter, the selected action level, the probability of deciding that the parameter exceeds the action level, the statistical hypothesis, and the boundaries of the **grey region** (defined in chapter 7). To this end, USEPA (2000) says that

> *elements of a systematic documented planning approach include … determination of the quantity of data needed and specification of performance criteria for measuring quality (DQO Step 6)*

and that an outcome of the DQOs process should be the specification of some measure of the desired quality of a decision rule, which also takes uncertainty into account.

Accordingly, when conducting or reviewing DQOs, it is imperative to distinguish between quality as “suitable quality to support the required decision”, on the one hand, and the assessment of data usability in the context of DQIs and MQOs on the other.

The use of the term ‘quality' by the USEPA is illustrated by its definition of data quality assessment (DQA), described as data analysis and interpretation, in Section 1 and Figure 7 of this document. The five steps of statistical DQA are described by USEPA (2006, G-9R) as:

- review the project objectives and sampling design – undertaken by reviewing the systematic planning objectives to ensure the context of the investigation is understood. This review also allows an assessment of the quality of the data to be made, in terms of addressing the objective of the investigation as well as its quality for use
- conduct a preliminary data review – this involves a preliminary review of the data, including an assessment of quality assurance (QA) reports to identify any anomalies. The data should also be assessed for its distribution and patterns, and to identify any potential outliers
- select the statistical method – this will be guided by the previous two steps. The choice of statistical method will be based on the objectives of the investigation and the dataset
- verify the assumptions of the statistical method – review the statistical method used to ensure that any assumptions made are justified. (For example, if the dataset contains outliers, some statistical methods may lead to biased conclusions.)
- draw conclusions from the data – draw the conclusions based on the findings of the data in line with the objectives of the investigation. The conclusions should include an assessment of the sampling design and whether or not it can be used in other scenarios. Conclusions should be documented and coherently justified so that all stakeholders can understand how they were reached.

The DQOs process, as designed and implemented by the USEPA, is a component of a multi-stage project life cycle that primarily addresses the sampling design and statistical aspects of a proposed environmental evaluation. The DQA component is a stage after the implementation of the data collection: it addresses the statistical rigour of the investigation and the achievement of the project objectives, and relies on both the data collected as part of the study or investigation and the DQOs developed during the planning phase of the study or investigation (as shown in Figure 7).

Logically, if the DQOs are not correctly developed, the whole process cannot result in defensible products and decisions, as the later steps are founded on the assumption of appropriate systematic planning, with the DQOs again incorporated into the assessment stage as part of arriving at a defensible decision.

Decision problems and estimation problems

The seven-step DQOs process, as shown in Figure 8, is a method for systematic planning that includes options for the type of problem to be addressed, based on the intended use of the data to be collected. The two primary types of intended use are classified as **decision making** and **estimation**.

Decision making is defined as making a choice between two alternative conditions. USEPA (2006, G-4) describes it as follows:

> *This is where statistical methods help a decision maker structure the decision problem. The methodology of 'classical' Neyman–Pearson statistical hypothesis testing provides a framework for setting up a statistical hypothesis, designing a data collection program that will test that hypothesis, evaluating the resulting data, and drawing a conclusion about whether the evidence is sufficiently strong to reject or (by default) accept the hypothesis, given the uncertainties in the data and assumptions underlying the methodology. The DQO Process has been designed to support a statistical hypothesis testing approach to decision making.*

Estimation is used when the objective of an investigation is to evaluate the magnitude of some environmental parameter or characteristic, noting that the resulting estimate may be used in further research, as an input to a model, or perhaps eventually to support decision making. USEPA (2006, G-4) notes that:

> *the defining characteristic of an estimation problem versus a decision-making problem is that the intended use of the estimate is not directly associated with a well-defined decision.*

To illustrate the two types of problems, consider the project requirement to compare metal concentrations in soil at a site with the ambient background concentrations (ABCs). The NEPM (2013, B5b) describes that the preferred method for estimating ABCs is by direct measurement at a clean reference site, with a soil type comparable to that of the site being examined. While no specific decision can be made as part of estimating the ABCs, some statistical rigour is desirable in estimating the metal concentrations. This is considered an estimation problem, with the number of samples required in part determined by the variance in the metals data.

Once suitable ABCs have been determined, the site data needs to be compared to the ABCs. As a well-defined decision is required – whether the site concentrations are greater than, or not greater than, the ABCs – this is a decision problem. The project requirement could also include such things as comparing groundwater monitoring wells upgradient of a source to wells downgradient of a source, or comparing background surface water quality to the water quality of a release (or potentially contaminating discharge) into the surface waters. However, decision problems, such as comparing site data to specified levels, are the more common type of problem.

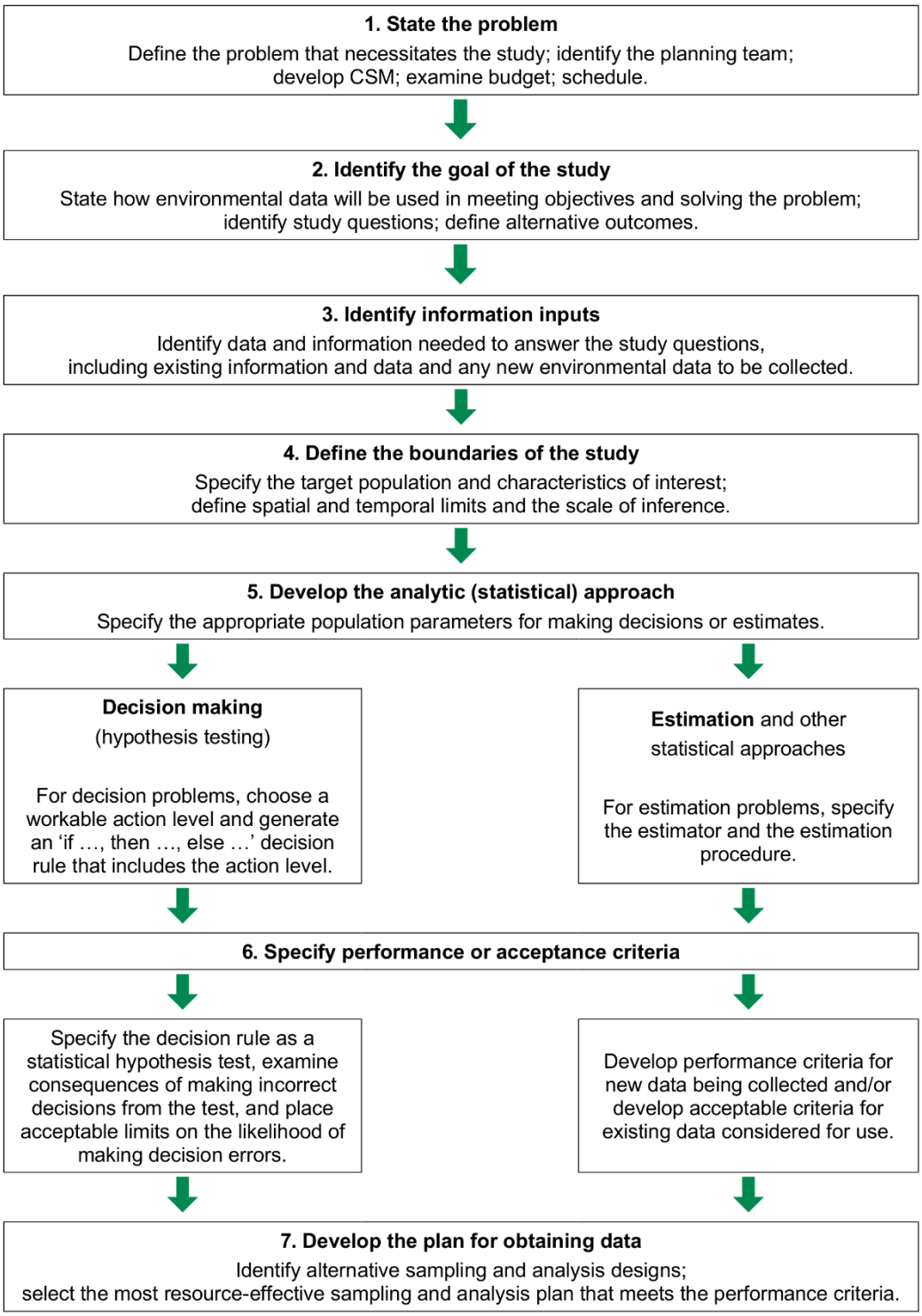

Figure 8 Overview of the USEPA DQOs process

Modified from USEPA (2006, G-4)

For complex problems, such as multiple contaminant types and a number of impacted media, more than one decision may be required, or estimates of multiple parameters may need to be combined. These multiple decisions or estimates may combine or impact on each other in addressing the problem. In addition to CSMs, the DQOs process recommends the use of flowcharts, logic diagrams, influence diagrams, etc., to illustrate, document and manage these problems.

For addressing multiple but specific technical questions, the use of modules is recommended, grouped by logical categories depending on the magnitude of the problem. Where multiple media are involved, this may be the logical grouping, e.g. soil, groundwater and soil gas. Where multiple contaminant types exist, this may be the logical grouping, e.g. volatiles, metals, semi-volatiles and asbestos. For sites that are large with varied site histories, modules for specific decision areas may also be warranted.

The activities to be undertaken as part of step 5 for decision problems are:

- specify the population parameter considered to be important to make inferences about the target population (e.g. mean, 95% upper confidence limits (UCL) of the arithmetic mean, median or percentile)
- choose an action level (using information identified in step 3) that sets the boundary between one outcome of the decision process and an alternative, and verify that there are sampling and analysis methods that have detection limits below the action level
- construct the theoretical 'if ..., then ..., else ...' decision rule by combining the true value of the selected population parameter, the action level, the scale of decision making (Step 4), and the alternative actions (step 2).

For decision problems, the outputs for this step are:

- identification of the population parameters most relevant for making inferences and conclusions on the target population
- the 'if ..., then ..., else ...' theoretical decision rule based upon a chosen action level.

For estimation problems, step 5 involves specifying the estimator by combining the selected population parameter (e.g. mean) with the scale of the estimation and other population boundaries from step 4, then applying the estimation procedure (e.g. 95% confidence interval).

Specify performance or acceptance criteria

Step 6 of the DQOs process establishes quantitative criteria known as performance or acceptance criteria, or data quality objectives (DQOs). The DQOs vary depending on the type of problem being addressed:

- for decision problems, the DQOs are typically tolerable limits on the probability or chance (risk) of the collected data leading to making an erroneous decision (e.g. confidence levels)
- for estimation problems, the DQOs are typically an acceptable uncertainty, e.g. the width of an uncertainty band or interval, associated with a point estimate at a desired level of statistical confidence (e.g. confidence intervals).

USEPA (2006, G-4) notes that performance criteria represent

> *the full set of specifications that are needed to design a data or information collection effort such that, when implemented, generate newly-collected data that are of sufficient quality and quantity to address the project's goals*

while acceptance criteria are

> *specifications intended to evaluate the adequacy of one or more existing sources of information or data as being acceptable to support the project's intended use.*

Accordingly, the DQOs process should be used to generate **performance criteria** for new environmental data and **acceptance criteria** for existing information and data. Where existing data and information do not meet the acceptance criteria, they may need to be classified as estimates, and new information and data may need to be obtained, subject to the specified performance criteria.

Develop the plan for obtaining data

Step 7 of the DQOs is where a resource-effective, field investigation sampling and analysis design is to be developed, to generate data that satisfies the decision performance criteria identified in step 6, and the other requirements specified in the preceding steps of the DQOs. It is usual to iterate between steps 6 and 7 when assessing and refining the design parameters against the project objectives and constraints. The output of step 7 is the sampling and analysis design that is documented in the sampling and analysis quality plan (SAQP).

For most field investigations, a probabilistic sampling approach is necessary to provide a scientific basis for extrapolating the results from samples to the entire site or decision area. USEPA (2000) comments that

> *[b]y combining an effective probabilistic data collection design with a statistical hypothesis test, the decision maker will be able to optimize resources such as funding, personnel, and time while still meeting DQOs.*

For common probabilistic designs, information regarding the expected variability of the contaminants is necessary, as determining a minimum sample size relies on an estimate of total variability in the data to be collected (USEPA 2006, G-4). Such estimates may be determined from existing data from the site (or from similar sites). If no existing data are available, you may need to make limited field investigations to determine a preliminary estimate of variability, unless the site history is well documented and has no data gaps.

Information derived from the systematic planning that is used as input to the sampling and analysis design process includes:

- the target population and spatial/temporal boundaries of the study (DQO step 4)
- the preliminary estimation of variance (DQO step 4)
- the purpose of the data collection – hypothesis testing, estimating a parameter with a level of confidence, or detecting hotspots, or a combination (DQO step 5)
- the statistical parameter of interest, such as the mean, the 95% UCL of the arithmetic mean, the median, percentile, trend, slope, etc. (DQO step 5)
- limits on decision errors and precision, in the form of false acceptance and false rejection error rates and/or the overall precision specifications (DQO step 6).

Step 7 includes the development of alternative data collection designs, to assess which design best limits the total study error to tolerable levels to satisfy the decision performance criteria. To generate alternative designs, aspects to be considered include:

- type of samples collected
- sampling design
- sample selection technique
- number of samples
- spatial/temporal locations of samples
- field sampling or analytical methods used
- number of analyses per sample
- number of replicate analyses performed on samples.

USEPA (2000, G-4HW) notes that

> *[d]esigns that balance the number of field samples with the number of laboratory analyses should be considered*

and further comments that

> *[t]wo mathematical expressions are necessary for optimizing each data collection design alternative in relation to the decision performance criteria. First, a tentative method for analysing the resulting data (e.g. a student's t-test or a tolerance interval) should be specified, along with any available sample size formulas corresponding to the proposed method. This information will be used to solve for the minimum sample size that satisfies the*

decision maker's limits on decision errors. Second, a cost function that relates the total number of samples to the costs of sampling and analysis should be developed. This information will be used to compare the cost-effectiveness of different sampling designs.

It is also recommended that a sensitivity analysis be performed on the alternative designs, to determine if changes to design assumptions significantly affect the design's ability to achieve the expected decision error limits, and the associated impacts on costs or resources required. For example, if contaminant variability is higher than estimated, will the proposed number of samples meet the performance criteria? Or will more samples be required, leading to a higher cost?

Once a data collection design has been selected, the design parameters and key assumptions must be documented, so that the collected data can be analysed and interpreted to determine whether the data are of the type, quality and quantity required to achieve the project objectives. Within the USEPA's environmental data life-cycle process, this occurs during the assessment stage as part of the data analysis and interpretation.

While the choice of the sampling and analysis design will have an impact on the data quality indicators (DQIs), and the DQIs should be considered as part of step 7, the DQOs do not specifically address the DQIs or their acceptance criteria. In much the same way as step 5 of the DQOs is conducted under the assumption that one has "access to perfect information on unlimited data" (USEPA 2006, G-4), the DQOs assumes that all of the data collected is usable, at least until step 7 and the subsequent development of the SAQP, as discussed below.

Data quality indicators and measurement quality objectives

Data quality is a measure of the degree of acceptability or usability of sampling data for a particular purpose. It relates to both sampling errors and measurement errors: as USEPA (2006, G-4) notes, sampling error is generally much larger than measurement error, and consequently needs a larger proportion of resources to control.

Figure 9 shows an example of how total study error can be broken down into components that are associated with the various activities as part of environmental sampling and analysis. While interrelated, the activities associated with sampling error are addressed through the DQOs process and sampling design, and the activities associated with measurement error are addressed through the DQIs. The magnitude of total study error should be controlled by generating an appropriate sampling design and choosing suitably accurate measurement techniques.

In regard to measurement errors, certain qualitative and quantitative characteristics of the collected data, (that is the data quality attributes), can be defined and measured. DQIs are the quantitative and qualitative measures, or indicators, of the principal data quality attributes.

The principal data quality attributes are precision, accuracy[3], representativeness, comparability, completeness, and sensitivity (PARCCS), with precision, accuracy/bias and sensitivity being defined and measured in quantitative terms, and representativeness, comparability and completeness having more qualitative definitions. MQOs are the acceptance criteria or goals for the data quality attributes, as measured by the project DQIs.

[3] USEPA (2001,G-5i) describes that the 'A' in PARCCS refers to accuracy instead of bias, and that this substitution of 'A' for 'B' occurs because PARCCS is a historically recognised and familiar acronym, and some analysts believe accuracy and bias are synonymous, although accuracy is actually comprised of random error (precision) and systematic error (bias).

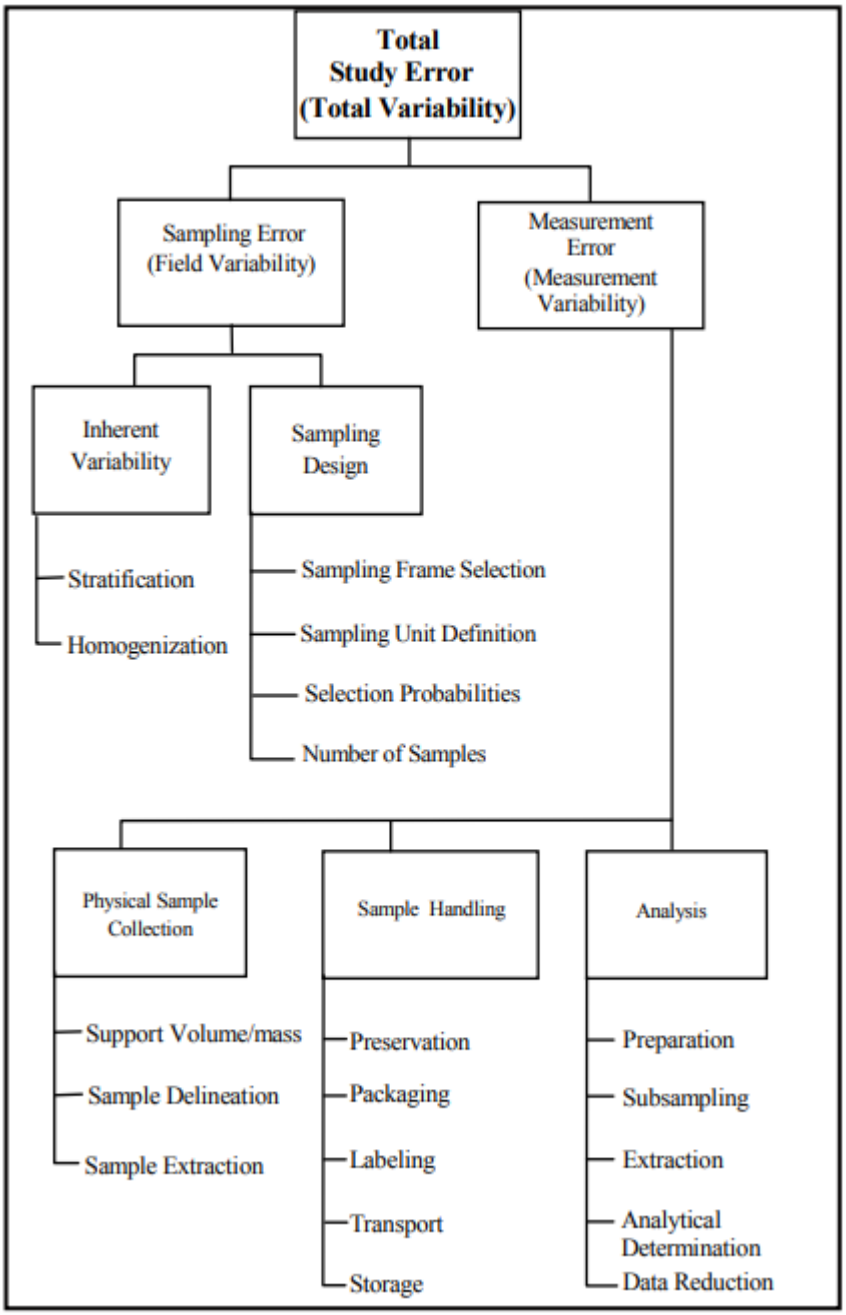

Figure 9 Total study error by components

From USEPA (2006, G-4)

USEPA (2001, G-5i) states that DQOs are qualitative and quantitative study objectives for the collection of environmental data, and that

> *[h]istorically, DQIs sometimes have been incorrectly equated with DQOs, which are specifications for decision making.*

DQIs are not the focus of this guidance but they are important inputs to the sampling design process, as they indicate whether the resulting data is expected to meet the DQOs. The process of establishing measurement quality objectives (MQOs) – the goals set for the DQIs – is an integral part of designing the study.

After collecting the data you should determine its adequacy or usability, as part of the data-verification and data-validation component of the data life-cycle process. Use a weight of evidence/multiple lines of evidence approach and take into account both the project specific requirements and the stage of the data collection event. Data quality requirements for final data (characterisation or validation), will generally be more stringent than for preliminary data.

USEPA (2001, G5i) notes that:

> *the highest interest is in whether the data set will support a decision with the desired degree of certainty. It is important to consider the performance and representativeness of the measurement effort prior to reaching conclusions regarding data adequacy; however, at this point it is less critical to determine if each and every goal set for given DQIs (i.e. the MQOs) was achieved. If adequate sensitivity was achieved, and bias is 'under control,' the key issues revolve around whether an adequate number of samples was obtained, given the observed measurement, spatial and temporal variability, and given the actual magnitude of the measurements made (relative to levels of concern). If a data collection effort fails to generate adequate data, then interest in DQIs is heightened.*

In this context, an adequate 'number of samples' must explicitly include an evaluation of the sample representativeness. If the samples cannot be shown to be representative of the condition of the site or decision area, in the context of the decision to be made, evaluation of the measurement quality in isolation cannot demonstrate if the data is of a suitable quality to support the required decision (Crumbling 2001).

Project level planning

At a broader, 'whole of project' level, assessment of site contamination projects also requires systematic planning, to ensure an appropriate level of project planning and documentation to manage the overall project. Projects often consist of a number of separate investigations, and always involve sequential steps of assessment and management, from the preliminary investigation through to remediation, and ultimately validation and completion.

The systematic planning process for sampling should identify the objectives of the site investigations and establish the types of information needed to make the various environmental decisions required. The DQOs process is often used: the NEPM (2013, B2) notes that the "DQO process is applicable at both the project level (for example, is the site suitable for development?) and at the investigation level".

As the DQOs process was designed for addressing multiple but specific technical questions, the later specific design steps of the DQOs do not generally apply to the broadscale, project level planning. The DQOs process can be applied in various ways at different stages of the project. USEPA (2000) points out that

> *during early site assessment phases, where investigators generally examine existing site information and conduct site reconnaissance, planning teams can benefit from the qualitative DQO steps, but may have to allow for a more liberal interpretation of the quantitative steps.*

Table 5 gives the recommended approach for applying the DQOs process across all levels of site contamination assessment projects.

Table 5 Recommendations for implementation of DQOs

Project requirement	Applicable DQOs step
Project level	All steps, with only generalised information in steps 5, 6 and 7. Document in a project-level SAQP.
Individual investigations	All steps, with steps 5, 6 and 7 fully addressed for simpler investigations. Document in an investigation-level SAQP, or a project-level SAQP with specific investigation requirements added.
Complex investigations with multi-contaminant types, media, or site histories, including risk assessment investigations or offsite investigations of surface waters, sediments, biota etc.	Address project-level requirements at project level and use modules to address specific technical questions. Document in either a project-level or investigation-level SAQP, with specific module requirements added.

References

Crumbling DM 2001, *Applying the Concept of Effective Data to Environmental Analyses for Contaminated Sites* (EPA 542-R-01-013), U.S. Environmental Protection Agency, Washington DC.

National Environment Protection Council (NEPC) 2013, *National Environment Protection (Assessment of Site Contamination) Measure 1999,* National Environment Protection Council, Canberra. Includes Amendment Measure 2013 (No. 1), as:

- Schedule B(2), Guideline on Site Characterisation
- Schedule B(5b), Guideline on Methodology to Derive Ecological Investigation Levels in Contaminated Soils.

US Environmental Protection Agency (USEPA) 2000, *Data Quality Objectives Process for Hazardous Waste Site Investigations (QA/G-4HW)*, EPA/600/R-00/007, USEPA, Washington DC.

US Environmental Protection Agency (USEPA) 2000, Guidance on Technical Audits and Related Assessments for Environmental Data Operations (QA/G-7), EPA/600/R-99/080, USEPA, Washington DC.

US Environmental Protection Agency (USEPA) 2001, *Guidance on Data Quality Indicators (QA/G-5i)*, USEPA, Washington DC.

US Environmental Protection Agency (USEPA) 2002, *Guidance on Choosing a Sampling Design for Environmental Data Collection for Use in Developing a Quality Assurance Project Plan (QA/G-5S)*, EPA/240/R-02/005, USEPA, Washington DC.

US Environmental Protection Agency (USEPA) 2002, *Guidance on Environmental Data Verification and Data Validation (QA/G-8)*, EPA/240/R-02/004, USEPA, Washington DC.

US Environmental Protection Agency (USEPA) 2002, *Guidance for Quality Assurance Project Plans (QA/G-5)*, EPA/240/R-02/009, USEPA, Washington DC.

US Environmental Protection Agency (USEPA) 2002, *Overview of the EPA Quality System for Environmental Data and Technology*, EPA/240/R-02/003, USEPA, Washington DC.

US Environmental Protection Agency (USEPA) 2006, *Data Quality Assessment: A Reviewer's Guide (QA/G-9R)*, EPA/240/B-06/002, USEPA, Washington DC.

US Environmental Protection Agency (USEPA) 2006, *Data Quality Assessment: Statistical Methods for Practitioners (QA/G-9S)*, EPA/240/B-06/003, USEPA, Washington DC.

US Environmental Protection Agency (USEPA) 2006, *Guidance on Systematic Planning Using the Data Quality Objectives Process (QA/G-4)*, EPA/240/B-06/001, USEPA, Washington DC.

US Environmental Protection Agency (USEPA) 2007, *Guidance for Preparing Standard Operating Procedures (SOPs) (QA/G-6),* EPA/600/B-07/001, USEPA, Washington DC.

Appendix B: Data-quality objectives: worked example

This worked example is a decision problem, based on the following hypothetical scenario.

A site operated as a sheepskin processing facility between the 1950s and the early 1970s. It had a combined storage shed/office/amenities building and a workshop that incorporated asbestos-containing cement sheeting. The sheepskins were dried on racks housed in timber and corrugated-iron sheds, and also directly on the ground.

Arsenic (As) was used as a biocide to treat the skins, and a 500-litre vat for mixing and storing it was located in one of the drying sheds. The As was in powder form and was mixed with water onsite, with the treatment solution applied using a network of irrigation pipes below the celling. The spent As solution was discharged to the land surface.

Anecdotal information also suggests some fuel was stored in the site, although it has not been confirmed if this was above ground or below ground. It is likely that other chemicals and fuels (e.g. paints, solvents, greases, etc.) were used in the workshop area. It is not known if a significant amount of waste was buried onsite, although some wastes have been dumped, e.g. empty 200-litre drums. Several areas of what appears to be disturbed natural soils, and some building and demolition wastes beneath former structures, have been identified.

Many investigations have been conducted in the area of the old processing facility, located at the crest of a small hill: they have included judgmental and systematic soil sampling, and judgmental groundwater sampling. These have found no organic contaminants or asbestos fibres but have identified a number of soil locations where elevated As occurs, as well as some copper (Cu) and zinc (Zn) (assumed to be related to galvanised materials in the sheds) and some lead (Pb) (assumed to be related to Pb paint flakes). No groundwater impacts were identified. The old processing facility has been considered to be sufficiently characterised to not require further investigations.

An area of filling adjacent to the former processing centre is approximately 80 m x 160 m and is understood to consist of between 2 m and 3 m of fill material. This area requires characterisation to determine the land-use suitability, or the appropriate waste classification. The intended land use is residential with accessible soil (HIL-A), and it is proposed to develop the decision area into 400 m^2 residential blocks, although the specific lot layout has not yet been determined.

The DQOs process output for the proposed investigation for the filled portion decision area is shown in Table 6.

Table 6 DQOs process steps and their outputs

No.	DQO process step	Outputs of DQOs process step
1	**State the problem** – assemble an effective planning team, describe the problem and examine the resources for investigating the problem.	–
1.1	Write a brief summary of the contamination problem.	An area of filling of approximately 80 m x 160 m and between a depth of 2 m and 3 m from the surface requires characterisation to determine the land use suitability or the appropriate waste classification.
1.2	Identify members of the planning team.	Landowner/developer, planning consultant and site contamination assessment consultant.
1.3	Develop/refine the conceptual site model (CSM), including a summary of the exposure scenarios.	Contaminants – potentially metals (As, Cu, Pb and Zn), organics (TRHs, BTEX, PAHs and OCPs/OPPs), and asbestos fibres/fragments. Sources – buried building and demolition and wastes, and former chemical wastes and drums. Receptors – site maintenance workers and trespassers, as the site is fenced and secured. If developed, site workers (surface and subsurface), residents and visitors (adults and children). Pathways – dermal contact, inhalation of dust and ingestion have been identified as the pathways of concern. Further assessment of groundwater and/or soil gas will be considered based on the findings of this investigation.
1.4	Specify the available resources and constraints, such as relevant deadlines for the study, budget, availability of personnel and schedule.	The site contamination assessment consultant has the available capacity to conduct the investigation using appropriate subcontractors (excavator and laboratory). While the developer seeks close out of the issue within the next three months, there are no practical constraints, as the land is identified as high-value and sufficient budget is available. Additional investigations, or remediation will be conducted as required.
2	**Identify the goals of the study** – identify the principal study question(s) and potential alternative actions (with implications), and combine these to make statements on the decision problem.	–
2.1	Identify the principal study question(s).	Is the fill material suitable for a residential land use based on contaminant levels and aesthetic concerns? If not, what disposal options are available, i.e. what is the waste classification?
2.2	Identify the alternative outcomes or actions that could result from resolution of the principal study question(s).	The alternative outcomes will be: • the fill material is suitable for residential land use (HIL-A) or • the fill material is not suitable for the proposed land use and needs to be partially or fully removed from site to allow development.

No.	DQO process step	Outputs of DQOs process step
2.3	For decision problems, combine the principal study questions and the alternative actions into decision statements.	If the contamination status of the fill material is acceptable, the material can remain onsite. If the contamination status of the fill material is unacceptable, consider the remediation hierarchy.
3	**Identify information inputs** – identify the information needed to formulate and investigate the problem and confirm that appropriate sampling and analytical methods are available.	–
3.1	Identify the information that will be required to resolve the decision statements/estimation, including existing information and new environmental data, and identify the sources for each item of information required.	Soil data collected as part of this investigation, including field samples and analytical samples. No previous investigation of the area has been conducted, although information from the investigation of the adjacent facility will inform this investigation.
3.2	Identify the information needed to establish the action level.	Investigation criteria will be sourced from: • NEPM (2013) Schedule B1, HILs for residential with accessible soil • NSW EPA (2014) Waste Classification Guidelines.
3.3	Confirm that appropriate sampling and analytical methods exist to provide the necessary data.	Sampling and analytical methods will be consistent with existing guidance, including the NEPM (2013, B2 and B3). Analytical laboratories will be NATA accredited and use analytical methods based on NEPM, USEPA and APHA methods.
4	**Define the boundaries of the study** – define the target population and the spatial and temporal boundaries associated with the population; examine any practical constraints to collecting data, and factors that affect the selection of the unit that defines the scale of sampling and the scale of decision making or estimation.	–
4.1	Define the target population of interest and its relevant spatial boundaries.	The area of filling is approximately 80 m x 160 m and is between a depth of 2 m and 3 m from the surface. The decision area is approximately 12,800 m^2 and contains an estimated 32,000 m^3 of fill. It is believed to be reworked natural material, derived from the levelling of the process area. The natural soil is silty to sandy clay with frequent weathered parent material (quartzite and phyllite) gravel. Uncontaminated soils in previous site investigations were found to be fairly homogenous in regard to metal concentrations, i.e. expected relative standard deviation (RSD) < 50%.
4.2	Define what constitutes a sampling unit.	Sampling units will consist of: • field samples of appropriately described and logged samples which are field screened • analytical samples of the laboratory-specified sample jar quantity.

No.	DQO process step	Outputs of DQOs process step
4.3	Specify temporal boundaries and other practical constraints associated with sample/data collection.	To achieve the three-month schedule for problem resolution, the field investigation should start within two weeks of the investigation plan (SAQP and commercial) being accepted. There are no site access restrictions for personnel once they are inducted and project-approved. The decision area is open with a light grass covering only and directly accessible without obstructions.
4.4	Specify the smallest unit on which decisions or estimates will be made.	The decision is to be based on the complete decision area. However, following data analysis, some form of segregation may be considered, i.e. some of the decision area may be suitable for HIL-A and some may require offsite disposal.
5	**Develop the analytic (statistical) approach** – develop a logical 'if …, then …, or …' statement that defines the conditions that would cause the decision maker to choose among alternative actions.	–
5.1	Specify the statistical parameter that characterises the population of interest, such as mean, median, maximum, 95% upper confidence limit (UCL) of the arithmetic mean, proportion, etc.	The 95% UCL of the arithmetic mean will be the key statistical parameter. The data evaluation will include that: • the 95% UCL arithmetic mean to be ≤ criterion • no individual sample to exceed 250% of the criterion • the sample standard deviation to be < 50% criterion. Additional considerations will include aesthetic requirements, including no odours or staining, no waste materials and no monolithic deposits as per the NEPM (2013, B2).
5.2	Specify the action level for the decision.	To determine if the material is suitable for the HIL-A land use, analytical action levels are to be based on the NEPM HILs (2013, B1). If the material is not suitable for the HIL-A land use, the material will be classified in accordance with EPA (2014) for offsite disposal. Samples will be held at the laboratory for additional analyses, including leachate analysis following TCLP extraction, if required.
5.3	Confirm that measurement detection will allow reliable comparisons with the action level.	Samples will be submitted to NATA-accredited laboratories. The laboratories' analytical LORs are suitably below the adopted criteria. Note: to achieve an acceptable limit of reporting for asbestos fines and fibrous asbestos, the method may not be NATA-accredited but undertaken using in-house methods for quantification.
5.4	Combine the outputs from the previous DQOs steps and develop an 'if ..., then ..., else ...' theoretical decision rule based on the chosen action level.	**If** the statistical parameters (or aesthetics) of the sampling data exceed the applicable action levels, **then** offsite disposal of the fill material will be required, **otherwise**, **if** the statistical (and aesthetic) parameters are below the applicable action levels, **then** the fill material will be determined to be suitable for a HIL-A land use.

No.	DQO process step	Outputs of DQOs process step
6	**Specify performance or acceptance criteria** – specify probability limits for false rejection and false acceptance decision errors.	–
6.1	Specify the decision rule as a statistical hypothesis test.	The null hypothesis is that the fill material is contaminated and exceeds the adopted criteria. The alternative hypothesis is that the fill material is not contaminated above the adopted criteria.
6.2	Examine consequences of making incorrect decisions from the test.	Possible decision errors include: • the fill material being accepted as suitable for a HIL-A land use when it is not, thereby potentially risking human health or environmental impacts • unnecessary disposal of the fill material offsite, imposing needless financial and resource burdens on the development project and resulting in an inappropriate waste classification.
6.3	Place acceptable limits on the likelihood of making decision errors, including acceptable alpha (α) and beta (β) risk levels.	Stated hypotheses: • null hypothesis (H_0): the 95% UCL, and other requirements, are > the action level; and • alternate hypothesis (H_A): the 95% UCL, and other requirements, are ≤ the action level. Potential outcomes include Type I and Type II errors: • Type I error of determining the fill material is acceptable for the proposed HIL-A land use when it is not (wrongly rejects true H_0). • Type II error of determining the fill material is unacceptable for the proposed HIL-A land use when it is (wrongly accepts false H_0). For performance criteria, the acceptable limits on the likelihood of making decision errors to be applied are: • alpha risk (Type I error) of $\alpha = 0.05$ • beta risk (Type II error) of $\beta = 0.2$. No previously collected data is available for use, therefore acceptance criteria are not required.
7	**Optimise the design for obtaining data** – identify a resource-effective sampling and analysis design for generating data that is expected to satisfy the DQOs.	–

No.	DQO process step	Outputs of DQOs process step
7.1	Document the final sampling and analysis design, along with a discussion of the key assumptions underlying this design.	To allow statistical inference, a probabilistic systematic strategy is to be adopted. As the proposed development is based on 32 residential lots of 400 m^2, 32 sample locations were selected such that the density equates to one sample location per lot. Using a regular square grid size of 20 m, the grid will consist of 4 x 8 cells, with the sample locations within each cell to be selected using a random number generator. The grid lines will be designated A to D from north to south (short axis) and 1 to 8 from west to east (long axis), such that the first node will be A1, through to D8. There will be 32 sample locations. Test pits will be excavated at each location to the underlying natural material as identified by the remnant A-horizon, at a depth of 2–3 m. Two field samples will be collected at each sample location at the surface (a depth of 0.01 m) and an approximate depth of half the total test pit depth, such that there will be 64 field samples. Sixteen analytical samples are to be analysed initially, with the remaining field samples to be held at the laboratory. It was assumed that a relative standard deviation of about 75% could be expected and based on a maximum probable error (MPE) of between 30% and 50%, 16 samples were calculated as appropriate for analysis using the MPE method for determining the number of samples required. Based on the size of the decision area, this sampling design results in: • one sample location per forecasted residential block (400 m^2) • one field sample per 500 m^3 • one analytical sample per ~2,000 m^3. This design is theoretically capable of detecting a minimum hotspot diameter of 23.6 m.
7.2	Detail how the design should be implemented, together with contingency plans for unexpected events.	The field methods for sample collection, handling, and analysis (at analytical laboratories) are described in the project-level standard operating procedures (SOPs). Contingencies include collecting additional samples from material that is significantly different from the reworked natural, and conducting additional analyses where field indicators (staining, odours, field screening results) suggest other contaminants.
7.3	Determine the quality assurance and quality control (QA/QC) procedures that are to be performed to detect and correct problems to ensure defensible results.	The required field QA, and the field and laboratory QC, are described in the project-level SOPs. These include both the data quality indicators (DQIs) and the associated measurement quality objectives (MQOs).
7.4	Document the operational details and theoretical assumptions of the selected design in the SAQP.	Theoretical assumptions include: • the fill material consists of reworked natural material, and only minor wastes exist (if any) • surficial impacts from overland flow from the adjacent facility and burial of wastes are the modes of contamination expected • the fill material is relatively homogenous • the remnant A-horizon will be readily discernible from buried grass and organic soil.

The resulting detected metal data from the ‘implementation’ of this investigation is summarised in Table 7.

Table 7 Summary of analytical results – metals in soil (mg/kg)

Sample/descriptor	Arsenic	Chromium	Copper	Lead	Nickel	Zinc
LORs	5	2	5	5	2	5
Analytical						
Analytical sample B2-01	103	12	34	20	18	11
Analytical sample B2-02	50	21	30	7	2	10
Analytical sample D2-01	43	26	83	17	14	35
Analytical sample D2-02	9	10	29	14	5	12
Analytical sample A4-01	203	4	260	18	12	232
Analytical sample A4-02	54	5	55	17	9	41
Analytical sample C4-01	341	19	401	133	7	543
Analytical sample C4-02	34	17	46	16	10	13
Analytical sample B6-01	71	18	24	14	5	9
Analytical sample B6-02	14	6	8	17	12	5
Analytical sample D6-01	62	11	51	15	3	36
Analytical sample D6-02	6	4	18	16	24	10
Analytical sample A8-01	27	17	61	16	4	24
Analytical sample A8-02	7	10	38	20	13	10
Analytical sample C8-01	24	15	39	12	6	8
Analytical sample C8-02	13	16	17	14	19	7
Descriptive statistics						
Number of samples	16	16	16	16	16	16
Number of detects	16	16	16	16	16	16
Percentage non detects	0%	0%	0%	0%	0%	0%
Maximum	341	26	401	133	24	543
Third quartile	64.3	17.3	56.5	17.3	13.3	35.3
Median value	38.5	13.5	38.5	16.0	9.4	11.5
First quartile	13.8	9.0	27.8	14.0	5.2	9.8
Minimum	6	4	8	7	2	5
Arithmetic average	66.3	13.2	74.6	22.9	10.2	62.9
Geometric average	35.2	11.4	43.5	17.3	8.3	20.0
Mode	–	10	–	17	12	10
Variance	7,792.2	42.4	10,988.8	872.1	39.7	19,410.1
Standard deviation	88.3	6.5	104.8	29.5	6.3	139.3
Coefficient of variation (CV)	1.3	0.5	1.4	1.3	0.6	2.2
Inferential statistics						
Standard error of the mean ($SE_{\bar{x}}$)	22.1	1.6	26.2	7.4	1.6	34.8
Relative standard deviation (RSD)	133.1%	49.4%	140.5%	129.1%	61.9%	221.6%
Margin of error (MoE)	47.0	3.5	55.9	15.7	3.4	74.2

Sample/descriptor	Arsenic	Chromium	Copper	Lead	Nickel	Zinc
Maximum probability error (MPE)	70.9%	26.3%	74.9%	68.8%	33.0%	118.1%
95% LCL$\bar{x}$ two-sided Student's t	19.3	9.7	18.7	7.1	6.8	-11.4
95% UCL$\bar{x}$ two-sided Student's t	113.4	16.7	130.5	38.6	13.5	137.1
95% LCL$\bar{x}$ one-sided Student's t	105.0	16.0	120.5	35.8	13.0	123.9
ProUCL determination	120.5	16.0	135.2	55.1	13.0	214.7
Method recommended	Gamma	Student's t	H-UCL	Chebyshev	Student's t	Chebyshev
Criteria and number of samples						
HIL-A land use (NEPC 2013, B1)	100	100	6,000	300	400	7,400
Number of samples CRV method	43.9	1.4	1.4	1.4	1.4	1.4
Number of samples CRV method	44	2	2	2	2	2
Number of samples MPE method	15	18	16	16	14	15

Notes

LORs = limits of reporting

For determination of descriptive statistics, see Appendices A to D of Part 2 of these guidelines (*Sampling Design Part 2 – Interpretation*).

SE$\bar{x}$ – see Appendix I of Part 2 of these guidelines.

RSD – see Appendix A of Part 2 of these guidelines.

MoE – see Appendix I of Part 2 of these guidelines.

MPE – see Appendix I of Part 2 of these guidelines.

For determination of lower confidence limit (LCL) and upper confidence limit (UCL), see Section 5 of Part 2 of these guidelines (*Sampling Design Part 2 – Interpretation*).

ProUCL = USEPA's ProUCL, Version 5.1

For determination of number of samples, see Appendices E and F of this document (i.e. Part 1 of the guidelines).

Appendix C: Determining sampling grids for hotspot detection

This appendix provides the methods for determining the required grid size, for square grids, to detect hotspots of a specified size. The method for determining the approximate number of sampling locations, based on the hotspot shape and size, is also provided. However, as the number of sampling locations required is in part based on the geometry of the site or decision area, the actual number of sampling locations required is dependent on applying the specified grid size to the actual site or decision area.

Determination

For determination of grid size:

Equation 1

$$G = \frac{r}{k}$$

For determination of the number of sampling locations:

Equation 2

$$n = \frac{A}{G^2}$$

For determination of the critical size of hotspots:

Equation 3

$$r = k \times G$$

Where:

- **G** is the grid size, i.e. the distance between nodes of grid
- **r** is the radius of a circular hotspot (for intermediate-shaped and elliptical hotspots, halve the length of the major axis)
- **k** is a statistical constant, dependant on the shape of the hotspot and the required confidence level
- **n** is the number of sampling locations
- **A** is the area of the site or decision area.

The values for k at 95% confidence level were determined from Figure 10.3 in Gilbert (1987), as:

- 0.59 for circular hotspots (ratio is 1:1)
- 0.69 for intermediate shaped hotspots (ratio is 4:3)
- 0.9 for elliptical hotspots (ratio is 2:1).

Gilbert (1987) notes that for elliptical targets, the curves in Figure 10.3 are “average curves over all possible orientations of the target relative to the grid”.

To determine the required grid size, choose the expected size (r) and shape (k) of the hotspot, then determine the required grid size from Equation 1. Use Equation 2 to determine n, the number of sampling locations required, and Equation 3 to determine r, the minimum hotspot size that can be detected.

If the contaminant is known or suspected to exhibit periodic spatial variations, the sampling pattern should be oriented such that it will not be in or out of phase with the known or suspected periodic spatial variations.

References

Gilbert RO 1987, *Statistical Methods for Environmental Pollution Monitoring*, John Wiley & Sons Inc., Brisbane.

Appendix D:
Summary of existing guidance for sample design

Table 8 summarises sampling design information from guidance made and approved under Section 105 of the *Contaminated Land Management Act 1997*, and from other guidance documents. The specified guidance should be referred to for details of the sampling strategies, locations, sampling densities, and potential contaminants of concern (PCoCs).

Table 8 Existing guidance for sampling design

Situation or land use	Guidance	Medium	Sampling design information
All land uses	National Environment Protection Council (NEPC) 2013, *National Environment Protection (Assessment of Site Contamination) Amendment Measure 2013 (No. 1)*, Schedule B2, National Environment Protection Council, Canberra.	Soil, soil gas and groundwater	Provides judgmental and probabilistic sampling design information for various media, including stockpiles.
Banana lands	Environment Protection Authority (EPA) 1997, *Contaminated sites: guidelines for assessing banana plantation sites*, EPA 97/37, NSW EPA, Sydney.	Soil	Provides information to investigate and assess potential contamination on current and former banana growing lands, including PCoCs. A systematic/grid-based sampling strategy is recommended, with variable sampling densities based on the former use, the current land use, and the stage of the investigation. A judgmental sampling design is recommended during validation of any excavations.
Cattle-tick dip sites	McDougall KW & Macoun TW 1996, *Guidelines for the Assessment and Clean Up of Cattle Tick Dip Sites for Residential Purposes*, NSW Agricultural in conjunction with CMPS&F Environmental, Wollongbar NSW.	Soil	Provides information to assess and remediate sites containing former cattle tick dips, including an overview of the PCoCs and the areas of highest potential contamination. A stratified systematic sampling design is recommended, with sampling of the sub-area based on the likelihood for contamination from past use.

Situation or land use	Guidance	Medium	Sampling design information
Excavated natural material (ENM) order	Environment Protection Authority (EPA) 2014, Resource Recovery Order under Part 9, Clause 93 of the Protection of the Environment Operations (Waste) Regulation 2014, *The excavated natural material order 2014*, NSW EPA, Sydney.	Soil	Provides information to allow the adequate assessment and classification of ENM for resource recovery. The order stipulates sampling strategies based on in-situ or stockpiled material. For in-situ a systematic/grid-based sampling strategy at specified depth intervals is required. When stockpiled, the number of samples by volume is specified, noting judgement must be used to ensure that samples taken are representative of the material.
Ground gas	Environment Protection Authority (EPA) 2020, *Assessment and Management of Hazardous Ground Gases: Contaminated Land Guidelines*, EPA 2019P2047, NSW EPA, Sydney.	Soil gas	Provides information to assist with the investigation of sites with potential hazardous bulk and trace ground gases. Includes judgmental sampling design information to apply for site specific scenarios.
Groundwater	Department of Environment and Conservation (DEC) 2007, *Contaminated sites: Guidelines for the Assessment and Management of Groundwater Contamination*, DEC 2007/144, DEC NSW, Sydney.	Groundwater	Provides information to conduct groundwater investigations, including a description of relevant concepts to allow for an adequate sampling design program to be developed.
Gasworks	Department of Environment and Conservation (DEC) 2005a, *Information for the Assessment of Former Gasworks Sites*, DEC 2005/237, DEC NSW, Sydney.	Soil and groundwater	Provides information relating to former gasworks sites and the potential for contamination of site areas and PCoCs. Recommends stratifying the site and using a systematic sampling design to ensure sufficient sampling density for each area of concern. Describes that groundwater monitoring well locations should consider the site-specific complexities of the hydrogeology.
Land farming	Environment Protection Authority (EPA) 2014b, *Best Practice Note: Landfarming*, EPA 2014/0323, NSW EPA, Sydney.	Soil	Provides information on best practice land farming techniques and recommends a systematic sampling design that is adequate to provide a statistically reliable result.

Situation or land use	Guidance	Medium	Sampling design information
Orchards and market gardens	Department of Environment and Conservation (DEC) 2005b, *Contaminated Sites: Guidelines for Assessing Former Orchards and Market Gardens*, DEC 2005/195, DEC NSW, Sydney.	Soil	Provides information relating to former orchards and market gardens sites, the potential for contamination of specific site areas and relevant PCoCs. Recommends a systematic grid-based sampling plan across the cultivated areas of the site, targeting the surface soils, with a higher sampling density for areas where localised contamination is likely to have occurred i.e. chemical storage sheds and tractor turning circles.
Resource recovery order/exemptions	Environment Protection Authority (EPA) 2018, *Guidelines on Resource Recovery Orders and Exemptions: For the Land Application of Waste Materials as Fill*, EPA 2017/P0392, NSW EPA, Sydney.	Soil and fill	Provides information for resource recovery order/exemptions to allow beneficial reuse of waste products, e.g. fuel, fill, fertiliser, etc. The guideline provides a minimum number of samples which must be collected and specifies that the "sampling plan must have a clear, defensible rationale", implying the use of probabilistic systematic sampling designs.
Service stations and underground petroleum storage systems (UPSS)	EPA (2014c) *Technical Note: Investigation of Service Station Sites*, EPA 2014/0315, NSW EPA, Sydney **and** Department of Environment, Climate Change and Water (DECCW) 2009, *Guidelines for Implementing the Protection of the Environment Operations (Underground Petroleum Storage Systems) Regulation 2008*, DECCW 2009/653, DECCW NSW, Sydney.	Soil and groundwater	Provides information to investigate and assess contamination at service stations or locations with UPSS. Judgmental sampling design is recommended, targeting soil and groundwater in areas of infrastructure and known contamination.
Stockpiles	National Environment Protection Council (NEPC) 2013, *National Environment Protection (Assessment of Site Contamination) Amendment Measure 2013 (No. 1)*, Schedule B2, National Environment Protection Council, Canberra.	Soil and fill	Section 7.5 of the NEPM (2013, B2) provides information for assessing stockpiles of homogenous soil or fill of ≤ 200 m^3. Recommends a minimum number of samples to undertake an initial assessment of a stockpile, with either a judgmental or probabilistic sample design recommended, based on the specific circumstance.

Situation or land use	Guidance	Medium	Sampling design information
Surface water	Australian and New Zealand Environment and Conservation Council (ANZECC) and Agriculture and Resource Management Council of Australia and New Zealand (ARMCANZ) 2000, *Australian and New Zealand Guidelines for Fresh and Marine Water Quality*, paper no. 4, ANZECC and ARMCANZ, Canberra. Available at: www.waterquality.gov.au/anz-guidelines.	Surface waters	Provides detailed guidance for the management and assessment of waters in Australia and New Zealand. Information is provided on how to develop an appropriate surface water sampling program.
Vapour intrusion	Environment Protection Authority (EPA) 2020, *Assessment and Management of Hazardous Ground Gases: Contaminated Land Guidelines*, EPA 2019P2047, NSW EPA, Sydney **and** Department of Environment, Climate Change and Water (DECCW) 2010, *Vapour Intrusion: Technical Practice Note*, DECCW 2010/774, DECCW NSW, Sydney.	Soil gas and volatiles	Provides information for assessment of sites that have potential vapour intrusion issues. Judgmental and probabilistic sampling design information for the various vapour intrusion investigation methods, including conceptual information to determine the number of sample locations and frequency.
Vertical mixing of soil on former broadacre agricultural land	Environment Protection Authority (EPA) 1995, *Contaminated Sites: Guidelines for the Vertical Mixing of Soil on Former Broad-Acre Agricultural Land*, EPA 2003/28	Soil	Provides information regarding use of vertical mixing techniques of former agricultural land. No specific sampling design information is provided beyond sample depths.

Appendix E: Determining the number of samples by the CRV method

The number of samples needed to show that the mean concentration of a contaminant is below a defined action level or criteria can be determined using the combined risk value (CRV) method. This method can be used for sites or decision areas for all media, where probabilistic sampling has been undertaken.

The determination derived from the Student's t-test formula for hypothesis testing, with the alpha (α) value for a Type I error, or false rejection of the hypothesis, and the beta (β) value for a Type II error, or false acceptance of the hypothesis, is used to determine the CRV. In the assessment of site contamination, the null hypothesis (H_0) is always that the contaminant concentrations exceed the action levels or criteria, and where H_0 is not rejected, there is only a potential for a Type II or false acceptance error rate, and this sample size formula can be used to determine if the error rate has been satisfied.

This method can be used to design a sampling program, either using previous data or estimates to determine $\bar{x}$ and s, or retrospectively to demonstrate sufficient statistical power or otherwise. Where the determination results in low values of n, including ≤ 1, this suggests that the minimum detectable difference Δ (uppercase Greek letter delta) is overly large, and additional statistical analysis is required to determine or justify the number of samples.

Determination

1. The number of samples using the CRV method is determined by:

$$n = \frac{\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2 * s^2}{(C_s - \bar{x})^2}$$

Where

- **n** number of samples
- **Z** standard normal distribution (z curve)
- **Z_{1-a}** Z value for α
- **$Z_{1-\beta}$** Z value for β
- **C_s** criterion/action level
- **$\bar{x}$** sample mean
- **s** sample standard deviation.

2. The risk values are selected for $Z_{1-\beta}$ and $Z_{1-\alpha}$ from USEPA 1989.

When comparing to action levels or criteria, the recommended values are 0.05 α risk and 0.2 β risk, corresponding to confidence levels of 95% and 80% respectively. Using a 0.05 α risk value of 1.645 and a 0.2 β risk value of 0.842, the CRV is 6.2.

Where increased certainty is required, such as determining if costly remedial works are necessary, consultants are encouraged to examine the use of more conservative β values, which will result in an increased CRV. As β corresponds to the risk of falsely accepting the H_0, that is, that the site is contaminated, further sampling reduces the chance of Type II errors, which could potentially lead to the rejection of H_0, making the decision that the site is not contaminated. Generally, the cost of unnecessary remediation will far outweigh the cost of the additional sampling and analysis.

Worked example

The metals data in mg/kg from Table 7 in Appendix B is used in this example to confirm that the number of samples collected for arsenic (As) and chromium (Cr) is appropriate, i.e. that the statistical power of the test is sufficient.

1. Select the confidence level and the power of the test. For α = 0.05 and β = 0.2, the solution is:

$$n = \frac{6.2 * s^2}{(C_s - \bar{x})^2} + 1.4$$

2. The number of samples required using the CRV method is determined at α = 0.05 and β = 0.2 as follows.

Arsenic

3. For As, $\bar{x}$ = 66.3, s = 88.3 and HIL-A = 100 mg/kg:

$$n = \frac{6.2 * 88.3^2}{(100 - 66.3)^2} + 1.4$$

$$n = 43.9$$

Rounding to the nearest whole number, 44 samples are required to characterise the site or decision area for As, based on the large standard deviation. As the maximum concentration of As exceeds HIL-A by more than 250%, additional investigation is required to further characterise the distribution of As. The large number of samples required, and the large value of s, suggests that further characterisation should seek to segregate the decision area into different sub-populations, either in plan or by depth, for the design of further investigations and consideration of remedial options.

Chromium

3. For Cr, $\bar{x}$ = 13.2, s = 6.5 and HIL-A = 100 mg/kg (Cr^{6+}):

$$n = \frac{6.2 * s^2}{(C_s - \bar{x})^2} + 1.4$$

$$n = \frac{6.2 * 6.5^2}{(100 - 13.2)^2} + 1.4$$

$$n = 1.4$$

Rounding to the nearest whole number, two samples are required to characterise the site or decision area. This is not surprising, based on the small standard deviation and mean. However, with so few samples, it is not possible to estimate the true values of the critical parameters of the contaminant distribution, such as $\bar{x}$ and s. Rather, the use of the CRV method and the derivation of n less than the

number of samples collected, suggests that the false rejection (α) error rate has been satisfied, and that in the case of Cr, it is reasonable to reject H_0 (i.e. that the site or decision area is contaminated with Cr).

References

US Environmental Protection Agency (USEPA) 1989, *Methods for Evaluating the Attainment of Cleanup Standards, Volume 1: Soils and Solid Media*, EPA 230/02-89-042, USEPA, Washington DC.

US Environmental Protection Agency (USEPA) 2006, *Guidance on Systematic Planning Using the Data Quality Objectives Process (QA/G-4)*, EPA/240/B-06/001, Appendix: Derivation of Sample Size Formula for Testing Mean of Normal Distribution Versus an Action Level, USEPA, Washington DC.

Appendix F: Determining the number of samples by the MPE method

The number of samples needed to show that the average concentration of a contaminant is within a specific range, such as a confidence interval, can be determined using the maximum probable error (MPE) method. This can be thought of as a specified statistical precision around a point estimate and can be used for any medium and any probabilistic sampling design.

This method can be used when addressing estimation problems as defined within the data quality objectives (DQOs) process, as it allows a desired precision to be specified outside a strict hypothesis-testing framework.

This method uses the margin of error (MoE), the standard deviation (s), and a critical value at a specified confidence level. The MoE can be thought of as the 'radius' to, or half the width of, the diameter of the confidence interval. Initially, $Z_{1-\alpha/2}$ is used for a first determination, the result of which defines the degrees of freedom for selection of a value for $t_{1-\alpha/2,n-1}$. Subsequent iterations are conducted until the number of samples calculated stabilises.

In the case of the MPE method, as the equation approaches n = n, it cannot be used to retrospectively demonstrate sufficient sampling specifically, but rather it provides a guide to an appropriate number of samples based on the variability of the data (standard deviation), and the required precision of the data MoE. If the data shows too large an MPE (> 35–50%) for reasonable relative standard deviations (RSDs) (65–150%), then the data is probably not sufficiently precise.

Table 9 shows various values for n, calculated using USEPA (2015): they are illustrated in Figure 10. As the RSD increases, and higher precision (lower MPE) is required, the number of samples required increases.

Determination

1. The number of samples is calculated by the MPE method as:

$$n = {Z_{1-\alpha/2}}^2 * \frac{s^2}{MoE^2}$$

Where:

n number of samples

$Z_{1-\alpha/2}$ Z from the standard normal distribution

$t_{1-\alpha/2,n-1}$ critical value

s sample standard deviation

MoE margin of error (= $t_{1-\alpha}$ * $SE_{\bar{x}}$)

$SE_{\bar{x}}$ standard error of the mean (= s/√n).

2. The MoE and s can be standardised as relative values by dividing by $\bar{x}$, giving the maximum probable error (= MoE/$\bar{x}$) and the relative standard deviation (RSD) (= s/$\bar{x}$), which is also known as the coefficient of variation (CV). Using the standardised MPE method, the required number of samples is calculated by:

$$n = {t_{95\%}}^2 * \frac{RSD^2}{MPE^2}$$

Either method can be used as long as the variables are consistent, i.e. s and MoE are expressed in mg/kg, or RSD and MPE are given as percentages.

Worked example

The metals data in mg/kg from Table 7 is used in this example to determine if sufficient samples have been collected for copper (Cu), both at the surface and at depth.

1. The number of samples required is initially determined using Z, as:

$$n = {Z_{1-\alpha/2}}^2 * \frac{s^2}{MoE^2}$$

2. For the surface fill, at a 95% confidence level, Z = 1.96, s = 137 and MoE = 114.5.

$$n_1 = 1.96^2 * \frac{137^2}{114.5^2}$$

$$n_1 = 5.5$$

3. Rounding to the next whole number, 6, the degrees of freedom is 5. Using $t_{1-\alpha/2,n-1}$ = 2.571, the next determination is:

$$n_2 = {t_{1-\alpha/2}}^2 * \frac{s^2}{MoE^2}$$

$$n_2 = 2.571^2 * \frac{137^2}{114.5^2}$$

$$n_2 = 9.5$$

4. This process is continued until at n_4, n stabilises at 8, which was the number of samples collected.

5. The same process is used for the depth fill, which also stabilises at n = 8, which was the number of samples collected.

As discussed, this approach cannot be used to confirm retrospectively if an appropriate number of samples was collected. However, by examining the RSD and the MoE achieved by the number of samples that were collected, you can determine if a sufficient number of samples was collected to meet the project requirements in regard to the desired quantity and quality of the data.

In the present example, we can compare the surface fill and depth fill.

The RSD is 115% for the surface fill and 52.8% for the depth fill, and the MPE is 96.1% for the surface fill and 44.2% for the depth fill. Table 9 shows that these values give a sample number of 8 (by interpolation).

For the surface fill with an RSD of 115%, to achieve an MPE of 50%, 23 samples would be required (by interpolation).

In the case of the depth fill, based on the homogenous nature of the material, as indicated by the low RSD (~50%), and the precision of the data (MPE of ~45%), it is likely that the dataset would be suitable for a decision.

For the surface fill, Cu is well below the HIL-A of 6,000 mg/kg, and so it would probably be considered that sufficient samples have been collected from the surface fill to make a decision. However, where the dataset exhibits high RSDs and approaches the criteria or action levels, the MPE method provides a tool for assessing the quantity and quality of the data for making decisions.

References

Gilbert RO 1987, *Statistical Methods for Environmental Pollution Monitoring*, John Wiley & Sons Inc., Brisbane.

Provost LP 1984, Statistical Methods in Environmental Sampling, in GE Schweitzer and JA Santolucito (eds), *Environmental Sampling for Hazardous Wastes*, American Chemical Society, Washington DC.

US Environmental Protection Agency (USEPA) 2015a, *ProUCL Version 5.1.002: Technical Guide: Statistical Software for Environmental Applications for Data Sets with and without Nondetect Observations*, EPA/600/R-07/041, USEPA, Washington DC.

Table 9 **Number of samples (n) required to estimate mean, based on the MPE method**

RSD %	Maximum probable error %											
–	10	15	20	25	30	35	40	45	50	55	75	100
10	6	4	3	3	3	3	3	3	3	3	2	2
15	11	6	5	4	3	3	3	3	3	3	3	3
20	18	9	6	5	4	4	3	3	3	3	3	3
25	26	13	8	6	5	4	4	4	3	3	3	3
30	37	18	11	8	6	5	4	4	4	4	3	3
35	49	23	14	10	8	6	5	5	4	4	3	3
40	64	30	18	12	9	7	6	5	5	4	4	3
45	80	37	22	15	11	9	7	6	6	5	4	3
50	98	45	26	18	13	10	8	7	6	6	4	3
55	119	54	31	21	15	12	10	8	7	6	4	4
60	141	64	37	25	18	14	11	9	8	7	5	4
70	191	86	49	33	23	18	14	12	10	9	6	4
80	248	112	64	42	30	22	18	15	12	11	7	5
90	314	141	80	52	37	28	22	18	15	13	8	6
100	387	173	98	64	45	34	26	21	18	15	9	6
110	467	209	119	77	54	40	31	25	21	18	11	7
120	556	248	141	91	64	48	37	30	25	21	12	8
130	652	291	165	106	75	55	43	34	28	24	14	9
140	755	337	191	123	86	64	49	40	33	27	16	10
150	867	387	219	141	98	73	56	45	37	31	18	11
175	1,179	525	297	191	133	98	76	61	49	41	23	14
200	1,539	685	387	248	173	128	98	78	64	53	30	18

RSD = relative standard deviation ($s/\bar{x}$, where s is the sample standard deviation and $\bar{x}$ is the sample arithmetic mean).

MPE = maximum probable error ($MoE/\bar{x}$, where MoE is the margin of error ($= t_{95\%} * s/\sqrt{n}$)).

Shaded values represent the general range of n required for characterising homogenous material within the same decision area.

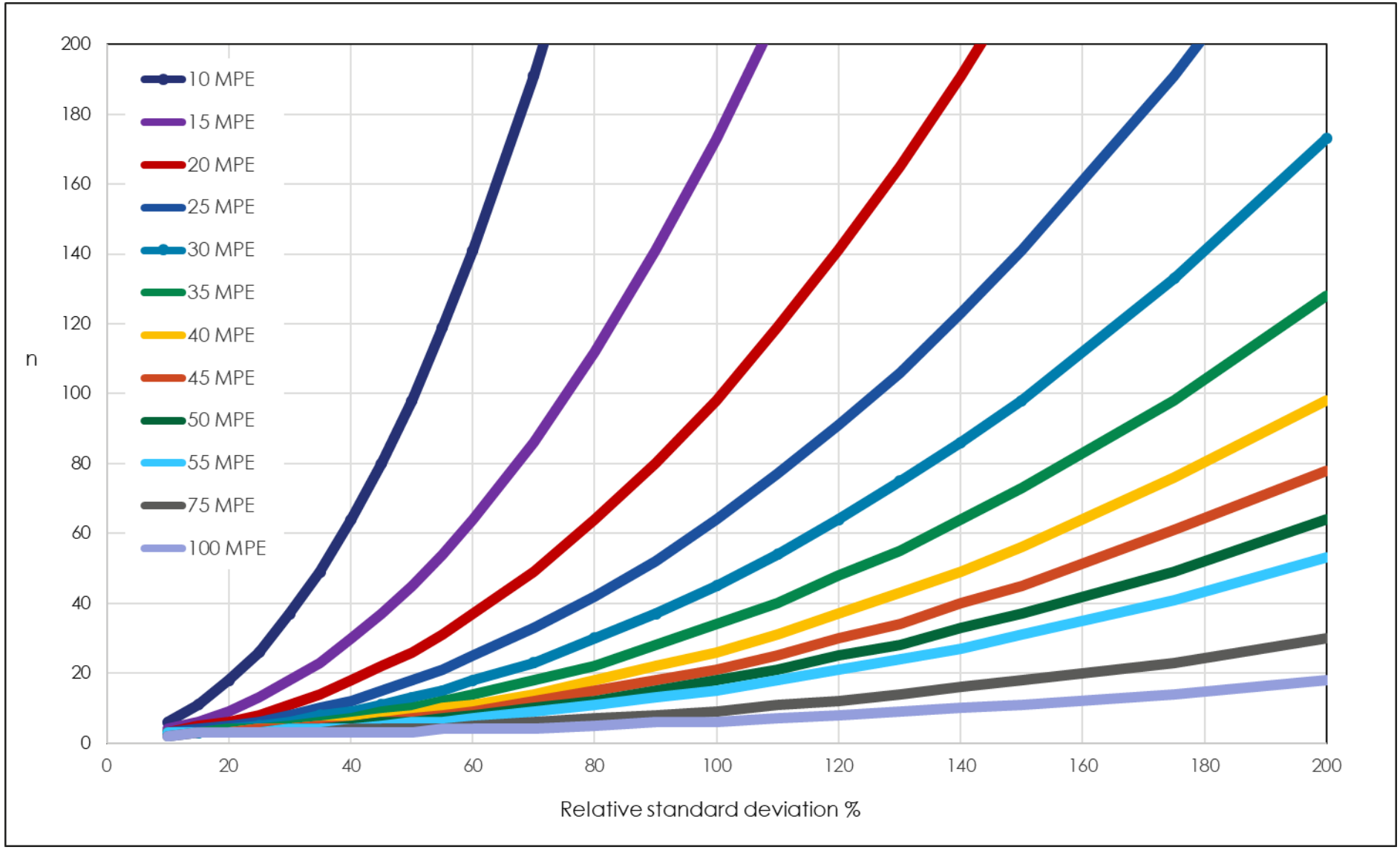

Figure 10 Number of samples (n) required to estimate mean, based on the MPE method

Source: Marc Salmon/Easterly Point Environmental Pty Ltd

Appendix G:
Further methods for consideration

The following section provides a summary of some of the investigation methods commonly used by USEPA, and also introduces geospatial statistics.

Incremental sampling methods

ITRC (2012) describes an **incremental sampling method** (ISM) as "a structured composite sampling and processing protocol that reduces data variability and provides a reasonably unbiased estimate of mean contaminant concentrations in a volume of soil". A successful ISM program relies on a systematic sampling approach that integrates a thorough CSM and DQOs, to identify the decision units. A site may be broken up into a number of decision units, which MDEQ (2015) notes can be based on:

- areas that establish exposure areas
- contaminant transport and exposure pathways
- site contaminant distributions
- geological and other physical characteristics.

Between 30 and 100 systematic, random increments (samples) are collected from each decision unit. These are then combined, using a specific method, with a representative sample collected for laboratory analysis. Figure 11 provides an example of a systematic sampling program, showing the density of sample coverage.

Due to the sheer volume of samples or increments collected, dedicated field and laboratory analysis methods must be specifically designed for an ISM program; ISM is therefore not used in Australia. However, ISM is widely used in North America, and is an accepted sampling method of USEPA, as it addresses a number of limitations that are prevalent in traditional discrete sampling methods. This section provides only a scraping of the surface of ISM. If you need further information about this method, see ITRC (2012) and MDEQ (2015).

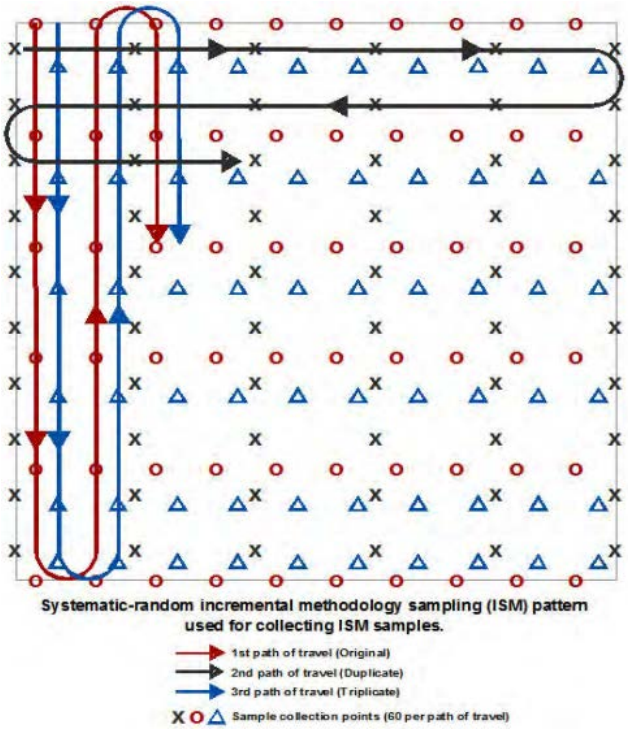

Figure 11 Example of an ISM sample design

Source: Adapted from ITRC (2012).

Triad approach

The Triad approach is a decision-making framework developed by the USEPA, to manage decision uncertainties in environmental data. It draws on new advances in science and technology and allows projects to proceed rapidly and data to be collected cost-effectively (Crumbling et al. 2004).

Traditionally, site characterisation has proceeded through many stages of investigation, with an emphasis on the assessment of the analytical data based on the relationship between the data quality and the analytical quality. However, this approach can sometimes be quite static, repetitious, time-consuming and expensive. In contrast, the Triad approach focuses on a more adaptive approach to site characterisation, using real-time decision-making tools to guide the field activities, creating more flexibility and reducing overall costs and resources.

Preliminary evaluations suggest that incorporating the Triad approach into the decision-making framework can save up to 50% of the cost of more traditional approaches to site characterisation (Crumbling 2001).

The Triad approach has three main elements. While these elements are not new concepts in the site-investigation process, what **is** new is how they are synthesised to "plan, implement and improve data collection from contaminated sites" (Clements et al. 2009). The three elements are:

Systematic planning

This is the most important element of the Triad approach and ensures high decision confidence. Systematic planning includes the application of the data quality objectives (DQOs) process and development of a conceptual site model (CSM). These planning tools can then be used to inform stakeholders by providing a clear understanding of the site, the uncertainties identified, and the required data objectives. This stage of the process also allows for stakeholder involvement and consensus regarding the desired project outcomes, including any end goals and/or exit strategies, which are clearly defined prior to the commencement of field work. This contrasts with the traditional approach, where a decision is made after the site investigation has been conducted, based solely on the results of analytical data (Crumbling et al. 2004).

Dynamic work strategies

This is the element of the framework that allows for projects to be completed much faster and at considerably less cost than can be done using the more traditional, static work strategies. Work planning documents are prepared in such a way that they allow for flexibility in the project planning as data from field measurements becomes available. For example, a sampling, analysis and quality plan (SAQP), may include contingencies that allow field activities to be modified, even while field work is still occurring (Clements et al. 2009). This allows the CSM to be a 'dynamic' document that can be refined as more site information and data become available.

Real-time measurement systems

By reviewing field screening data and analytical data in real-time, decisions such as remediation strategies and adaptive sampling plans (e.g. revised sampling locations, sample quantities, and/or analytical strategy) can be made while the fieldwork team is still onsite. This element of the Triad also allows for the data to be shared among all stakeholders as soon as it is generated, creating transparency, which helps to establish trust and good working relationships with regulators and stakeholders while also informing the decision-making process (Crumbling et al. 2004).

Traditionally site investigations have focused on the quality of the analytical data. i.e., the analytical data is considered to be 'definitive data' and of a 'high quality', while real-time data and field screening methods are considered to generate 'screening data', i.e., 'inferior' quality data (Crumbling 2001). In fact, the quality assurance conducted as part of the Triad approach has the potential to be "more relevant and supportive of defensible project decisions than [that done] under traditional scenarios" (Crumbling et al. 2004). The Triad process also improves project quality by recognising the potential impacts of uncertainties in site heterogeneity, which are often overlooked in traditional site assessments and project planning.

Figure 12 shows examples of real-time measurement technologies.

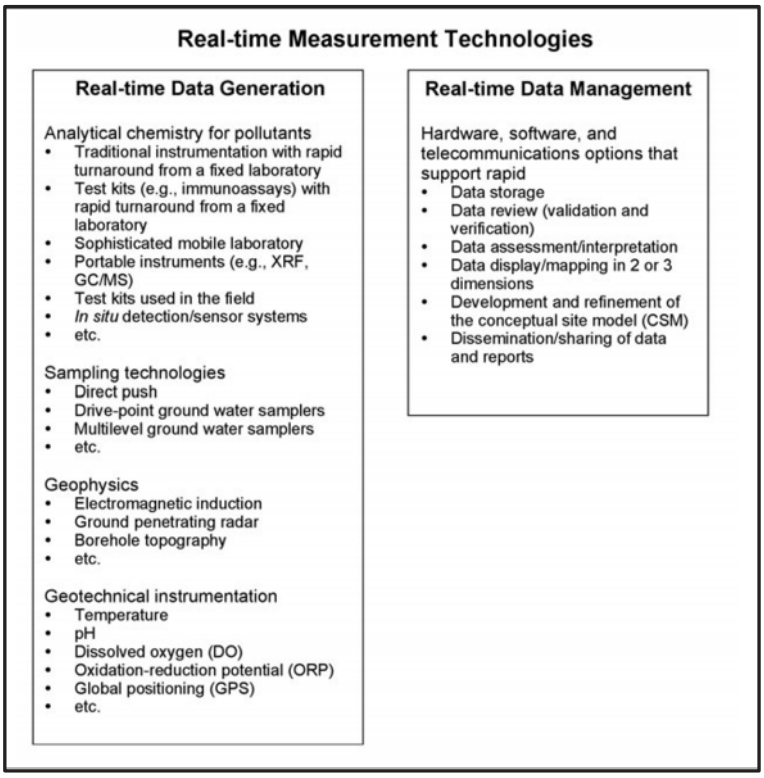

Figure 12 Real-time measurement technology

Source: Crumbling et al. 2004

Summary

Staged investigation processes came into use at a time when technology, science and consultant experience in relation to contaminated sites were limited. Now new technologies and science have emerged and consultants better understand contamination scenarios, environmental fate and transport processes, so it may be time to revise traditional approaches to site characterisation. The Triad approach increases decision confidence; provides an adaptive approach that focuses on real-time decision making to guide field activities and the development of a dynamic CSM; and can be a more cost- and resource-effective alternative to the traditional multi-stage investigation process. However, for the Triad approach to be effective, all of its key concepts must be used.

Geospatial statistics

Geostatistical data analysis is based on **bivariate statistics** theory. Bivariate statistics allow for the assessment of the relationships between two variables. The bivariate statistical approach for assessment of concentration data was developed in the 1960s for use in the minerals exploration industry (Krige 1981). This approach examines concentrations and variations, in relation to their spatial distribution. Given the similar objectives for the assessment of contamination at a site, this method has application in the assessment of contaminant concentration and location data. Bivariate geostatistics assess not only the distribution of the contaminant but also the spatial variance in concentration (Goovaerts 1997; Webster & Oliver 2001; Nielsen & Wendroth 2003).

Variograms

The **variogram** is the basic tool of geostatistics (Royle 1980) and expresses the spatial correlation between adjoining samples. Variograms are essentially scatter plots of distance between sample locations and variance of sample values to establish whether there is a predictable change in variance with distance.

A variogram is constructed by calculating the mean squared difference (variance) between sample values over incremental sample spacing. Methodologies for calculation of the mean squared difference and construction of variograms are outlined by a number of authors (Henley 1981, Krige 1981, Rendu 1981 and Royle 1980) and are therefore not outlined in detail here.

Variograms present a number of key geostatistical properties, which are shown in Figure 13. The range of influence is the distance over which the samples values are related. The total variance of the samples can be split into a random and a spatial component. Random variance is also known as the **nugget effect**.

Development of variograms generally requires specialist software or spatial data assessment software.

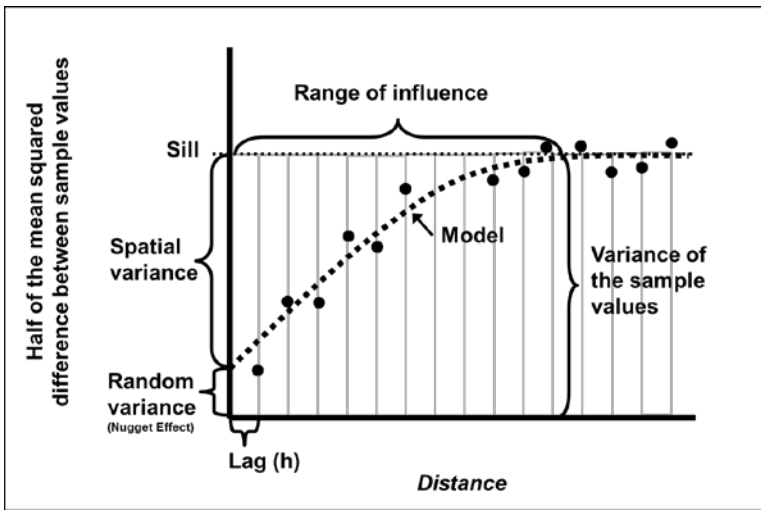

Figure 13 Features of a variogram

Source: Beck, Mikov and Curtis (2004)

Spatial interpolation methods

The data from the variogram is interpreted using the **kriging** method (Krige 1981; Randu 1981), which is essentially a weighted linear estimation technique (Royle 1980). Kriging provides an estimate of a value at a given location where no site-specific measurement has been made (Henley 1981). It can be used to predict the value (such as a concentration) at a location for which no data exists, or to predict the confidence in the interpretation by using **indicator kriging** (Krige 1981; Randu 1981; Isaaks & Srivastava 1989).

Kriging is generally regarded as the most reliable interpolation method for predicting values away from locations that were sampled, due to the absence of bias commonly associated with other interpolation methods.

Benefits of spatial geostatistics

Spatial geostatistics have a number of advantages over the univariate approaches commonly used, including:

- not being reliant on sample collection being unbiased, therefore allowing for the use of a single statistical method for all sampling data
- assessment of the spatial and random contribution to concentration variation
- providing a method to establish when a site is adequately characterised
- providing the most reliable interpolation method for spatial concentration data
- providing a reliable method for probabilistic mapping of occurrence of contamination
- providing a reliable method for probabilistic evaluation of volumes of contaminated media.

Spatial geostatistics can be considered for sites where more than 10 sampling locations have been completed, as the method requires a reasonable number of samples to be applied effectively.

Example application

The dataset used in this example was generated by a staged investigation at a large parkland and sporting ovals. The site was suspected to have been filled with soil derived from a gasworks. Data from the first stage of the investigation was used to develop a variogram for the polycyclic aromatic hydrocarbon (PAH) data to assess whether there were signs of a spatial relationship. Figure 14 shows the variogram derived from the first 26 samples analysed.

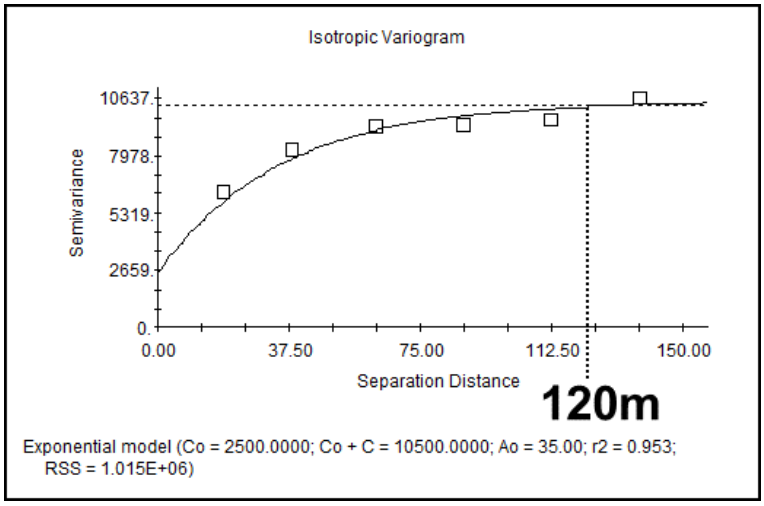

Figure 14 Variogram for PAHs from the first stage of the investigation

Source: Beck, Mikov and Curtis (2004)

The variogram derived from the initial samples shows that around 70% of the data variance is spatial and shows a range of around 120 m, while around 30% is random. The initial variogram was used to inform the design of the second sampling round by using indicator kriging to identify areas of low confidence in the spatial distribution of contamination. The second sampling round identified 10 sampling locations that would assist in improving the confidence in spatial interpretation. The data generated by the initial and second sampling round was the used to develop a second variogram using log transformed, which is shown below.

The second variogram (Figure 15) showed a notable decrease in the random variance to around 10% of the total variance, while the range remained relatively similar.

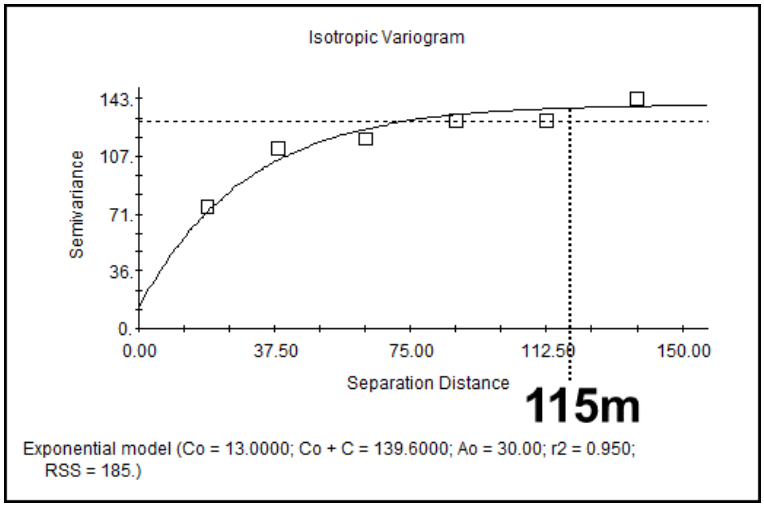

Figure 15 Second variogram for PAHs

Source: Beck, Mikov and Curtis (2004)

The second variogram was used to design the third sampling round, in which samples were collected from a further 12 locations. The variogram developed from the dataset after the third sampling round is shown in Figure 16.

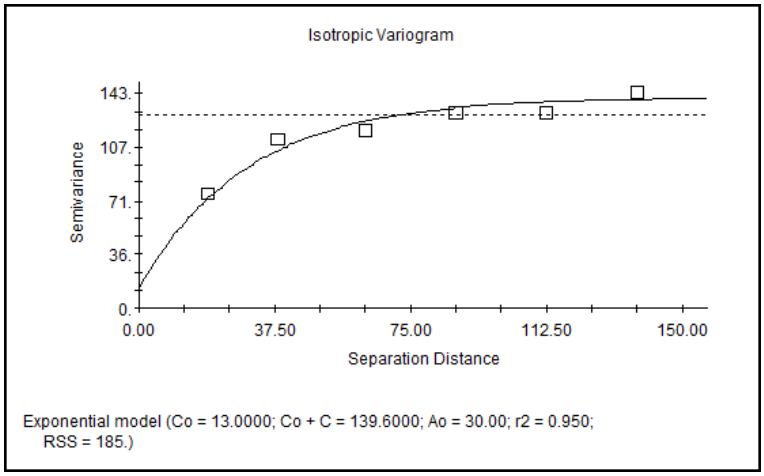

Figure 16 Third variogram for PAHs

Source: Beck, Mikov and Curtis (2004)

The third variogram was almost identical to the second, suggesting that further sampling would not improve characterisation of the site. However, the indicator kriging showed **that over 85% of the site was covered, at a confidence level of 80% or higher.**

References

Beck PH, Mikov A and Curtis M 2004, Using spatial geostatistics to evaluate the risk contamination occurring between sampling locations. Presented at *Contaminated Site Remediation Conference*, 16–18 September 2004, Adelaide.

Clements L, Palaia T & Davis J 2009, *Characterisation of Sites Impacted by Petroleum Hydrocarbons: National Guideline Document*, CRC Care Technical Report no. 11, CRC for Contamination Assessment and Remediation of the Environment, Adelaide.

Crumbling DM 2004, The Triad Approach to Managing the Uncertainty in Environmental Data, in: *NEMC 2004: The 20th Annual National Environmental Monitoring Conference*. Available at: http://nemc.us/docs/2004/2004_nemc.pdf.

Crumbling DM 2001, *Applying the Concept of Effective Data to Environmental Analyses for Contaminated Sites* (EPA 542-R-01-013), U.S. Environmental Protection Agency, Washington DC. Available at: http://cluin.org/download/char/ effective_data.pdf.

Crumbling DM, Hayworth JS, Call BA, Davis WM, Howe R, Miller DS & Johnson R 2004, The Maturing of the Triad Approach: Avoiding Misconceptions, *Remediation: The Journal of Environmental Cleanup Costs, Technologies & Techniques*, 14(4), pp. 81–96.

Goovaerts P 1997, *Geostatistics for Natural Resources* Evaluation, Oxford University Press, New York, p. 496.

Henley S 1981, *Nonparametric Geostatistics*, Elsevier Applied Scientific Publishers, Barking UK.

Interstate Technology and Regulatory Council (ITRC) 2003, *Technical and Regulatory Guidance for the Triad Approach: A New Paradigm for Environmental Project Management*, SCM-1, ITRC, Washington DC. Available at: www.itrcweb.org/SCM-1.pdf.

Interstate Technology and Regulatory Council (ITRC) 2012, *Technical and Regulatory Guidance, Incremental Sampling Methodology*, ISM-1, ITRC, Washington DC. Available at: www.itrcweb.org/ism-1/.

Isaaks EH & Srivastava RM 1989, *An Introduction to Applied Geostatistics,* Oxford University Press, New York.

Krige DG 1981, *Lognormal-de Wijsian Geostatistics for Ore Evaluation*, South African Institute of Mining and Metallurgy, Johannesburg, South Africa.

Michigan Department of Environmental Quality (MDEQ), 2015, *Incremental Sampling Methodology and Applications,* Remediation and Redevelopment Division Resource Materials, Michigan.

Nielsen D & Wendroth O 2003, *Spatial and Temporal Statistics: Sampling Field Soils and Their Vegetation*, Catena-Verlag, Reiskirchen, Germany, p. 614.

Rendu JM 1981, *An introduction to geostatistical methods of mineral evaluation*, South African Institute of Mining and Metallurgy, Johannesburg, South Africa.

Royle AG 1980, *Geostatistics*, McGraw-Hill, New York.

Webster R & Oliver MA 2001, *Geostatistics for Environmental Scientists*, Wiley, Chichester UK, p. 265.