Guideline 3 – Establishing baseline groundwater quality

Guidelines for groundwater quality monitoring of regulated activities series

ENVIRONMENT PROTECTION AUTHORITY

EPA.SA.GOV.AU

Guideline 3 – Establishing baseline groundwater quality
Guidelines for groundwater quality monitoring of regulated activities series
Author: Taryn Collaton, Principal Scientific Officer (Groundwater)

The Environment Protection Authority gratefully acknowledges Tamie Weaver (ERM) and Dana Windle (ERM) for their thorough review of and contribution to this guideline.

For further information please contact:

Information Officer
Environment Protection Authority
GPO Box 2607
Adelaide SA 5001

Telephone: (08) 8204 2004
Facsimile: (08) 8124 4670
Free call (country): 1800 623 445
Website: https://www.epa.sa.gov.au
Email: epainf@sa.gov.au

ISBN 978-1-925521-07-8

April 2024

EPA FOR RECONCILIATION

The EPA acknowledges and respects the Aboriginal peoples of South Australia as the first peoples and nations of this State. We recognise them as the traditional custodians of land and waters in South Australia and that their spiritual, social, cultural and economic beliefs are of ongoing importance today. We recognise that they have made, and continue to make, a unique and irreplaceable contribution to the State.

Artwork: 'Caring for Country', courtesy of Arrernte man Scott Rathman, for the EPA.

Contents

Abbreviations **1**

Executive Summary **3**

1 Introduction **5**

- 1.1 Description 5
- 1.2 Application 6
- 1.3 Purpose and scope 6

2 Understanding the groundwater system **9**

- 2.1 Groundwater flow and surface water interaction 9
- 2.2 Groundwater quality 10
- 2.3 Target population 11

3 Data collection **13**

- 3.1 Monitoring well network 13
- 3.2 Duration and frequency 15
- 3.3 Analysis 18

4 Understanding the data **20**

- 4.1 Checking data quality 20
- 4.2 Calculating summary statistics 30
- 4.3 Examining graphical representations of data 33
- 4.4 Recalculating summary statistics 41

5 Reporting baseline groundwater quality **42**

Appendix A Example of hydrogeological conceptualisations **47**

Appendix B Example of baseline groundwater quality summary **48**

Appendix C Reporting checklist **54**

List of figures

Figure 1 Key steps to establish baseline groundwater quality 8

Figure 2 Graph showing impact of changes in sampling methodolgy 21

Figure 3 Graph showing impact of sampling frequency for seasonal trend detection 22

Figure 4 Graph showing impact of correlated data on the calculated median concentration 23

Figure 5 Graph showing a change in reporting limit that will not imapact statistical analysis of the data set 24

Figure 6 Graph showing a change in reporting limit that may impact statistical analysis of the data set 25

Figure 7 Graph showing impact of different techniques used to manage non-detect data in a data set 27

Figure 8 Time series graph showing the impact of outliers on the distribution and statistical analysis of a data set 29

Figure 9 Histograms showing the difference in data distribution detail attributable to choice of bin size 35

Figure 10 Quantile–quantile plots showing the characteristic visual appearance of normally distributed data 35

Figure 11 Time series plot showing the impact of reporting limited data on the inference of a trend 37

Figure 12 Box plots showing the identification of spatially distinct populations of groundwater at a site 39

Figure 13 Piper plots showing spatial homogeneity, heterogeneity and temporal changes in groundwater 40

Figure 14 Example 3-D graphic conceptualisation of a groundwater system 47

Figure 15 Example cross–section conceptualisation of a groundwater system 47

Figure 16 Basic conceptual site layout and cross–section of landfill to assist understanding of data in Table 2. 53

List of tables

Table 1 Example of summary statistics for ammonia (as N) concentrations for three separate target populations 32

Table 2 Example baseline groundwater quality summary 48

Table 3 Checklist for reporting baseline groundwater quality 54

Abbreviations

CBE	charge balance error
EDA	exploratory data analysis
EPA	South Australian Environment Protection Authority
EP Act	*Environment Protection Act 1993*
GMMP	groundwater monitoring and management plan
IQR	interquartile range
ITRC	Interstate Technology and Regulatory Council
PCA	potentially contaminating activities
POC's	pollutants of concern
QA/QC	quality assurance/quality control
RL	reporting limit
ROS	Robust Regression on Order Statistics
TDS	total dissolved solids

Executive summary

This guideline is the first to be published by the South Australian Environment Protection Authority as part of the series titled ‘Guidelines for groundwater quality monitoring of regulated activities’. The series of guidelines is intended to provide guidance in relation to groundwater monitoring for regulated activities in South Australia.

The guideline series will comprise the following publications:

1. Groundwater monitoring bore network design.
2. Groundwater sampling.
3. Establishing baseline groundwater quality.
4. Establishing groundwater quality assessment criteria.
5. Developing a groundwater monitoring and management plan.
6. Groundwater quality assessment reporting.

This guideline is the third in the series, but the first to be published, as it is considered to offer the most immediate opportunity to improve the efficacy of groundwater monitoring for regulated activities. The establishment of baseline groundwater quality is an important step to enable the early identification of new groundwater pollution as a result of a regulated activity. It is applicable to all licensed sites that are required to undertake groundwater monitoring, regardless of whether or not there is existing groundwater pollution. The early identification of new groundwater pollution associated with a regulated activity will enable management measures to be put in place to:

- prevent significant pollution from occurring at a site where groundwater pollution has not previously been identified
- avoid *further* pollution from occurring at a site where groundwater pollution already exists.

The broad steps to establish baseline groundwater quality that are outlined in this guideline are:

1. Develop a conceptual understanding of the groundwater system.
2. Collect baseline groundwater quality data.
3. Interrogate the data to identify any patterns, trends or anomalies.
4. Report a baseline groundwater quality data set.

Other guidelines in this series will be progressively published with an anticipated completion date in 2026.

1 Introduction

Groundwater is a valuable natural resource throughout South Australia. It represents the largest source of freshwater in the state and is critical to the health of ecological communities and the viability of the agricultural, pastoral, mining and tourism industries. Groundwater is also used for domestic water supply, irrigated horticulture and a variety of industrial applications, as well as sustaining important Cultural Flows of First Nations (Landscape South Australia 2023). As the demand for groundwater is projected to increase in the future, the careful protection of this valuable resource is critical.

Like all water resources, groundwater is susceptible to pollution. Certain industries and land uses, as well as poor waste management practices, can result in pollutants seeping into the underlying groundwater and adversely impacting its quality. Groundwater pollution can have a significant effect on human health, ecosystems and socio-economic development.

Activities that have the potential to pollute groundwater quality are regulated by the South Australian Environment Protection Authority (EPA) and are typically required to implement a range of pollution prevention measures. Regular groundwater quality monitoring, assessment and reporting remains an important tool to validate the effectiveness of those measures.

It is important to note that the primary objective for groundwater quality monitoring of regulated activities is distinct from monitoring undertaken for the assessment of site contamination. Site contamination assessments typically focus on an assessment of risk associated with exceedances of default guideline values[1]. In contrast, the key objective for groundwater quality monitoring of regulated activities is simply to determine whether or not the activity is impacting the existing groundwater quality, resulting in new groundwater pollution[2], or adding to existing groundwater pollution. Impacts to groundwater quality can typically be measured as a statistically significant change to the existing groundwater quality[3]. Consequently, a key requirement to effectively identify impacts to groundwater quality (as a result of an activity) is to establish a good understanding of the existing, or baseline, groundwater quality.

1.1 Description

Baseline groundwater quality is simply described as the existing groundwater quality (in the area of interest), with consideration of the variability expected of a dynamic system. It ideally represents the groundwater quality for a period of time (at least two years) prior to the commencement or continuation of a regulated activity, such that it enables the future assessment of potential new impacts to groundwater quality as a result of that activity.

Baseline groundwater quality is distinct from a determination of ‘background concentrations’, which is a requirement for site contamination assessments. Background concentrations represents groundwater quality associated with naturally occurring and ambient chemical substances, but it must be free of influence from

---

[1] Default guideline values typically relate to the environmental values of the groundwater being assessed, as defined in clause 6 of the Environment Protection (Water Quality) Policy 2015. Examples of relevant guidelines include the *Australian & New Zealand Guidelines for Fresh and Marine Water Quality* (ANZG 2018), *Australian Drinking Water Guidelines* (NHMRC, NRMMC 2011), and *Guidelines for Managing Risks in Recreational Water* (NHMRC 2008).

[2] Pollution is defined in the *Environment Protection Act 1993.* It is important to note that pollution is not defined by the presence of actual or potential harm, and as such it is distinct from site contamination.

[3] Refer to Guideline 4 – Establishing groundwater quality assessment criteria (to be published).

pollution caused by a defined incident, activity or source[4]. Baseline groundwater quality does not have the same constraint, and simply represents the condition of groundwater prior to the commencement of compliance monitoring for the purpose of informing regulatory decision making under the *Environment Protection Act 1993* (EP Act). This can include groundwater quality associated with naturally occurring and ambient chemical substances, as well as the impacts of pre–existing pollution on groundwater quality from point and diffuse sources. Potential sources of pre–existing pollution include:

- the regulated activity
- other activities or incidents at the site of the regulated activity
- an incident, activity or diffuse source of pollution (either current or historical) in the vicinity of the site of the regulated activity.

1.2 Application

The establishment of baseline groundwater quality is becoming increasingly recognised as an essential requirement for assessing impacts to groundwater quality. A well–established groundwater quality baseline can be used to confidently attribute pollution to an activity. In Europe and the United States, baseline groundwater quality is used to understand pollution and impose regulatory limits (IDEQ 2014, Stefania *et al* 2018). Within Australia, the establishment of baseline groundwater quality is rapidly becoming standard practice in the petroleum industry (Bondu *et al* 2021, DMP & DW 2016, DPI 2014); and the mining sector is increasingly recognising the important role baseline groundwater quality plays in successful site closure (Gemson *et al* 2019, McCullough 2016).

Accordingly, for sites with a regulated activity that are required to undertake groundwater monitoring, the EPA expects that baseline groundwater quality is clearly understood and documented. For new activities, baseline groundwater quality should be established through regular monitoring prior to the commencement of operations. For existing activities, where regular groundwater quality monitoring has been undertaken, a suitable baseline[5] may be able to be established using historical data[6]. For existing activities without adequate historical data, it is anticipated that new groundwater quality monitoring will be required in order to establish a suitable baseline. It is also acknowledged that there will be some limited circumstances where baseline groundwater quality cannot be established, such as for sites where groundwater quality is subject to significant continued variation (eg an increasing trend) as a result of impacts from existing pollution. In these circumstances a risk-based assessment process may need to be adopted[7].

1.3 Purpose and scope

The EPA recommends that baseline groundwater quality is described through the establishment of a baseline groundwater quality data set, and the purpose of this document is to provide guidance for the establishment of that data set for the site of an existing or proposed regulated activity. It is important to note that this document only provides basic guidance, outlining the minimum steps that are recommended in the establishment of a robust baseline groundwater quality data set for a typical site. Large or complex sites may require additional steps outside of the scope of this guideline.

---

[4] Refer to SA EPA (2018).

[5] A suitable baseline is one that has a sufficient level of confidence to meet the monitoring objectives (refer to section 3).

[6] Provided that the historical data is still representative of current conditions.

[7] Refer to Guideline 4 – Establishing groundwater quality assessment criteria (to be published).

The intended stakeholders for this guideline are specifically EPA licence holders who are required to undertake groundwater monitoring as a condition of their licence[8]. However, the guideline may also be utilised by other regulators as a suitable reference document for activities that are subject to groundwater monitoring requirements under different legislation. This may include activities regulated under the *Mining Act 1971* or the *Petroleum and Geothermal Energy Act 2000*.

For EPA licence holders who are required to undertake groundwater monitoring as a condition of their licence, the establishment of a baseline groundwater quality data set will be implemented through its inclusion in a groundwater monitoring and management plan (GMMP) and related licence conditions. Once established, the baseline groundwater quality data set can be used to enable clear site–specific assessment criteria to be set for the identification of pollution associated with the regulated activity[9], as well as documenting clear actions to be undertaken if the criteria are exceeded[10].

For activities subject to groundwater monitoring requirements under different legislation, further advice about the process to report the established baseline groundwater quality data set should be sought from the relevant authority.

This guideline is not intended for use by the general public. The guideline is a technical document targeted at personnel with suitable qualifications or experience in groundwater quality assessment and data interpretation. It is anticipated that, where necessary, the services of a suitably qualified or experienced specialist consultant will be procured to interpret and implement this guideline.

The guidance provided in this document assumes that the monitoring bore[11] network used to obtain all relevant data is fit for purpose[12], and that all relevant data have been obtained using sampling and analysis methods that are of an acceptable quality and meet the objective(s) of the monitoring[13]. DES (2021) note that “appropriate bore construction, sampling procedure, monitoring design, laboratory analysis methodology and QA/QC [quality assurance/quality control] processes are essential in providing good quality data” (p 9).

The key steps to establish baseline groundwater quality are outlined in Figure 1, along with the corresponding section of the guideline where further information can be found.

---

[8] Schedule 1 of the EP Act lists activities of environmental significance that require an environmental authorisation in the form of a licence.

[9] Refer to Guideline 4 – Establishing groundwater quality assessment criteria (to be published).

[10] Refer to Guideline 5 – Developing a groundwater monitoring and management plan (to be published).

[11] The term groundwater ‘bore’ or ‘well’ is used interchangeably throughout this document.

[12] Refer to Guideline 1 – Groundwater monitoring bore network design (to be published).

[13] Refer to Guideline 2 – Groundwater sampling (to be published).

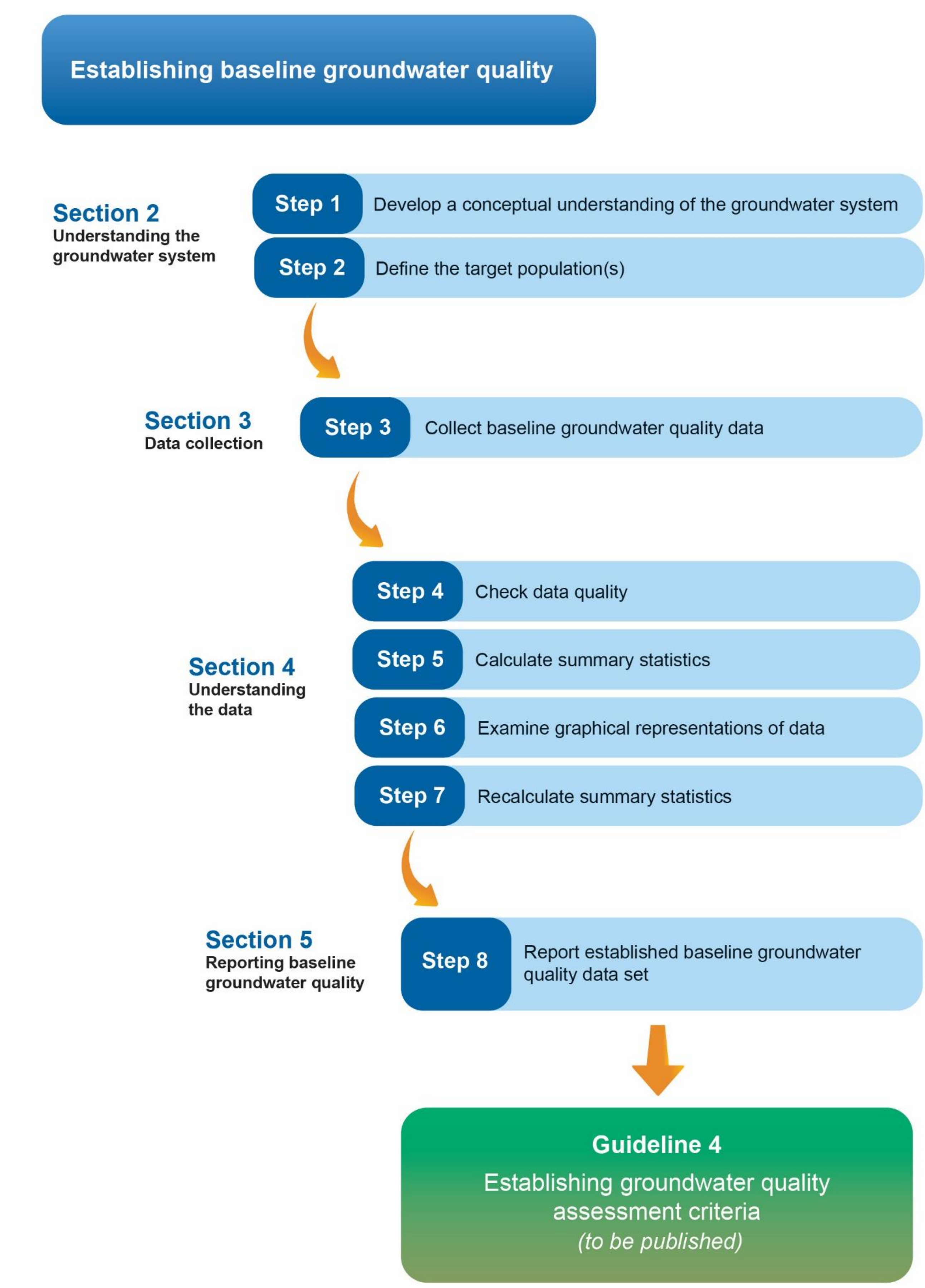


Figure 1 Key steps to establish baseline groundwater quality

2 Understanding the groundwater system

Any assessment of groundwater quality should begin with a solid conceptual understanding of the groundwater system. This includes developing a clear understanding of key elements, such as:

- hydrostratigraphy, including the classification of relevant units[14] and an estimation of their hydraulic properties
- groundwater flow, including an estimation of hydraulic gradients, groundwater flow directions and groundwater velocities
- surface water interaction, including the identification of any significant sources of recharge to, or discharge from, groundwater
- structural or boundary features that could influence groundwater flow directions, rates or quality
- influences on groundwater quality, both natural and anthropogenic.

While it is expected that the initial hydrogeological conceptualisation will have been developed prior to the installation of the groundwater monitoring network, that understanding will be further informed as monitoring data becomes available. A current conceptual understanding of the groundwater system, taking into account all available monitoring data, should be used to inform the establishment of a baseline groundwater quality data set.

It is important to note that a hydrogeological conceptualisation is distinct from a conceptual site model typically used for site contamination assessments, which focusses on a risk framework in the context of source–pathway–receptor linkages. Instead, a conceptual understanding of the groundwater system for the purpose of establishing a baseline groundwater quality data set, should focus on describing the hydrogeological framework and the processes governing groundwater flow, quality, and surface water interaction[15].

2.1 Groundwater flow and surface water interaction

The determination of baseline groundwater quality should be based on a solid understanding of groundwater flow and its interaction with surface water. It is important to acknowledge the dynamic nature of groundwater systems and consider changes to groundwater flow and surface water interaction based on spatial and temporal variability.

2.1.1 Spatial variability

When developing a conceptual understanding of the groundwater system, the focus should specifically be on understanding the groundwater system and flow at a *local* (project) scale, within the context of the *regional* hydrogeology. A local groundwater flow system may be distinct from flow at the regional scale due to natural conditions, such as a geological boundary feature or a groundwater recharge area. A local groundwater flow system may also be induced by anthropogenic activities, such as groundwater abstraction, agricultural irrigation, or deep excavation works.

---

[14] This includes aquifers and confining layers such as aquitards.

[15] Including the identification of groundwater sources and sinks.

When developing a conceptual understanding of the groundwater system, it is not acceptable to simply assume that the groundwater system at a local scale is consistent with the groundwater system at a regional scale without sufficient supporting information.

2.1.2 Temporal variability

Groundwater movement and surface water interaction can change over a range of temporal scales, particularly due to climate variability. This can include:

- Short-term changes in response to episodic events, such as a major storm or flood event,
- periodic changes in response to seasonal climate variability, such as changes to rainfall or surface water flows between summer and winter
- long-term changes as a result of natural climate variability or anthropogenic-induced climate change, such as sustained increased temperatures or decreased rainfall.

Short- and long-term changes in groundwater movement and interaction with surface water can also be induced by anthropogenic activities. Examples include intensive irrigation and dewatering activities.

When establishing a baseline groundwater quality data set, it is important to ensure that, as far as practicable, all of the data used is representative of the existing condition of the groundwater system being assessed with respect to groundwater movement and surface water interaction. This is particularly important where historical groundwater quality data is used that pre-dates any significant changes to groundwater hydraulics that have occurred at or within the vicinity of the site[16].

2.2 Groundwater quality

The determination of baseline groundwater quality should also be based on a solid understanding of the processes that control groundwater quality, both natural and anthropogenic (Bondu *et al* 2022). It is important to note that groundwater quality is not as dynamic as groundwater flow and surface water interaction, and changes to groundwater quality are typically slower to occur when compared with groundwater hydraulics.

2.2.1 Natural

Natural groundwater chemistry can vary widely (temporally and spatially) as a function of the many complex geological, geochemical, hydrogeological and climatic factors which control chemical evolution[17] (Shand *et al* 2007). It is important that the key factors controlling chemical evolution are adequately understood, so that significant variation in baseline groundwater quality due to natural hydrogeochemical processes can be predicted and identified.

2.2.2 Anthropogenic

The key anthropogenic process contributing to variability in baseline groundwater quality is pollution. Before establishing baseline groundwater quality, there is a need to identify all major sources of pollution (both point and diffuse) impacting groundwater within the area of interest. Polluted groundwater should only be included in the establishment of baseline groundwater quality if the plume associated with the primary source(s) of

---

[16] For example, changes in surface water interaction due to decreased rainfall as a result of climate change, or changes in groundwater flow due to intensive irrigation activities.

[17] The chemical evolution of groundwater occurs via a complex sequence of mineral reactions, gaseous exchange, sorption reactions, mixing and dilution (Shand *et al* 2007).

pollution has been determined to be stable[18] such that ongoing impacts to groundwater quality are predictable and unlikely to significantly increase in the future[19]. Conversely, polluted groundwater should not be included in the establishment of baseline groundwater quality when the plume associated with the primary source(s) of pollution has not been determined to be stable, as this may result in impacts to groundwater quality that are unpredictable and can potentially increase in the future[20]. It is also recommended that groundwater from sources for which there are current trends of increasing concentrations of pollutants of concern (POCs)[21] be excluded from use in the establishment of a baseline groundwater quality data set. It is important to note that this would not preclude the use of other bores to establish a baseline data set for those POCs.

2.3 Target population

A target population for baseline groundwater quality can be described spatially and temporally as:

- a distinct volume of groundwater (delineated horizontally and vertically) within the extent of the groundwater system being assessed, and
- groundwater representing a defined interval of time, which may include cyclic intervals[22].

As part of the establishment of a baseline groundwater quality data set, it is important to ensure that all target populations within the area of interest are identified and documented. Target populations will initially be inferred by the conceptual understanding of the groundwater system, and either verified or updated following the exploratory data analysis process (see section 4).

In some instances, the baseline groundwater quality within the extent of the groundwater system being assessed may not exhibit significant spatial variability. In those circumstances, the baseline groundwater quality may be represented as a single target population using all of the available data. In other instances, spatial variability (vertical or horizontal) of baseline groundwater quality within the system may be identified, leading to a need for multiple target populations represented by subsets of data that make up the overall baseline groundwater quality data set. A common example is the variability of baseline groundwater quality between hydrogeological units, leading to a distinct target population (and subset of data) for each unit. Other examples include variability in baseline groundwater quality within the extent of the groundwater system being assessed due to the presence of structural features, significant sources of groundwater recharge, or existing groundwater pollution. This may lead to multiple target populations (and subsets of data) being required to establish an effective baseline groundwater quality data set.

Target populations can also be described temporally. A common example is where a significant variation in groundwater quality occurs between high and low rainfall seasons. In those circumstances, there may be a need for multiple target populations to be identified and documented to ensure an adequate baseline groundwater quality data set from which to establish assessment criteria for seasonal compliance monitoring.

It is also important to note that wells comprising the target populations may be variable between analytes (both spatially and temporally). For example, relative spatial homogeneity of nitrate concentrations may be identified across all wells at a site, with only a subset of those wells identifying impacts from petroleum hydrocarbons. In this instance the target populations for nitrate and petroleum hydrocarbons will be different.

---

[18] In extent and concentration.

[19] Regardless of whether or not the primary source(s) of pollution is still active/uncontrolled.

[20] Regardless of whether or not the primary source(s) of pollution is inactive/controlled.

[21] Appendix A of Northern Territory Environment Protection Authority (2017) provides a list of common pollutants associated with a variety of industries, activities and land uses. It is important to note that it is only a guide and not an exhaustive list.

[22] For example, seasonal target populations.

While the data from all of the wells can be used to establish baseline groundwater quality for a single target population for nitrate, at least two target populations (represented by subsets of data) will be required to establish baseline groundwater quality for petroleum hydrocarbons[23]. In some instances, such as for sites with significant spatial variability of existing groundwater pollution, a target population for a specific analyte may comprise data from only a single well.

A conceptual understanding of the groundwater system is critical for determining the nature and stability of the target population(s) of groundwater measurements through representative sampling (ITRC 2013)[24]. A hydrogeological conceptualisation should be clearly documented as part of the process to establish a baseline groundwater quality data set. This can be done graphically, in writing, or using a table or flowchart, depending on the specific groundwater system. An example is provided in Appendix A.

---

[23] A target population comprising wells with identified petroleum hydrocarbon impacts, and a target population comprising wells with no identified petroleum hydrocarbon impacts.

[24] A representative sample is one with key statistical characteristics that parallel the characteristics of the target population (ITRC 2013).

3 Data collection

Sites with activities that are operational and/or have undertaken regular groundwater monitoring may already have sufficient data to establish a robust baseline groundwater quality data set[25]. This can be determined through analysis of the data (refer to section 4). For those sites without sufficient data to establish a robust baseline groundwater quality data set, additional groundwater monitoring will be required.

3.1 Monitoring well network

Sites requiring additional groundwater monitoring can broadly be categorised into two groups:

- sites with a new regulated activity that has not yet commenced operations (new activities)
- sites with an existing regulated activity where operations have already commenced (existing activities).

The key difference between the two categories is that sites with an existing regulated activity (where operations have already commenced) will need to consider the regulated activity as an additional potential source of pollution when establishing a baseline groundwater quality data set. The hydrogeological conceptualisation will still be critical to inform data collection requirements for both categories.

Groundwater from all monitoring wells, and any other suitable wells[26] or other sampling locations (eg springs, seepage), can be used to collect baseline groundwater quality data.

The well network for baseline groundwater quality data collection should comprise each well, or groupings of wells, within each target population that describes the baseline groundwater quality. Wells should be selected such that they enable consistent sampling methodology to be employed across the network (see section 4.1.1), as well as the collection of representative samples of groundwater from the identified target population(s). It is important to note that any well used for the collection of baseline groundwater quality should be constructed in accordance with the Minimum Construction Requirements for Water Bores in Australia (NUDLC 2000)[27].

Ideally, data should be collected from as many wells (or other sampling locations) as possible within the area of interest, regardless of whether or not those wells are intended to be used for monitoring in the future. This will ensure sufficient data are available to establish a robust baseline groundwater quality data set and identify any spatial heterogeneity (including vertical heterogeneity) to inform whether or not further delineation of distinct target populations within the area of interest is required[28].

---

[25] It is essential to ensure that any existing groundwater monitoring data intended to be used to establish a baseline groundwater quality data set represents the target population(s).

[26] Suitable wells are those that are in good condition and enable the collection of a sample that is representative of the target population. For example, wells with a pump and headwork installed to enable high flow abstraction may not be suitable for the collection of baseline groundwater quality data for volatile and semi-volatile analytes.

[27] Where a well is constructed using the open-hole method, care should be taken to ensure sampling is undertaken at a consistent depth within the production zone for each monitoring event (where possible).

[28] This will also be informed by a good conceptual understanding of the groundwater system and knowledge of different hydrogeological characteristics within the area of interest.

As a minimum, it is recommended that at least three wells[29] are monitored for each aquifer of interest, although more wells are likely to be required for large sites with an extensive compliance monitoring network, or sites where significant spatial heterogeneity or variable groundwater quality is identified[30].

The collection of data from as many wells (or other sampling locations) as possible within the area of interest will also assist in the identification of existing pollution. This may be associated with previous site uses, other current or historical activities within the vicinity of the site, or a regulated activity at the site (where operations have already commenced). The identification of existing pollution is particularly important where it includes POCs that are also associated with the activity being regulated. Where groundwater quality is already impacted, the impacts may be different at each well due to the location and proximity of each well to the pollution source(s). This may result in an increase in the number of target populations required to adequately describe the baseline groundwater quality.

For some large and complex sites (eg mining operations), the groundwater from privately owned production wells may be monitored during operations to identify impacts to specific groundwater users. Where it is intended to monitor groundwater from a private well during operations, baseline groundwater quality should first be established for that well[31].

While the focus of data collection for the establishment of a baseline groundwater quality data set will typically be on the uppermost water bearing unit, consideration should also be given to any underlying aquifers identified in the hydrogeological conceptualisation that may also be susceptible to pollution as a result of the regulated activity(ies). Examples of activities that have the potential to impact underlying aquifers include underground mining, landfilling into deep excavated cells, or the injection of water into deeper aquifers. Conversely, the potential for pollution of overlying aquifers, through upwards vertical leakage, may also need to be considered for certain activities and/or sites[32].

Where there is a potential source of pollution located up hydraulic gradient from a regulated activity, it may be useful to undertake additional groundwater monitoring to understand the quality of groundwater migrating across the up hydraulic gradient boundary of the site. Monitoring wells used for this purpose should be located up hydraulic gradient of the regulated activity, but down hydraulic gradient of the potential source of pollution. The number of up hydraulic gradient wells required will be based on site–specific considerations, including hydrogeological factors, the spatial heterogeneity of groundwater quality in the area[33], and the number and location of any potential up hydraulic gradient sources of pollution. It should be noted that the use of monitoring wells up hydraulic gradient from a regulated activity is not limited to circumstances where a potential source of pollution is identified. Up hydraulic gradient monitoring wells can also be used to monitor for any significant changes in baseline groundwater quality migrating across the up hydraulic gradient boundary of the site over time, including as a result of natural processes. This is likely to be most applicable to sites where a regulated activity is proposed to be undertaken for a significant length of time[34].

---

[29] Located to be as spatially representative of the entire site as possible.

[30] It is important to note that an increase in the number of wells monitored to establish a baseline groundwater quality data set, will typically result in an increased level of confidence in the baseline established, potentially reducing the identification of false positives and delivering a long-term cost benefit.

[31] It is important to note that any groundwater pollution identified in a private well may form information that is required to be included in the Public Register (section 83 or 83A of the EPA Act). For this reason, it is recommended that informed consent is obtained from well owners prior to any monitoring being undertaken on private land.

[32] For example, carbon capture and storage operations.

[33] This will be a particularly important consideration for sites with a large up hydraulic gradient site boundary.

[34] For example, a landfill or mine site.

3.2 Duration and frequency

The duration of baseline groundwater quality monitoring and frequency of sampling should be sufficient to establish confidence in the temporal variability of groundwater quality within the area of interest (DES 2021). In South Australia, this includes consideration of the effects of seasonality and longer-term climate variability (eg dry vs wet periods), as well as the ongoing impacts of climate change. While there is limited guidance in Australia on the frequency of sampling to identify seasonal variability, Barcelona et al (1989b) recommends that quarterly sampling is a good initial starting point, although bi–monthly monitoring may be more appropriate for reactive chemical constituents or in areas where groundwater quality is known (or suspected) to change over short time frames[35]. Barcelona et al (1989a) suggests that 3–6 samples per year provides a cost-effective monitoring frequency (with little sampling redundancy), and that sampling frequencies greater than monthly have the potential to result in considerable loss of information through serial correlation[36]. The duration of baseline groundwater quality monitoring required to establish confidence in the effects of longer-term climate variability, as well as climate change, will need to be assessed on an individual site/region basis[37].

The Interstate Technology and Regulatory Council (ITRC 2013) indicates that most guidelines on sample size for groundwater statistical tests recommend at least 8–10 samples when constructing (statistical) limits for a target population. However, ANZG (2018) points out that the precision with which percentiles for a data set are estimated depends heavily on the sample size. This means that for a relatively small sample size of 8–10 samples, only a limited range of percentiles around the 50th percentile (median) are likely to be able to be estimated with relatively good precision. In order to estimate more extreme percentiles with the same precision for each target population, a larger sample size is required. ANZG (2018) provides some guidance[38] for sample sizes required to estimate percentiles with relative precision for non-parametric[39] data sets. It suggests that to estimate the 20th or 80th percentile of a data set with an associated 95% confidence interval, a minimum sample size of 17 is required; and that to estimate the 10th or 90th percentile of a data set with the same precision, a minimum of 36 samples is required.

When undertaking groundwater quality monitoring to establish a baseline data set, it is important to consider the advantages associated with a larger sample size. For example, if regulatory assessment criteria are estimated from baseline percentile statistics, the level of confidence in those criteria will depend on the baseline groundwater quality sample size[40]. Small sample sizes often have a low level of confidence in the estimated percentiles, which means false negatives[41] and false positives[42] are more likely to be reported.

---

[35] More frequent and/or targeted monitoring of certain parameters (eg water levels, electrical conductivity), including through the use of data loggers, may also be beneficial in establishing a good conceptual understanding of the groundwater system, as well as the baseline groundwater quality.

[36] Serial correlation can occur when samples are collected too frequently, resulting in data points which are not independent and provide no new information about groundwater quality.

[37] Climate data across South Australia is available on the Bureau of Meteorology website http://www.bom.gov.au/climate/data/index.shtml. Historical groundwater monitoring data from, or in the vicinity of, a site (if available) may also provide important insights into any long-term climate variability.

[38] Based on the work of Goudey (2007).

[39] Non-parametric data sets have no identified or assumed population distribution.

[40] Refer to Guideline 4 – Establishing groundwater quality assessment criteria (to be published).

[41] A false negative can occur when an impact to groundwater quality is present, but the assigned assessment criteria indicate that an impact is not present.

[42] A false positive can occur when no impact to groundwater quality is present, but the assigned assessment criteria indicate that an impact is present.

False negatives can result in a groundwater quality impact not being detected, while false positives can result in unnecessary and costly compliance actions being initiated.

The EPA recommends that the number of data points (for each target population) required to establish a baseline groundwater quality data set is commensurate with the level of risk associated with the regulated activity(ies). As a minimum, groundwater should be monitored at least quarterly[43] for at least two years to assess for variability in quality and establish a baseline data set[44]. This should provide at least eight data points for each monitoring well (or other sampling location), which can potentially be combined into a larger single data set if no significant heterogeneity is identified between the groundwater quality from each well (or other sampling location).

For regulated activities which have a high level of risk (such as many landfills and mine sites) or are located adjacent to sensitive receptors (such as groundwater users or groundwater dependent ecosystems), a greater level of confidence in the baseline groundwater quality data set is recommended in order to establish effective assessment criteria for compliance monitoring. In such instances, the EPA recommends a minimum sample size of seventeen for each monitoring well (or other sampling location). This can be achieved through sampling for baseline groundwater quality more regularly (than quarterly)[45], and/or sampling for a longer period (more than two years).

Where a sampling period greater than two years is required to ensure the establishment of a baseline groundwater quality data set with a greater level of confidence, it may be appropriate to establish an interim baseline groundwater quality data set[46] which can then be used to develop interim groundwater quality assessment criteria prior to the commencement of the regulated activity. Where an interim baseline groundwater quality data set is established, it should be clearly reported as interim and include a clear timeframe for the collection of additional data to update the baseline groundwater quality data set to achieve the required level of confidence[47].

There may be exceptional circumstances where a baseline groundwater quality data set (and subsequent assessment criteria) is required to be established within a timeframe shorter than two years to facilitate the commencement of an activity[48]. Where this is the case, and there is sufficient information to support the requirement, the use of an interim baseline groundwater quality data set may be appropriate. A suitable interim data set may be achieved by undertaking baseline groundwater quality monitoring more regularly than quarterly to ensure at least eight data points are available for each sampling location within the

---

[43] In the same months each year (where possible).

[44] The literature suggests that long-term monitoring (more than two years) is required to adequately describe seasonal trends in groundwater. However, a minimum of two years is considered adequate for the purposes of establishing a baseline groundwater quality data set for compliance monitoring.

[45] If groundwater sampling is undertaken more frequently than monthly, it should be analysed for potential serial correlation.

[46] Based on two years of monitoring with at least eight data points for each monitoring well (or other sampling location).

[47] It will be prudent to ensure that any additional baseline groundwater quality monitoring undertaken after the commencement of the regulated activity only occurs at locations that are anticipated to be free from impacts associated with the regulated activity.

[48] For example, where there is an unavoidable short timeframe between the scoping and commencement of an activity.

timeframe required[49]. The interim baseline groundwater quality data set can then be used to develop interim groundwater quality assessment criteria for compliance monitoring, until a more robust baseline[50] can be established following the commencement of the regulated activity. It is important to note that this should only occur in exceptional circumstances[51]. Where an interim baseline groundwater quality data set is established using data collected within a timeframe shorter than two years, it should be clearly reported as interim, and a process outlined for the continued collection of additional data[52]. When baseline monitoring data spanning at least two years becomes available, the interim baseline groundwater quality data set should be updated as a matter of priority.

For some activities which are only anticipated to operate for a short duration (less than about two years), it may be appropriate to establish a baseline groundwater quality data set (and assessment criteria) using data collected over a reduced timeframe[53]. Where this is the case, it is still anticipated that at least eight data points will be available for each sampling location[54], or alternatively at least eight data points for each target population[55]. In these circumstances, it is anticipated that the established baseline groundwater quality data set would not be considered interim, and no further updates would be required.

It is anticipated that, for most activities, baseline groundwater quality monitoring will continue to be undertaken following the establishment of a baseline groundwater quality data set and the commencement of the regulated activity. The continued collection of baseline groundwater quality data will result in a larger sample size and a subsequent increase in the level of confidence of the associated assessment criteria[56]. This will assist in the reduction of false negatives and false positives when undertaking compliance groundwater quality monitoring. Careful consideration will need to be given to the monitoring locations for baseline groundwater quality data collection following the commencement of the regulated activity. Monitoring should only occur at locations that are anticipated to be free from any impacts associated with the regulated activity[57], or that can be inferred as being free from impacts associated with the regulated activity through the continued reporting of analyte concentrations within the range of the initially established baseline data set. The continued collection of baseline groundwater quality data is also an important process to establish any changes to the baseline over time. This may include natural impacts associated with hydrogeochemical processes, episodic events, and longer–term climate variability; as well as potential anthropogenic impacts from sources adjacent to operations associated with the regulated activity.

---

[49] Being careful to avoid serially correlated data.

[50] Using monitoring data collected across a timeframe of at least two years.

[51] For most activities, it is expected that baseline groundwater quality will be considered at the project scoping stage, so that sufficient time (at least two years) is allowed to undertake sampling to establish an adequate baseline groundwater quality data set prior to the commencement of operations.

[52] It will be prudent to ensure that any additional baseline groundwater quality monitoring undertaken after the commencement of the regulated activity only occurs at locations that are anticipated to be free from impacts associated with the regulated activity.

[53] Less than two years.

[54] Through sampling more often than quarterly.

[55] Where an initial data assessment has identified multiple sampling locations as being representative of a single target population.

[56] Refer to Guideline 4 – Establishing groundwater quality assessment criteria (to be published).

[57] For example, baseline groundwater quality monitoring at an up hydraulic gradient location to the regulated activity; or monitoring at a down hydraulic gradient location that is a sufficient distance from the potential source(s) of pollution so that any impacts to groundwater quality would not yet be identified at that location.

3.3 Analysis

Analytes monitored as part of baseline groundwater quality should include key POCs specifically associated with the regulated activity[58], as well as other water quality characteristics (pH, electrical conductivity[59], temperature, redox potential, dissolved oxygen) and filtered metal concentrations that can be used to enhance understanding of the groundwater system and assist in data quality checks. Major cation and anion species[60] should also be monitored in order to characterise groundwater and allow the ionic signature to be compared between wells (DES 2021)[61].

While groundwater quality monitoring at a regulated site during operations is expected to focus on the identification of impacts associated with the regulated activity(ies), it is recommended that POCs associated with all potentially contaminating activities (PCAs) undertaken at a site[62] are included in the list of analytes to be collected for baseline groundwater quality sampling[63]. A number of studies have reported that POCs which are initially perceived to be low risk at a site[64], and subsequently do not have an established baseline, can often result in significant data gaps when undertaking an assessment of impacts at a site prior to closure and the cessation of formal regulation (Weaver *et al* 2022, Sweeney *et al* 2022). This can result in protracted monitoring requirements and an increase in costs for licensees to achieve closure. When considering the potential to establish baseline groundwater quality for a broader range of POCs, it is important to consider that even a small data set collected over a shorter period of time (compared with the key POCs) may still provide a long-term benefit for achieving site closure.

When determining a list of analytes for the collection of baseline groundwater quality data, it is also prudent to consider analytes with naturally elevated concentrations in groundwater. For example, groundwater in the vicinity of a mine site may be naturally enriched in elements associated with the target mineral deposit. If those naturally elevated concentrations are not identified prior to the activity commencing, subsequent groundwater monitoring may erroneously associate them with the activity, resulting in unnecessary work to identify and manage any perceived risks. Once again, this can result in protracted monitoring requirements and an increase in costs to achieve closure.

It is recognised that for sites where the regulated activity has already commenced (existing activities), periodic groundwater quality compliance monitoring may have already been undertaken, resulting in baseline groundwater quality data being available for some analytes. In these instances, data may be available for some POCs associated with the regulated activity, but not for other analytes recommended for inclusion in the monitoring[65]. Where there is not sufficient data to establish a baseline groundwater quality data set for all

---

[58] The baseline data sets associated with these analytes can then be used to establish groundwater quality assessment criteria.

[59] It is recommended that pH and electrical conductivity (total dissolved solids) is assessed through a combination of field measurements and laboratory analysis.

[60] This includes sodium, potassium, calcium, magnesium, chloride, sulphate and bicarbonate/carbonate.

[61] The major ions can also be used to further develop the conceptual understanding of the groundwater system.

[62] Not just PCAs associated with the regulated activity(ies).

[63] Appendix A of Northern Territory Environment Protection Authority (2017) provides a list of common polltants associated with potentially contaminating activities, including non-prescribed activities (not formally regulated). It is important to note that it is only a guide and not an exhaustive list.

[64] Not typically associated with the formally regulated activity(ies).

[65] Such as those associated with other PCAs, or analytes that can be used to enhance understanding of the groundwater system and assist in data quality checks.

of the recommended analytes[66], it may still be appropriate to establish an interim baseline groundwater quality data set for a reduced set of analytes[67], which can then be used to develop interim groundwater quality assessment criteria for key POCs[68]. Where an interim baseline groundwater quality data set is established, it should be clearly reported as interim and a clear methodology included for the collection of additional data to update the baseline groundwater quality data set to include all of the recommended analytes.

---

[66] Or undertake all of the data analyses required (see section 4).

[67] Assuming at least two of the analytes are key POCs associated with the regulated activity.

[68] Where sufficient data is available for those POCs.

4 Understanding the data

In order to establish baseline groundwater quality a preliminary review of the data set should be conducted to gain familiarity with the data and identify any associated anomalies or limitations that need to be managed. This process is typically referred to as exploratory data analysis (EDA), and it is becoming increasingly recognised as an investment of project resources that can yield significant dividends in the future.

The EDA process consists of three key iterative steps:

1. checking data quality
2. calculating summary statistics
3. examining graphical representations of data.

A range of graphical examples of the EDA process (data quality checking and data visualisation) have been provided throughout section 4 of this guideline. Most examples are based on data delineated by wells. However, it is important to note that the same principles can be applied to data sets for any target population, including data sets consisting of pooled data from multiple locations/sources[69] to represent a single target population.

4.1 Checking data quality

A review of the data quality should focus on an assessment of its suitability for statistical analysis. The aim is to ensure each data set represents the target population with a degree of accuracy and precision required to achieve the monitoring objectives. This means that inconsistent sampling and analysis processes, as well as variations and errors in reporting, need to be identified and screened before statistical analysis is undertaken to establish a baseline groundwater quality data set.

All data intended to be utilised to establish a baseline groundwater quality data set for a site should be subject to a thorough data quality check.

Where multiple target populations have been identified through a hydrogeological conceptualisation, the data should be divided into subsets of data representing each target population, before commencing a review of the data quality (see section 2).

4.1.1 Sampling and analysis methodology

In order to be suitable for statistical analysis, all data points meant to represent a target population should have been drawn from that population using a similar collection and measuring process (ITRC 2013). Significant changes in sampling and analysis methodology within a data set (either for a target population or an individual well) have the potential to bias the results and introduce unacceptable levels of uncertainty. This is particularly true for historical data, which may have been collected and analysed using methods that are not consistent with current best practice[70]. It is important that any significant changes in sampling and analysis methodology within a data set are identified and documented. This information may provide an

---

[69] For example, multiple wells across a site.

[70] For example, samples collected using a bailer compared with current low flow methodologies; analyses conducted using a method with lower precision compared to current methods.

important explanation for data anomalies identified through the EDA process while still allowing the potential use of historical data.

In some instances, depending on data quality, it may be appropriate that only a subset of the available data is used for statistical analysis in order to establish baseline groundwater quality. For example, where a data set contains historical data collected using one method, which is not comparable to contemporary data collected using a different method, the historical data may need to be excluded from the data set for the purpose of statistical analysis. An example is provided in Figure 2. It is important to note that this may not preclude the use of the historical data for other aspects of the EDA process, in particular for the generation of graphical representations to visually examine the data (see section 4.3).

It should also be noted that changes in sampling methodology is a common issue encountered for the sampling and analysis of metals, as historical data does not typically clearly indicate if analyses were conducted on filtered (field or laboratory) samples. This needs to be considered when undertaking a data quality check of any metals analyses.

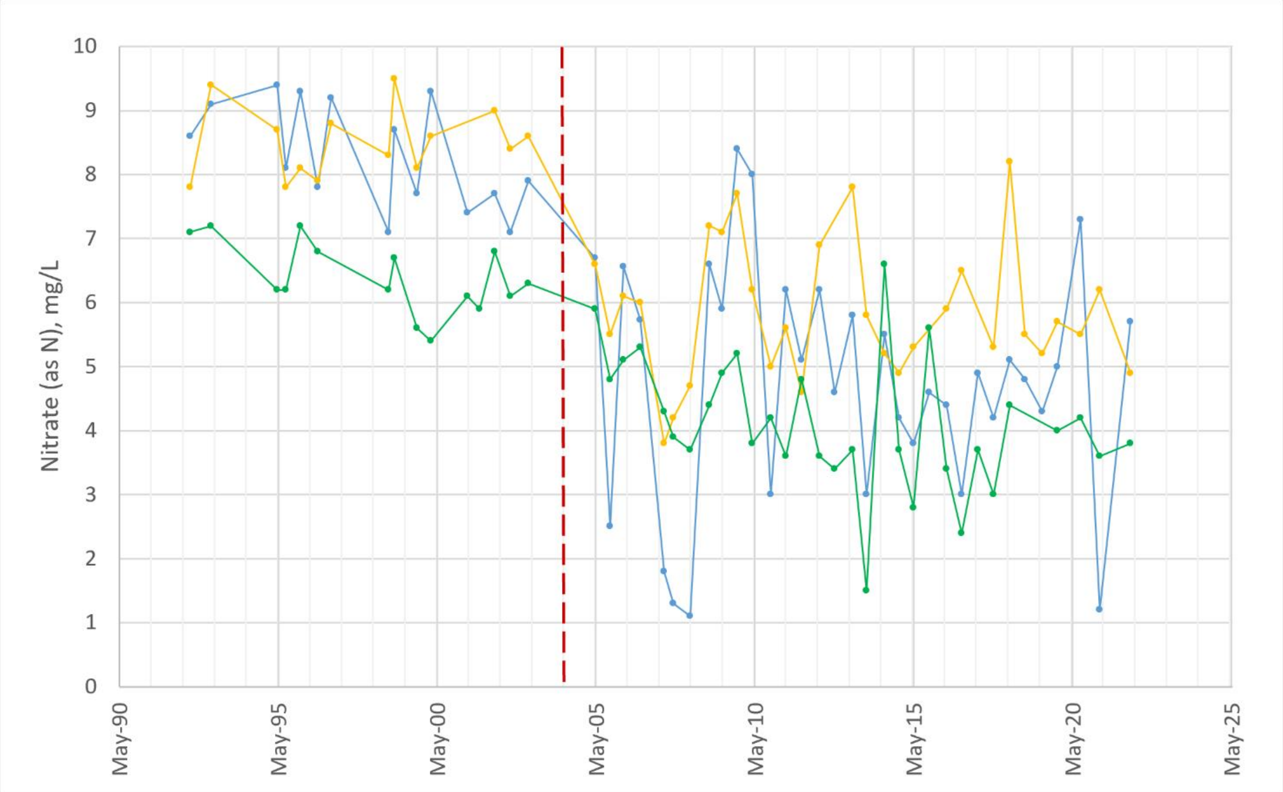


Figure 2 **Graph showing nitrate (as N) concentrations over 30 years for three wells representing the same target population. The red dashed line indicates a significant change in sampling methodology that was applied to all three wells at the same time. An initial visual inspection of the time series plot identifies a decrease in nitrate (as N) concentrations between 2003 and 2005, which may be associated with the change in sampling methodology. As the influence of the change in sampling methodology on these results is not quantified, it would be reasonable to only use the more contemporary data for all three wells to undertake statistical analysis of nitrate data for the target population.**

4.1.2 Sampling frequency

The frequency of sampling within a baseline data set can also potentially impact the results of statistical analysis. For example, seasonal trends can only be identified if data have been collected frequently enough to adequately characterise annual cyclic changes in groundwater quality. An example is provided in Figure 3 for data from a single well. Generally, quarterly sampling of groundwater quality is considered adequate to identify seasonal trends at most sites, although this may vary based on site-specific conditions. Conversely, if locations are sampled too frequently it can result in serially correlated data which can also impact the results of statistical analysis. Figure 4 provides an example of the potential impacts of correlated data. Generally, baseline groundwater quality data collected more frequently than every month should be tested for correlation[71], although this will vary based on site-specific factors, such as the local groundwater velocity[72].

Where duplicate or replicate sample data are available for a particular location and time, only one value should be included for data analysis. When establishing baseline groundwater quality, a conservative approach would typically be to use the lowest value.

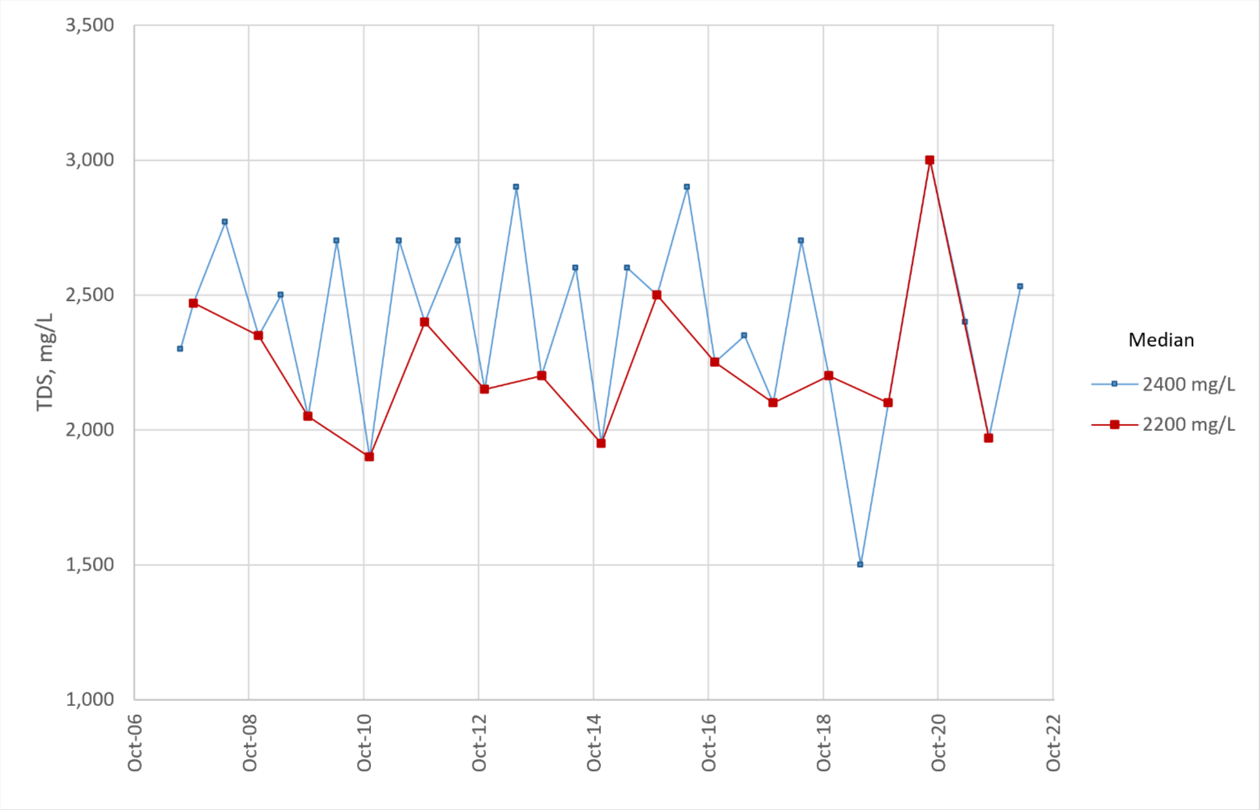


Figure 3 Graph showing total dissolved solids (TDS) concentrations for a single well over 16 years. The blue line shows the results for biannual sampling, which indicates a clear seasonal trend in the data. The red line shows how the time series would look if only annual sampling had been undertaken, with the seasonal trend not evident. The difference in the calculated median (50th percentile) for each data set demonstrates the importance of optimising the sampling frequency when collecting data to establish baseline groundwater quality.

---

[71] Pearson’s correlation coefficient can be calculated for parametric data, and Spearman’s correlation coefficient can be calculated for non-parametric data [refer to Helsel *et al* (2020) for further information]. Both calculations can be performed simply in Microsoft Excel.

[72] For example, it may be useful to undertake more frequent data collection of certain parameters (eg through the use of data loggers) in specific areas of some sites (eg areas of surface water recharge or seawater intrusion) to gain additional valuable information about the groundwater system and variability of baseline conditions.

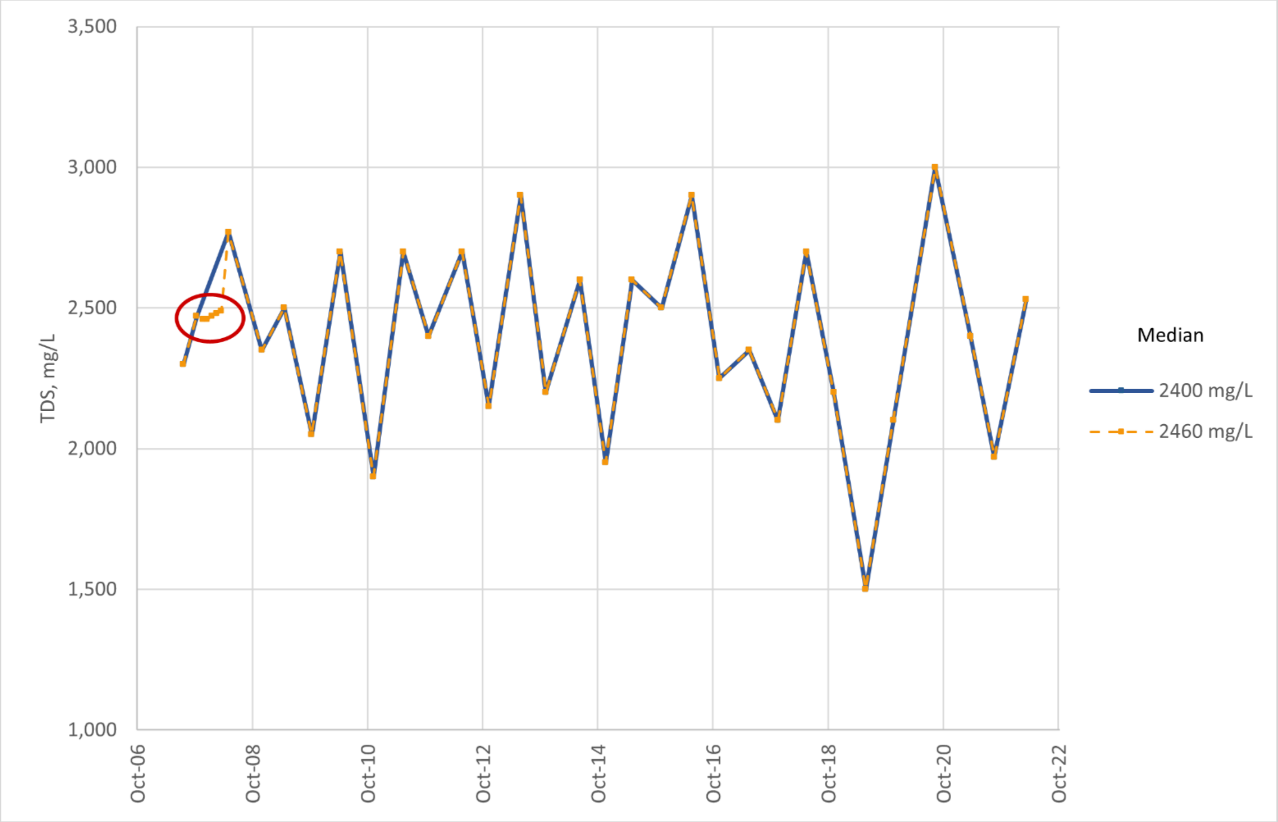


Figure 4 **Graph showing total dissolved solids concentrations over 16 years for the same well as Figure 3. The blue line shows the time series using data from bi-annual sampling. The yellow line shows the time series for the same bi-annual sampling, but also includes an additional six samples taken monthly between October 2007 and March 2008 (circled in red). The graph demonstrates that in this example the monthly samples, which show evidence of correlation, don't provide any valuable additional information about the groundwater quality (compared to the bi-annual sampling). The calculated median for each data set further shows the potential impact correlated data can have on the summary statistics for a data set.**

4.1.3 Reporting limits

The reporting limit (RL) is a laboratory established censoring limit. It typically represents the lowest concentration of an analyte that can be measured by a specific analytical method with a sufficient degree of confidence for each sample. Reporting limits can change for a variety of reasons, including changes in analytical methods, or sample dilution as a result of matrix interference[73]. Changes can also occur on an ongoing basis, or for a defined period of time within a monitoring program. For example, an ongoing change in reporting limit for a specific analyte may occur simultaneously for all wells in a target population based on a change in analytical methodology. Conversely, a temporary change could occur for a single well within a target population based on matrix interference for a single sampling event. When assessing a data set, it is important to establish whether or not the RL is consistent over time for each analyte across each well and target population.

---

[73] Matrix interference occurs when non-target analytes or physical/chemical characteristics of a sample prevent quantification of the analytes of interest.

Where an analyte's data set does not have consistent RLs, it is important to consider how each inconsistency could potentially change or bias the results of any statistical analysis. Any inconsistencies that have the potential to result in data anomalies through the EDA process should be clearly documented.

For example, if the data set for a particular analyte contains multiple RLs, but all of the concentrations reported for that analyte are greater than the relative RL, then the inconsistency will not impact any statistical analysis. An example is provided in Figure 5. Conversely, where a data set contains multiple RLs, and some of the analyte concentrations are reported as less than the RL (or *non–detect*), a significant amount of detail about the data may be lost. This can result in an impact to the statistical analysis of the data. An example is provided in Figure 6.

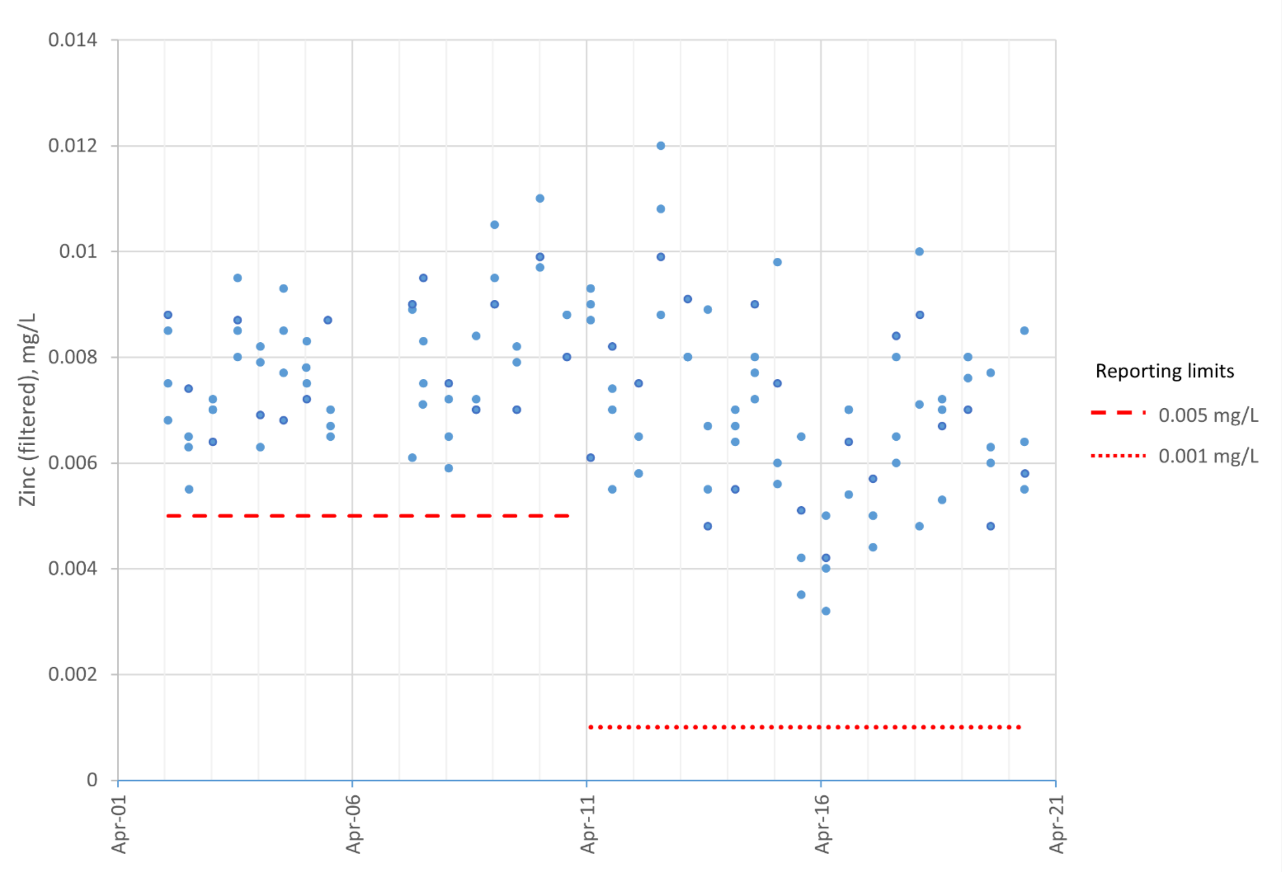


Figure 5 **Graph showing zinc (filtered) concentrations for a target population (comprising multiple wells) over 20 years. Despite the change in reporting limit from 0.005 mg/L to 0.001 mg/L, all the concentrations reported were above the relative reporting limit at the time. Consequently, the change in reporting limit will not impact the statistical analyses or graphical representations of the data.**

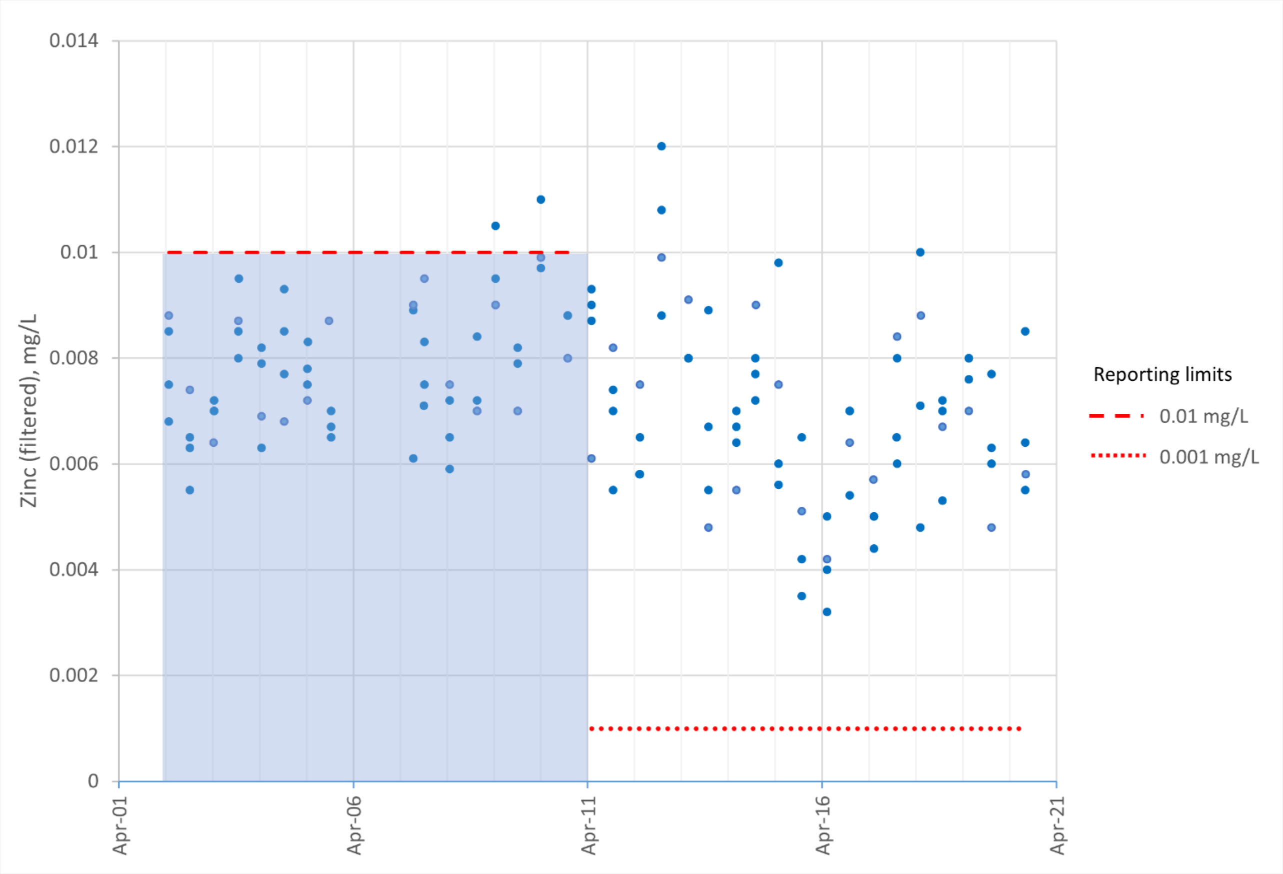


Figure 6 **Graph showing the consequence of a higher initial reporting limit for the zinc (filtered) data from Figure 5. With a reporting limit of 0.01 mg/L, all that is known about the zinc (filtered) concentrations between 2001 and 2010, is that they are <0.01 mg/L. The potential range of values is represented by the blue shaded area in the graph. The pale blue dots underneath show the significant amount of detail about the data that has been lost, which can consequently impact statistical analyses of the data.**

4.1.4 Non-detects

Concentrations of analytes reported below the RL designated by the laboratory are typically referred to as non-detects. For these samples, the only information known about the concentration of the analyte is that it lies somewhere between 0 and the RL. No statistical technique can fully compensate for the information loss due to data censoring as a result of non-detects. The larger the proportion of censored data and the larger the censoring limits (RLs), the greater the information loss and uncertainty (ITRC 2013).

While statistical tests designed for data with no assumed distribution (non-parametric tests) are not as impacted by a small number of non-detects as tests designed for data with an assumed distribution (parametric tests) [74], almost all statistical tests can be confounded by a large number of non-detects, especially if associated with varying RLs (ITRC 2013). As RLs can vary between laboratories, as well as through other factors, such as a change in analytical methodology or the composition of a specific sample, it is not uncommon to have a groundwater quality data set with considerable variability in RLs for non-detect values. As such, careful management of all censored data is required to minimise information loss and uncertainty associated with the data set.

---

[74] The most commonly inferred data distributions are normal, lognormal and gamma.

Historically, it has been common to manage non-detects in data sets by substituting them with another value, such as the RL, half the RL, or 0. However, this can adversely affect the results and introduce bias. Fortunately, there are a range of effective techniques available which are specifically designed to accommodate censored data. For the purpose of undertaking statistical analysis to establish a baseline groundwater quality data set, the following general strategies for managing non-detects can be implemented[75]:

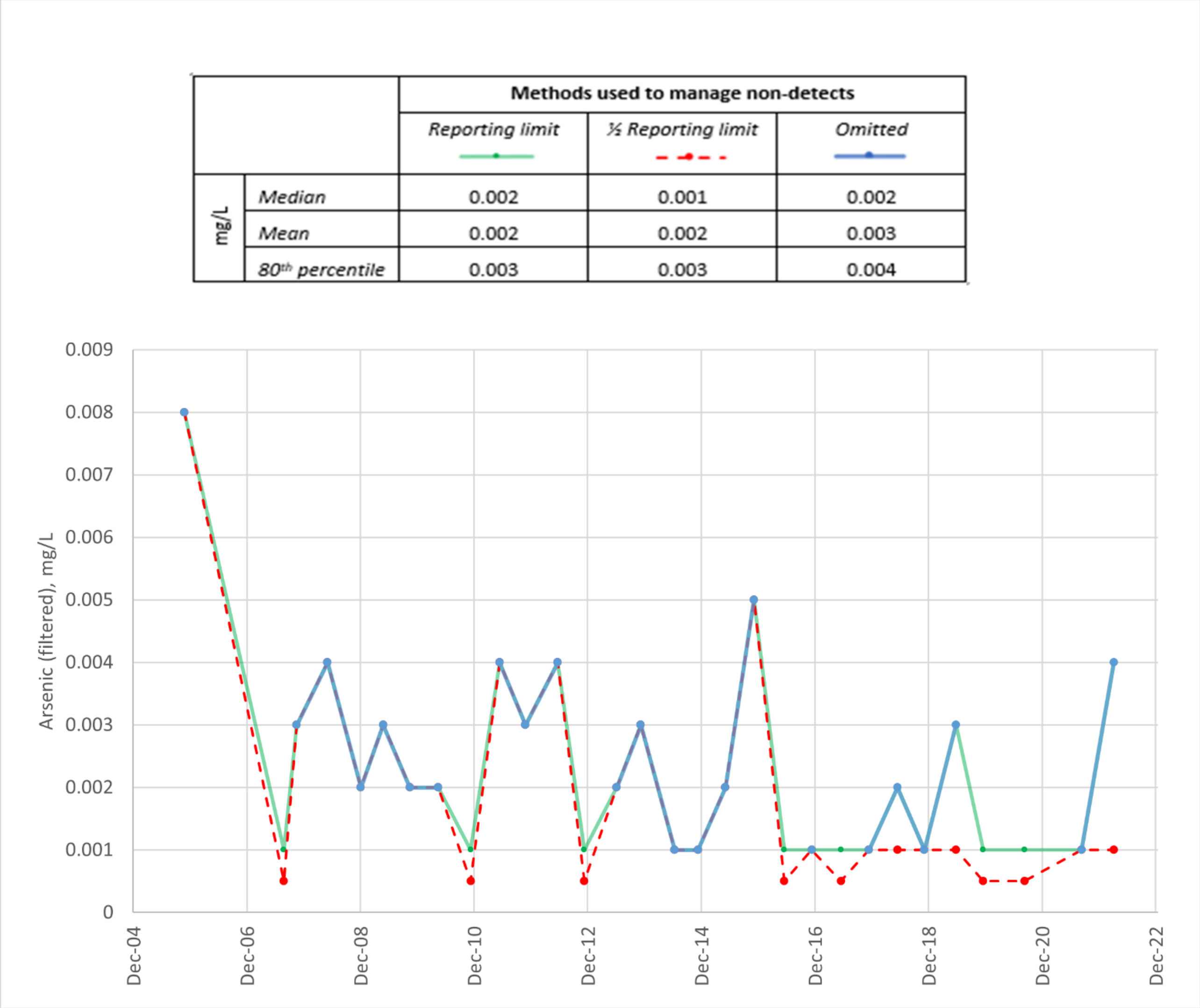
1. use a technique suitable for censored data to estimate summary statistics, such as Kaplan–Meier or Robust Regression on Order Statistics (ROS).
2. use a rank-based non-parametric test to analyse for trends, such as Mann–Kendall.
3. use statistical approaches specifically designed to accommodate censored data when comparing the parameters and distributions of two populations, such as the Tarone–Ware test or the rank-based Wilcoxon–Mann–Whitney test (equivalent to the Wilcoxon Rank Sum test).

It is important to understand the underlying assumptions, as well as strengths and weaknesses, of statistical tests before applying them to a data set. Some tests may have underlying assumptions that preclude their use with certain data sets, and some tests may have weaknesses that can reduce their effectiveness with certain data sets. For example, the Kaplan–Meier test can be used with data sets having a frequency of detection less than 50%, whereas Robust ROS requires a detection frequency greater than 50%. ITRC (2013) provides a basic summary of the assumptions, strengths and weaknesses of common statistical methods. Additional information may also be available in user guides provided with statistical software packages.

For the purpose of generating and examining graphical representations of the data, it is likely that non–detects will either need to be omitted or substituted with another value (such as the RL or half the RL). In these instances, it is important to understand the impact that omission or substitution of non-detects may have on the results generated, particularly for data sets with a higher percentage of non-detects. An example is provided in Figure 7 for data from a single well, but the same principles can be applied to data sets for any target population.

---

[75] All of the listed statistical techniques/tests in this guideline are available within the free ProUCL software statistical package published by the United States Environmental Protection Agency. The software is available to download at https://www.epa.gov/land-research/proucl-software, and the website includes guidance material for use of the software. More detailed information about the methods can also be found in ITRC (2013). The EPA recognises that there are a range of other statistical software packages available for use. For example, Shell Global Solutions and The University of Glasgow provide a free software toolkit for groundwater data analysis (GWSDAT), which is available at https://gwsdat.net/home/.



		Methods used to manage non-detects		
		Reporting limit	*½ Reporting limit*	*Omitted*
mg/L	*Median*	0.002	0.001	0.002
	Mean	0.002	0.002	0.003
	80th percentile	0.003	0.003	0.004

Figure 7 Graph showing arsenic (filtered) concentrations for a single well with different techniques used for the management of non-detect data (RL = 0.001 mg/L). The blue line represents the time series graph where non-detect values have simply been omitted; the green line depicts the use of the reporting limit as a value for all non-detects; the dashed red line shows the result of using half of the reporting limit as a substitute for all non–detects. The table above the graph presents some basic summary statistics for each of these scenarios and demonstrates the impact management of non-detects can have on a data set. For example, the 80th percentile of the data sets varies by up to 25%.

4.1.5 Outliers

Outliers represent data points with values that are significantly different to other observations within the data set, and are typically caused by:

1 a measurement or recording error

2 an observation for a different population to the rest of the data

3 a rare event or extreme variability (eg hotspot) associated with the target population (Helsel *et al* 2020).

It is important that outliers are identified and investigated prior to establishing a baseline groundwater quality data set, as they may be unrepresentative of the target population(s)[76]. Potential outliers in a data set for a target population, can often be identified visually through graphical representations of the data, typically quantile–quantile plots and boxplots (see section 4.3). A formal outlier test, such as Dixon's test or Rosner's test, can also be used to provide a statistical assessment of potential outliers in a data set, although their applicability is limited to normally distributed data (once any suspected outliers are removed).

Where outliers are verified to be the result of a measurement or recording error, the value should be corrected (where possible), or simply omitted (where correction is not possible) and the reason for change or omission clearly documented. Where an outlier is suspected to represent an observation for a different population to the rest of the data, it should only be removed if there is corroborating evidence to suggest it is unrepresentative of the target population[77]. In such an instance, a thorough understanding of the groundwater system is crucial and the basis for removal of the outlier should be clearly documented.

Where the cause of any outliers cannot be confidently determined, it is recommended that all statistical data analyses are conducted with and without the inclusion of the outliers to see how the different results compare. This will provide important information about the implications for how outliers are managed, particularly with regards to the upper and lower statistical limits of the calculated baseline groundwater quality data set.

While outliers should generally be kept as part of the data set unless there is reasonable evidence to suggest they do not represent the target population(s), an outlier with a much higher value than the rest of the data set will tend to increase the statistical limits of the calculated baseline groundwater quality data set, which can potentially diminish its effectiveness to identify pollution associated with a regulated activity. In such a case, it may be justified to remove any high-magnitude statistical outliers before establishing baseline groundwater quality, and instead account for potential observations associated with rare events or extreme variability through the establishment of relevant groundwater quality assessment criteria[78].

An example of the impact of a potential outlier for pH in a single well is provided in Figure 8. It should be noted that, as the pH scale is logarithmic, any statistical analyses of pH data should use the untransformed variable (ie hydrogen ion concentration)[79]. It is also important that any statistical analyses clearly document the use of the hydrogen ion concentration (rather than pH). For ease of visualisation, graphical and tabular representations of data can utilise pH as the variable.

---

[76] Further information about identifying outliers is provided in section 4.3.

[77] For example, concentrations for an analyte from a group of wells in one area of a site that are different from the concentrations for the same analyte from all other wells across the site.

[78] Refer to Guideline 4 – Establishing groundwater quality assessment criteria (to be published).

[79] $pH = -\log_{10}[H^+]$

		Comparison of basic statistical data (<u>calculated</u> using hydrogen ion concentration)	
		With potential outlier	*Without potential outlier*
pH units	*Minimum*	7.3	7.3
	Maximum	11.6	9.6
	Mean	8.5	8.4
	Median	8.2	8.2
	20th percentile	8	8
	80th percentile	9.1	8.8
	99th percentile	11.2	9.6

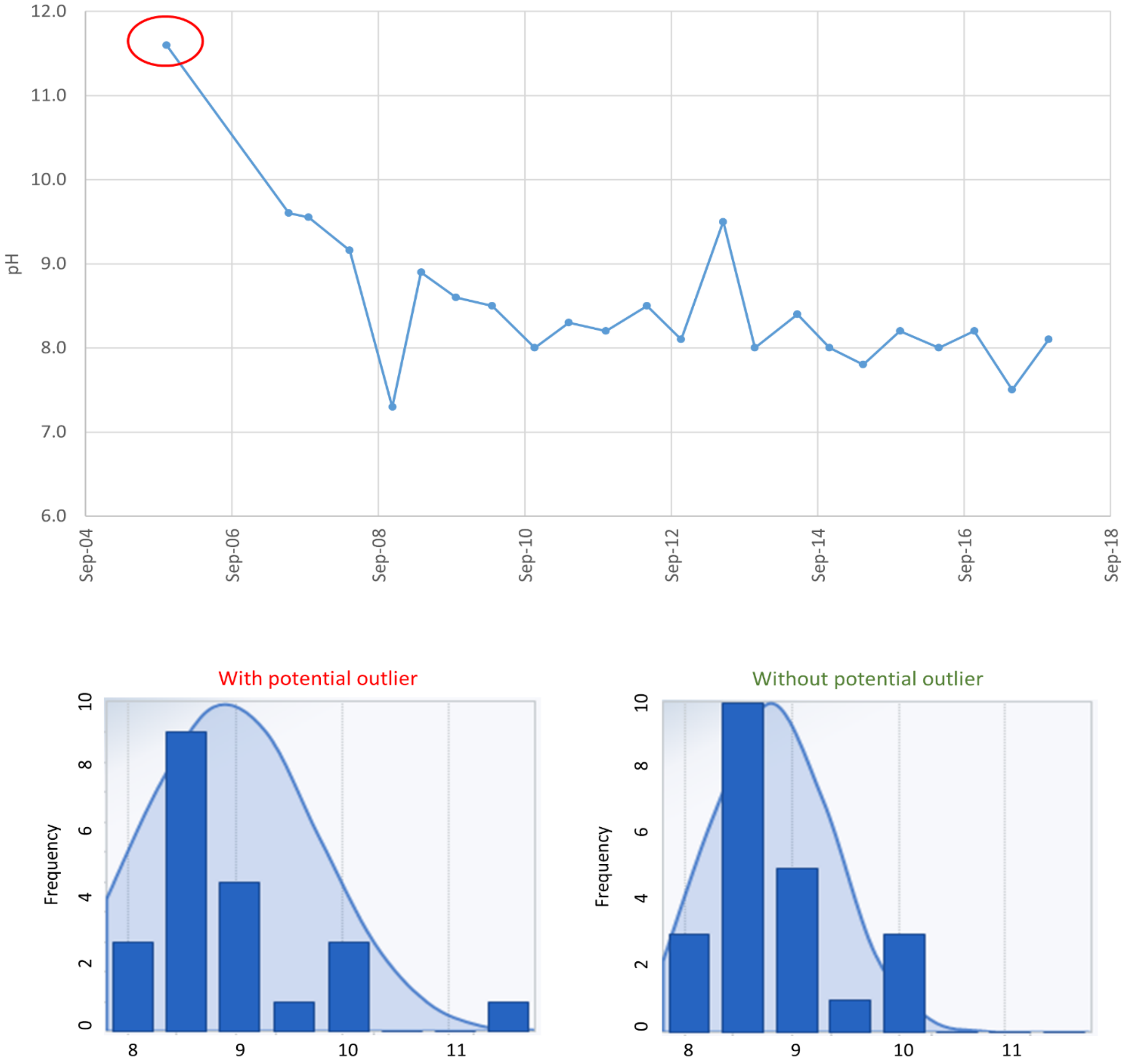


Figure 8 **Time series graph showing pH measurements for a single well between 2005 and 2017. The value highlighted in red is identified as a potential outlier via graphical analysis and also using a formal statistical outlier test. The histograms below the graph show the positive skew induced in the data set as a result of the inclusion of the potential outlier. The table above the graph demonstrates how the positive skew impacts the calculation of the mean and all calculated percentiles above the median (calculated using hydrogen ion concentration).**

4.1.6 Charge balance error

Before undertaking any statistical analysis, the charge balance error (CBE)[80] of all data sets should be calculated[81]. The CBE calculation tests the electrical neutrality of a water sample by comparing the number of positively charged ions in solution (cations) with the number of negatively charged ions (anions). Generally, a CBE threshold of ±10% is considered acceptable for groundwater samples (ALS 2022), and the application of this threshold has been successfully demonstrated for groundwater analyses in a number of studies (Guggenmos *et al* 2011, Guler *et al* 2002) where its use has ensured the retention of as much water quality data as possible by only excluding water samples with a significant charge imbalance from further analysis[82].

For the purpose of establishing baseline groundwater quality, the EPA recommends that a CBE threshold no greater than ±10% is adopted. Where a water sample has a CBE outside of the accepted threshold, further investigation should be undertaken to determine the potential cause(s)[83]. The most common cause of a CBE outside of the accepted threshold is the relative abundance of less common cations and anions in the water[84]. Where this is the case, the less common species should be included in the CBE calculation[85]. Other potential causes of the CBE calculation falling outside of the accepted threshold include the presence of organic salts in the water, and the use of ‘total’ analyses rather than ‘soluble’ (filtered) analyses[86]. Where the cause of a CBE outside of the accepted threshold cannot be identified or rectified (to within the accepted threshold), the data from that water sample should be removed from the data set before any further analysis is undertaken.

4.2 Calculating summary statistics

Once data quality has been thoroughly checked, summary statistics should be calculated. Summary statistics provide a ‘snapshot’ of a data set, giving an important insight into its centrality and spread. They also provide a useful insight into the types of graphical representations that should be examined to further interrogate the data and ensure that each target population is accurately represented (see section 4.3). It is important to note that summary statistics will need to be re–calculated following an examination of graphical representations of the data, as part of the iterative EDA process (see section 4.3).

Summary statistics should be calculated and reported for all relevant target populations associated with a baseline groundwater quality data set.

Any report of summary statistics should include basic descriptive parameters, such as the number of observations, the date range over which the observations have been collected, and the minimum and maximum values reported.

The centrality of a data set can be measured by calculating the mean and median. The median, which is equivalent to the 50th percentile, is the value that falls directly in the middle of the data when the measurements are ranked in order from smallest to largest (US EPA 2006a). The median can be used with

---

[80] Or ionic balance.

[81] CBE (%) = {[Cations (meq/L)] – [Anions (meq/L)] / [Cations (meq/L)] + [Anions (meq/L)} x 100

[82] A significant charge imbalance may be indicative of a sampling or analysis error.

[83] Hill Labs 2023 provides further information about factors which can affect the CBE.

[84] For example, metals, fluoride, nitrate/ammonia.

[85] In this situation it is also important to recognise that major ion plots, such as a Piper plot (see section 4.3.5), may not adequately represent the groundwater composition.

[86] The CBE calculation only considers anions and cations that are naturally in solution.

censored data and is not influenced by extreme values. Conversely, the sample mean, which is an arithmetic average, is influenced by censored data and extreme values (US EPA 2000). Despite this, it may still be useful to calculate the mean for data sets with non-detects and extreme values, as it can provide important information about the shape and spread of the data when compared to the median value.

The spread of a data set can be measured by calculating the standard deviation. Standard deviation measures the dispersion of data from the mean; with a small standard deviation indicating a cluster of data around the mean (US EPA 2000). Dividing the standard deviation by the mean of the data set provides a 'relative standard deviation' that can be used to compare dispersion across data sets. Standard deviation can be affected by a large number of non-detects or extreme values. Where this is an issue, the interquartile range[87] provides an alternative measure of spread of the data set that is not significantly influenced by a large number of non-detects or extreme values.

Percentiles provide a measure of the relative standing between observations in a data set, representing the value below which a certain percentage of observations will fall. Percentiles are commonly used for water quality assessment purposes. The specific percentiles that should be included in summary statistics to establish baseline groundwater quality will depend on site–specific factors, as well as the data quality objectives. The 50th percentile, or median, is typically useful for all data sets, and the 25th and 75th percentiles will be required if an interquartile range needs to be calculated. Other percentiles, such as the 80th and 90th percentiles may be used to establish a baseline groundwater quality data set and the subsequent establishment of assessment criteria for the identification of pollution associated with a regulated activity[88]. The 95th and 99th percentiles may occasionally be used for a risk–based assessment of groundwater pollution where a baseline cannot be reliably established.

When reporting summary statistics it is important to consider the precision with which the data should be reported based on the number of significant figures (accuracy) of values in the data set. Summary statistics should not be reported with a greater number of significant figures than the original data.

Table 1 provides an example of reported summary statistics for three target populations.

---

[87] The interquartile range is a measure of the difference between the 25th and 75th percentiles (or first and third quartiles) of the data set.

[88] Refer to Guideline 4 – Establishing groundwater quality assessment criteria (to be published).

Table 1 **Summary statistics for ammonia (as N) concentrations from 1992 to 2022 for three separate target populations. A wealth of information about each target population can be obtained from this information. For example, the mean for target population 1 is larger than the calculated median value. This indicates that the data set for the target population is right skewed. Based on the large maximum detect and non–detect values, as well as the incremental increase in value from the 75th to the 99th percentiles, this is likely due to a few values that are significantly higher than the rest of the values for the target population. This is also supported by the large value for the interquartile range[89]. For target population 2, the mean value is larger than the median, indicating that it is also right skewed. However, this time it is likely due to a single large value, based on the maximum detect value, and the large difference between the 95th and 99th percentiles[90]. For target population 3, the common value shared by the 20th, 25th, 75th and 80th percentiles, as well as the value of 0 for the interquartile range, indicates that most of the values are clustered at the lower end of the range. This is supported by the large percentage of non–detect values. However, the mean is slightly smaller than the calculated median value, indicating that the data for the target population is slightly left skewed. This is likely to be as a result of the majority of non–detects being reported to the maximum non–detect value (10), and a small number of non–detects being reported to the minimum non–detect value (0.01)[91].**

	Summary statistics for ammonia (as N) concentrations (in mg/L) taken from wells between 1992 to 2022		
	Target population 1	**Target population 2**	**Target population 3**
# Observations	49	30	22
# Detects	29	6	4
Minimum detect	0.17	0.34	20
Maximum detect	330	343	80
# Non– detects	20	24	18
Minimum non–detect	0.01	0.01	0.01
Maximum non–detect	200	50	10
% Non–detects	40	80	81
Mean	50	19	8.2

---

[89] This could also be visually displayed through a graphical representation of the data (eg histogram).

[90] The potential for this value to be an outlier should be investigated through graphical examination (eg Q−Q plot).

[91] This could also be visually displayed through a graphical representation of the data (eg boxplot).

	Summary statistics for ammonia (as N) concentrations (in mg/L) taken from wells between 1992 to 2022		
	Target population 1	**Target population 2**	**Target population 3**
Standard deviation	75	64	20
Interquartile range	40	7.5	0
10th percentile	10	9	1
20th percentile	10	10	10
25th percentile	10	10	10
50th percentile (median)	20	10	10
75th percentile	50	18	10
80th percentile	110	38	10
90th percentile	150	52	29
95th percentile	230	87	49
99th percentile	290	270	77

4.3 Examining graphical representations of data

Graphical methods are a key component of the EDA process and provide a quick visual summary of essential data characteristics (ITRC 2013). They can be used to identify patterns, trends and anomalies in data for target population(s) that may be more difficult to identify using purely numerical methods. Any patterns, trends or anomalies identified in the data can then be used to refine the target population(s) for the establishment of a baseline groundwater quality data set.

All data intended to be utilised to establish a baseline groundwater quality data set should be subject to interrogation using graphical methods to identify any patterns, trends or anomalies[92].

---

[92] Prior to the recalculation of summary statistics (see section 4.4).

Graphs also provide a valuable tool to quickly convey an immediate understanding of the important characteristics of the data to relevant stakeholders, including regulators (US EPA 2006a). This may be of particular use for compliance reporting[93].

In some instances, data which is not found to be suitable for statistical analysis within a target population, may still be able to be included in graphical representations in order to convey important visual characteristics of the data. For example, historical data obtained using one sampling methodology may not be statistically comparable to modern data obtained using an alternative methodology, but it can still be graphed to explore, for example, seasonal trends in the data set.

A range of simple graphical data analysis methods that can be used to look at the relationships between two variables, are outlined below. For more complex data sets, it may also be appropriate to consider the use of multivariate data analysis and visualisation techniques, which look at the relationships and correlations between multiple variables[94].

4.3.1 Histogram

A histogram is a graphical representation of data points organised into user–specified ‘bins’, which depict the frequency of data values within the range specified by each bin. Histograms provide a visual technique to identify the underlying shape, structure and skew of the data, as well as indicating whether or not the data are representative of a single target population or potentially multiple target populations (modality)[95]. The visual representation of data in a histogram is sensitive to the choice of bins. Too few (large) bins, or too many (small) bins, can impact the ability of the histogram to represent visually discernible patterns. An example is provided in Figure 9. While the optimal number of bins is unique for each data set and the features being investigated, a simple starting point is to use the square root of the sample size[96].

It is important to note that some of the detail associated with a data set is lost when plotting a histogram (due to the data being placed in ‘bins’). If a higher degree of visual detail is required from the data set, a stem and leaf plot can be used. Stem and leaf plots can sometimes be more informative than a histogram, particularly for small data sets (US EPA 2006b)

---

[93] Refer to Guideline 6 – Groundwater quality assessment reporting (to be published).

[94] Mehdi *et al* (2023) and Chai *et al* (2020) provide examples of the use of multivariate statistical techniques to assess groundwater quality.

[95] Modality could also be an indicator of poor monitoring well network design.

[96] For example, 36 data points would adopt six bins.

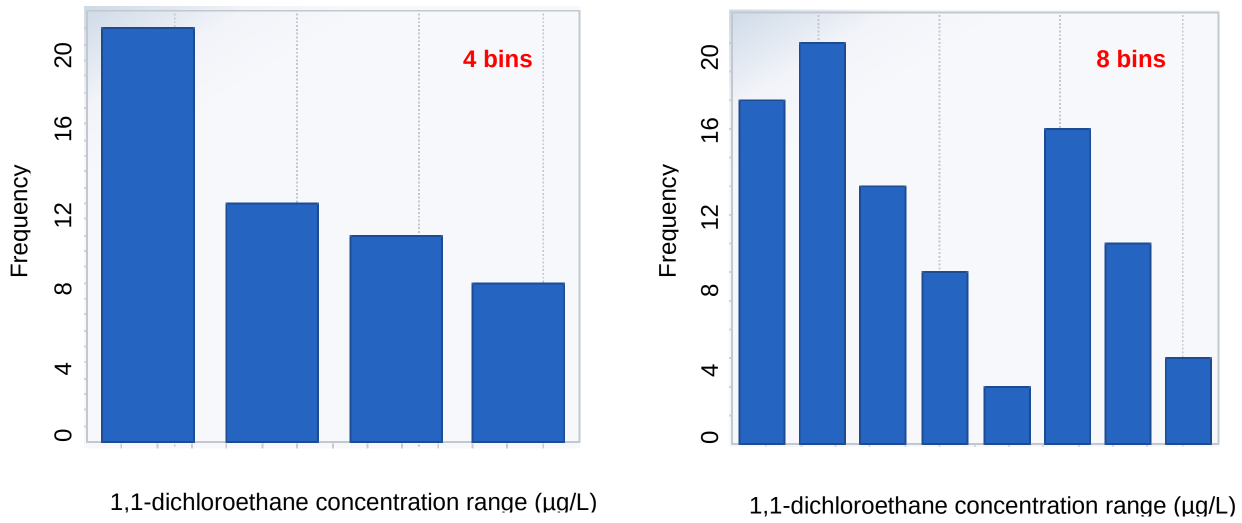


Figure 9 Histograms of 1,1–dichloroethane concentrations in groundwater based on 64 data points from an inferred single target population (based on the initial hydrogeological conceptualisation). The target population consists of data points from 8 wells across a site from 2017 to 2022. Both histograms represent the same data set, with a difference in visual appearance attributable to the choice of bin size. The histogram on the left (a) displays the data with 4 bins; the histogram on the right (b) displays the data with 8 bins (the square root of the sample size). The potential bimodal data distribution discernible in the histogram with 8 bins warrants further investigation to determine if there is a second target population included in the data set, potentially associated with a single well (or wells) or a temporal variation across the site.

4.3.2 Quantile-quantile (Q–Q) plot

Quantile–quantile (or Q–Q) plots can be used to examine the fit of data to a theoretical distribution (ITRC 2013)[97]. For example, to determine if data are normally distributed, all observations can be plotted against theoretical quantiles for a normal distribution, with the plot then examined to see how well it fits a straight line. An example is provided in Figure 10. If the data are normally distributed, all observations should lie approximately on a straight line. Potential outliers can typically be identified as single observations which do not fit the distribution of the rest of the data. It is important to note that the use of Q–Q plots to test for data distribution is qualitative only, and further testing should be undertaken to confirm any indicative data distribution[98]. Goodness–of–fit tests (eg Shapiro and Wilk) are available for use in most statistical software packages, including ProUCL.

---

[97] Further information on constructing Q-Q plots can be found in US EPA (2006b) or ITRC (2013).

[98] While most groundwater data sets are anticipated to be nonparametric, it is important to quantitatively determine whether or not a data set follows a defined population distribution. The ability to apply parametric tests to a data set conveys a higher statistical power, a lower false-positive error rate, and more confident conclusions overall (IDEQ 2014).

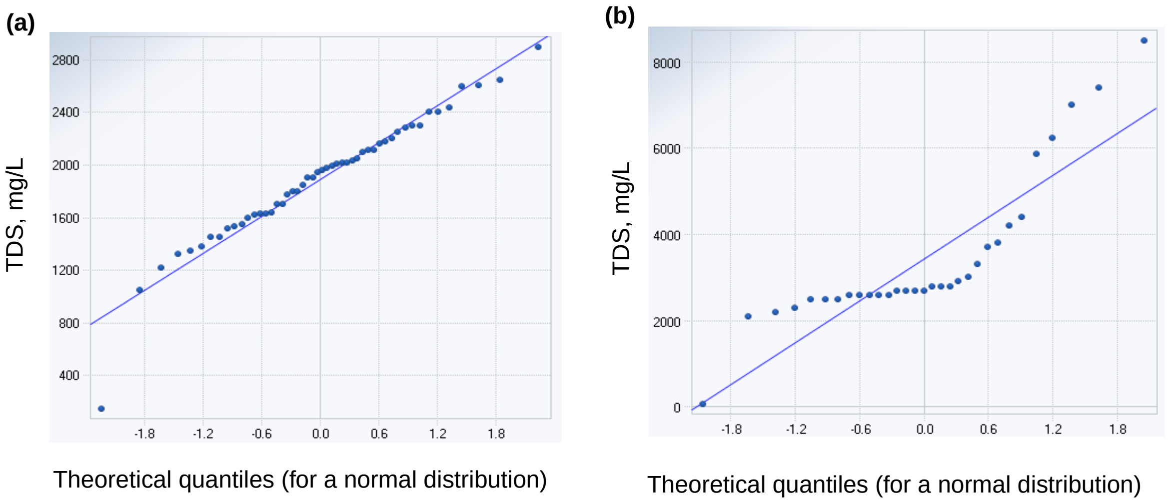


Figure 10 Quantile–quantile plots of TDS concentrations for two separate wells at a site from 1992 to 2022. The Q–Q plot on the left (a) demonstrates that the data is potentially normally distributed (although this should be confirmed statistically)[99]. The data point highlighted in red is an outlier that requires further investigation[100]. The Q–Q plot on the right (b) provides simple visual confirmation that the data are not normally distributed.

4.3.3 Time series plot

A time series plot is simply a plot of the data over time. It can be used as a qualitative tool for visually identifying temporal trends, such as seasonal fluctuations in groundwater quality annually, as well as changes over a longer term. When examining a time series plot it is important to ensure that the data used are collected frequently enough to evaluate any temporal trends of interest. For example, annual monitoring would not be sufficient to identify seasonal trends, but it may be sufficient to identify an ongoing upwards or downwards trend (ITRC 2013). The vertical axis of the time series plot can also be adjusted to assist in the visual identification of particular trends. It should also be noted that the use of a logarithmic scale on the vertical axis can potentially mask the identification of a trend that can be visually identified on a linear plot.

When constructing a time series plot it is important that censored data are not omitted. Instead, a value should be assigned to each non-detect data point (eg the RL). It is recommended that different symbols are used to depict censored data on the graph such that they are distinct from the other values. This will make it easier to visually identify any artificial trends in the data due to changing reporting limits. An example is provided in Figure 11.

---

[99] If confirmed as normally distributed then parametric statistical tests may be able to be applied to the data.

[100] Potentially including a formal outlier test.

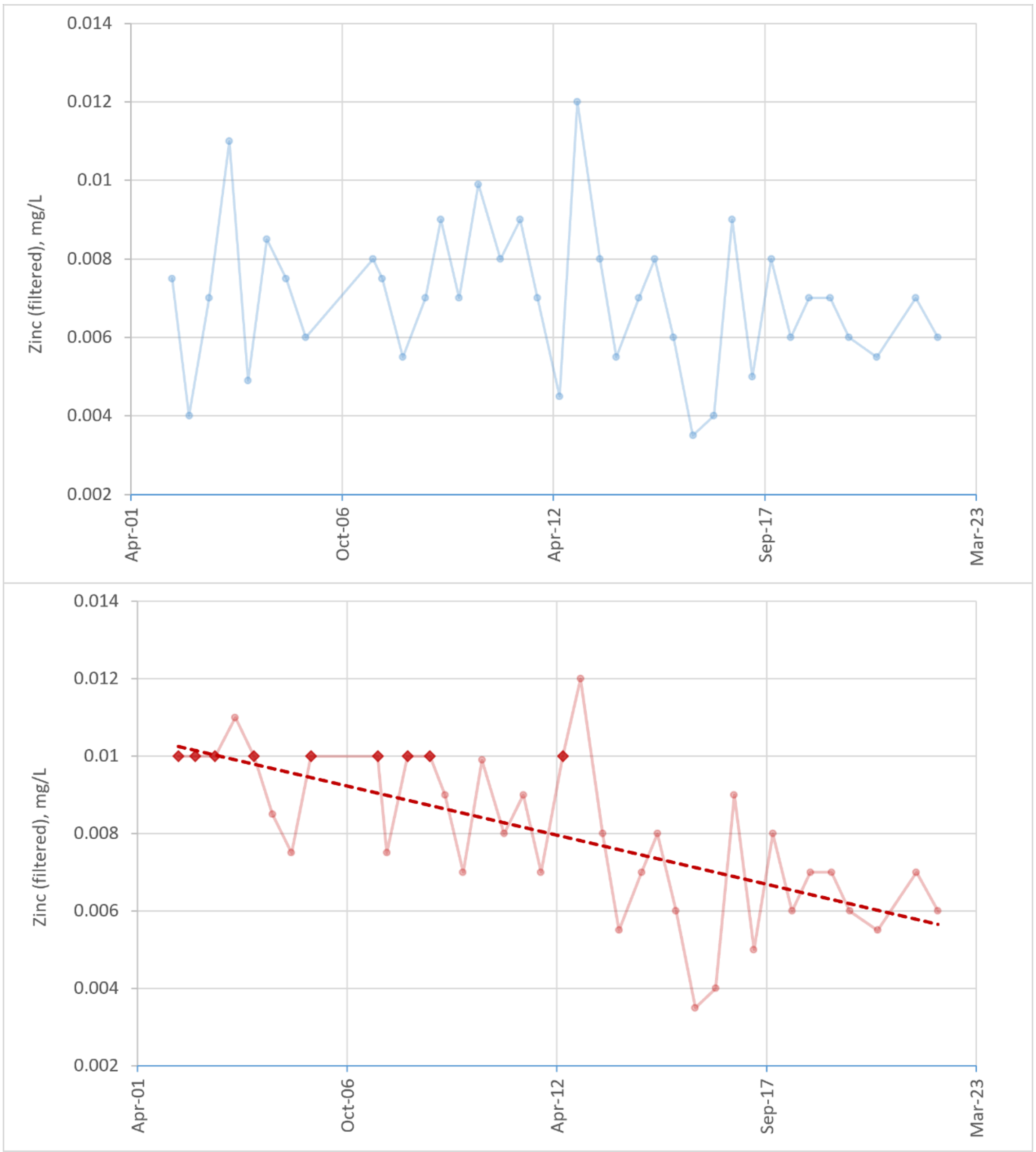


Figure 11 Time series plot (a) represents zinc concentrations for a single well from 2001 to 2023. There is no visually discernible trend in the data. Time series plot (b) represents the same data set but includes censored data (reported as RL = 0.01 mg/L) which constrains the level of information included in the first half of the plot. The dashed linear trendline demonstrates that the inclusion of censored data (reported as RL = 0.01 mg/L) induces a visually discernible trend in the data. However, when the censored data is highlighted (bold diamonds), it is implicated as a potential artificial cause of the trend. This example demonstrates the importance of adequately managing and visually identifying censored data (non–detects) when assessing for trends.

It is important to note that time series plots should only be used as a visual (qualitative) tool to identify potential trends. Formal verification (or otherwise) of a potential trend should always be undertaken through quantitative testing (eg Mann–Kendall trend test).

Seasonal trends in a data set should be assessed and managed before evaluating for monotonic trends. Seasonal trends may simply be identified through the use of a moving average, or through other quantitative methods that can be used to detect for seasonal trends in both parametric and non-parametric data sets. One-way ANOVA is a quantitative method that can be used for parametric data sets, and the Kruskal–Wallis[101] test can be used for non–parametric data sets[102].

There are also quantitative methods for the detection of monotonic trends for parametric and non–parametric data sets. For parametric data sets, a simple linear[103] regression analysis[104] can be used to generate a model ('line of best fit') of the data[105], with a corresponding R squared (R^2) value to describe how well the model fits the data[106]. A regression line with a positive slope indicates an increasing trend, and a negative slope indicates a decreasing trend. A t–test can be used to measure if the slope of a regression line is statistically different from zero[107]. When using this method to verify (or otherwise) a potential trend, care should be taken to ensure conclusions are not made based on a linear regression model with a poor fit to the data, or where outliers are present[108].

For non–parametric data sets, which will include most groundwater quality data sets, the Mann–Kendall trend test can be used to identify (or otherwise) a trend. This test is based on a calculation of the association between two variables (eg analyte concentration and date)[109]. The computation assumes that observations are independent, so it is important to ensure the data has no serial correlation before undertaking the analysis (see section 4.1.2). Where a monotonic trend is identified, the Theil–Sen estimation (or Sen's Slope) can be used to calculate the slope of the regression line through the data points[110]. In contrast to a simple linear regression model which uses a weighted mean ('least squares') to estimate the slope, the Sen's Slope is estimated using median values, which means it is not as impacted by outliers[111].

4.3.4 Box plot

A box plot, also referred to as a box and whisker plot, is a standardised method to graphically represent the locality and distribution of data for a single variable using quartiles (Q). It typically divides the data into four sections, each containing 25% of the data. A box plot consists of two parts, a box and a set of 'whiskers'. The box is bounded by Q1 (25th percentile) and Q3 (75th percentile) with a horizontal line in the middle to represent the median or Q2 (50th percentile). The whiskers represent user–defined data extremes, which are typically 1.5 times the interquartile range (IQR)[112]. When constructing box plots it is important that censored data are not omitted. Instead, a value should be assigned to each non–detect data point (eg half the RL).

Several box plots can be added to a single graph making it an excellent visual tool to compare spatial variability between data sets. For example, box plots representing baseline data from different wells at a site

---

[101] The Kruskal-Wallis test is also sometimes referred to as 'one-way ANOVA on ranks'.

[102] Both of these tests are available within the free ProUCL software statistical package.

[103] Non-linear regression analysis techniques are also available for data sets without a linear relationship.

[104] A simple linear regression analysis can be undertaken in Microsoft Excel.

[105] The data model is represented by the equation $y = ax + b$, where y is the analyte concentration, x is the time, b is the y-axis intercept, and a is the slope of the line.

[106] R^2 values range from 0 to 1, with 0 representing a very poor model, and 1 representing a perfect model.

[107] The t-test is available within the free ProUCL software statistical package.

[108] Linear regression is heavily influenced by outliers (IDEQ 2014).

[109] Kendall's tau (τ) measure.

[110] Further information about the technique, as well as additional reference material, can be found in ITRC (2013).

[111] Both the Theil-Sen and Mann-Kendall trend tests are available within the free ProUCL software statistical package. It is also noted that GSI Environmental provides a free software toolkit for Mann−Kendall analysis, which is available at https://www.gsienv.com/product/gsi-mann-kendall-toolkit/.

[112] The IQR is calculated as the difference between Q3 and Q2.

can be compared on one graph to identify whether they can potentially be pooled into a single spatial data set.

An example is provided in Figure 12. It is important to note that box plots only provide a visual (qualitative) tool to compare spatial variability between data sets. Quantitative testing should always be undertaken to justify the pooling of any data sets[113].

Box plots are also a useful tool for the identification of potential outliers, although it is important to note that it is only a qualitative method. Typically, data values that extend beyond 1.5 times the IQR are considered potential outliers and should be assessed using a quantitative (statistical) method.

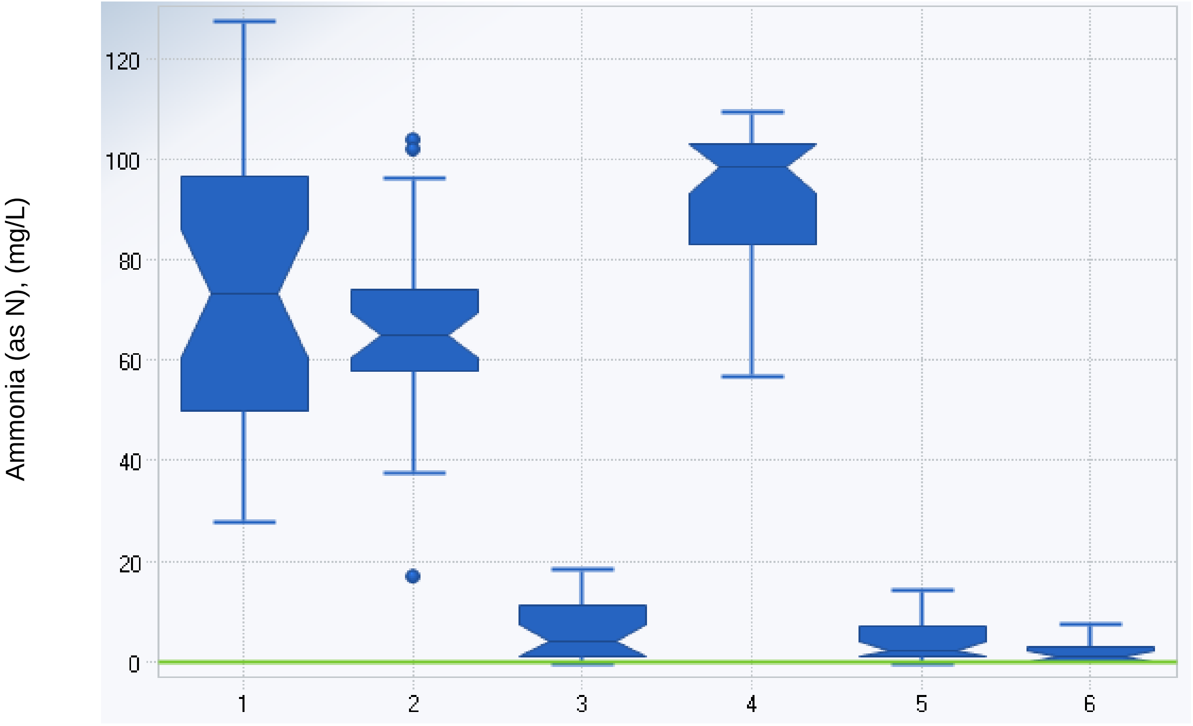


Figure 12 Graph of box plots for six different wells at a site showing ammonia (as N) concentrations collected between 2015 and 2022. The green line represents the maximum reporting limit. The wells on the graph appear to represent two (or three) spatially distinct populations of ammonia (as N) in groundwater at the site. Based on a visual inspection of the graph, the ammonia (as N) data for wells 3, 5 and 6 are likely to be suitable for pooling to represent one common target population. Wells 1, 2 and 4 may also be suitable for pooling, although it would be prudent to undertake another level of assessment of the data to ensure that well 4 does not represent a spatially distinct population of ammonia (as N) compared to wells 1 and 2. The data points for well 2 highlighted in red represent potential outliers that require further investigation.

---

[113] The ProUCL software statistical package includes a number of two sample hypothesis tests to compare data sets, including the Wilcoxon-Mann-Whitney test, which is suitable for use with nonparametric data sets containing non-detects.

4.3.5 Piper plot

A Piper plot (or trilinear diagram) is a graphical representation of the relative abundance of common (or major) ions in a water sample. It comprises three components; a ternary diagram in the lower left representing cations (calcium, magnesium, sodium + potassium), a ternary diagram in the lower right representing anions (chloride, sulphate, carbonate + bicarbonate), and a diamond plot in the middle which is a matrix transformation of the two ternary diagrams. The location of a water sample on the diamond plot identifies its hydrochemical facies. As multiple water samples can be plotted on the same diagram, Piper plots provide a powerful visual tool to compare the chemistry of water samples between sampling locations or target populations, and also to identify changes in chemistry over time at a single location or within a target population. Examples Piper plots are provided in figure 13.

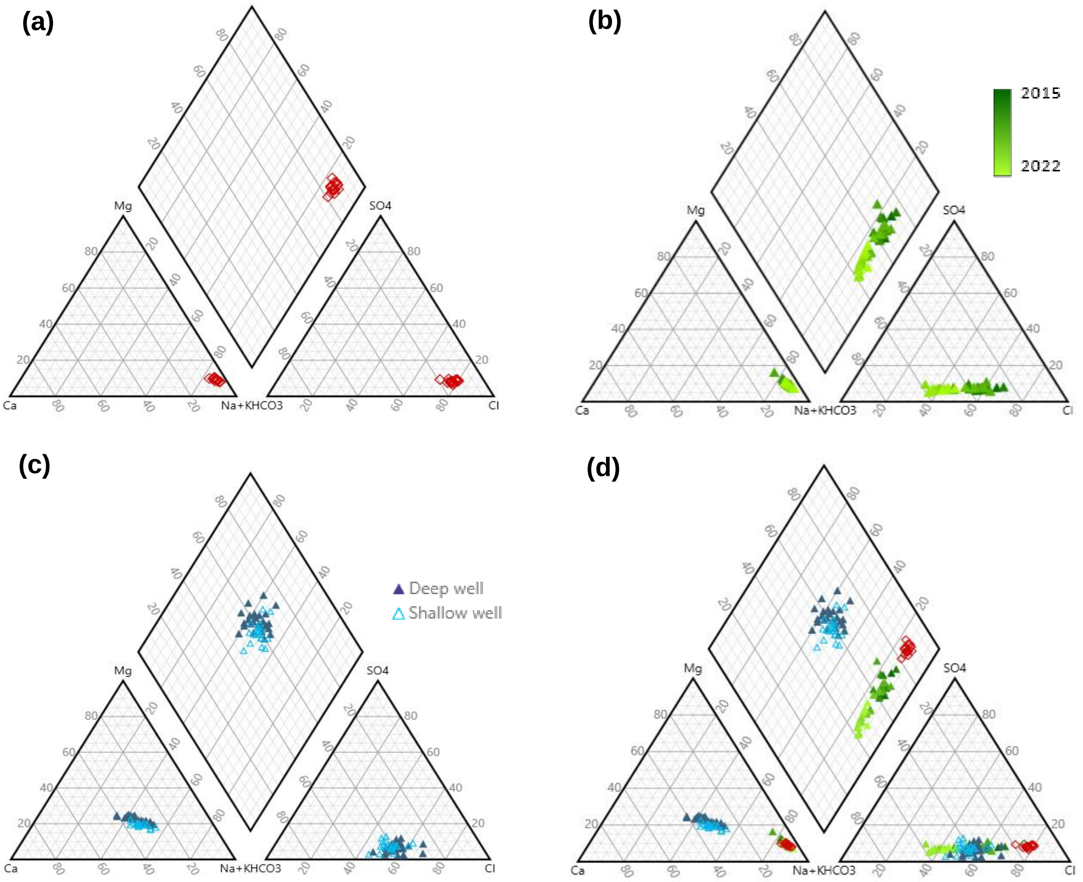


Figure 13 Piper plots for four different wells at a site which provide important information about groundwater chemistry between 2015 and 2022. (a) A groundwater well with very consistent major ion composition over time, representing sodium chloride type water. (b) A groundwater well with changing major ion composition over time, initially representing sodium chloride type water, with a gradually increasing bicarbonate component (and decreasing chloride component). (c) A deep and shallow well intersecting different aquifers at the same location. The similarity in ion composition indicates the potential for hydraulic connection between the aquifers (and data pooling into a single target population). (d) All four wells on the same Piper plot, highlighting the spatial heterogeneity in groundwater at the site. The well represented in green initially has a similar major ion composition to the well represented in red, but it becomes increasingly different over time. This may be indicative of an issue associated with sample collection and analysis for the well represented in green, or it may indicate a potential change in groundwater conditions (eg change in recharge in the vicinity of the well).

It should be noted that there are a number of other tools available for visualising major ion data. For example, Schoeller diagrams and Stiff diagrams provide alternative methods for visualising major ion data and can be used to identify spatial differences in a data set. Durov diagrams provide another method for visualising major ion data, and they also facilitate an assessment of the data with two other groundwater parameters, pH and TDS, which makes them particularly useful for more complex assessments.

It is important to recognise that all of these tools (Piper plot, Schoeller diagram, Stiff diagram, Durov diagram) only represent major ions present in a water sample. Where less common cations or anions are present in a water sample in relative abundance, these tools may not adequately represent the groundwater composition (see section 4.1.6).

4.4 Recalculating summary statistics

To finalise the establishment of a baseline groundwater quality data set, the summary statistics for each target population associated with the site should be recalculated in accordance with section 4.2, incorporating any changes to the data or target population(s) that were identified (and verified) during the EDA process. For example, this may include the removal of outliers verified through a formal test; the splitting of a single data set into multiple target populations due to temporal and/or spatial heterogeneity; or the combining of one or more sets of data into a single target population, typically due to spatial homogeneity.

It is also vital that the effects of any seasonal trends identified through the EDA process are removed from the data for each target population(s) *prior* to recalculating summary statistics. Failure to account for seasonal variability may lead to statistical conclusions that are inaccurate (IDEQ 2014).

5 Reporting baseline groundwater quality

EPA licence holders who are required to undertake groundwater monitoring as a condition of their licence will be required to do so through the implementation of a GMMP. This includes formal reporting of the established baseline groundwater quality data set. Guideline 5 – Developing a groundwater monitoring and management plan (to be published) will provide further details in relation to the development of a GMMP[114].

For activities subject to groundwater monitoring requirements under different legislation, the process to formally report baseline groundwater quality may be different. Regardless of the mechanism for reporting baseline groundwater quality, it is important to note that planning for the establishment of a baseline groundwater quality data set should begin at the project scoping phase, including engagement with the relevant regulator.

For EPA licence holders who are required to undertake groundwater monitoring as a condition of their licence, the established baseline groundwater quality data set should be reported in a GMMP associated with the licensed activity.

For activities subject to groundwater monitoring requirements under different legislation, further advice about the process to report the established baseline groundwater quality data set should be sought from the relevant authority.

The main body of the GMMP[115] should include:

1. a clearly documented hydrogeological conceptualisation (see Appendix A)
2. presentation of a summary of the baseline groundwater quality data for each target population associated with the site (see Appendix B).

An example of a baseline groundwater quality summary table is provided in Appendix B. The table should include key statistics of the data set for each target population associated with the site, clearly delineating the associated temporal and spatial conditions for the application of each target population. A brief summary of the key characteristics of the data associated with the established baseline groundwater quality, which have been identified through a hydrogeological conceptualisation and the EDA process, should also be included with the table[116].

It might also be appropriate for the baseline summary table to be included in other documents that refer to baseline groundwater quality for regulatory compliance[117].

While the summary table provides the key statistical information about baseline groundwater quality, particularly with regards to the development of assessment criteria[118], it is important that the process to

---

[114] Key elements to be included in a GMMP include the established baseline groundwater quality data set, site-specific assessment criteria established from the baseline groundwater quality data set, and clear actions to be undertaken if site-specific assessment criteria are exceeded.

[115] Or other document used to formally report the established baseline groundwater quality data set.

[116] Including clearly identifying target populations that represent groundwater impacted by specific sources of pollution.

[117] For example, an annual groundwater monitoring compliance report or environmental management plan.

[118] Refer to Guideline 4 – Establishing groundwater quality assessment criteria (to be published).

establish the baseline groundwater quality is also clearly reported in an appendix to the GMMP[119]. This should include:

1. a summary of the monitoring network and data collection methodology
2. presentation of the data quality checks performed, and methods used to manage any identified data quality issues
3. a tabular description of the summary statistics calculated for each target population following the data quality checks
4. presentation of graphical analyses performed to visually identify data characteristics
5. a summary of any quantitative analyses performed to formally verify the visually identified data characteristics
6. a tabular description of the summary statistics recalculated for each target population following completion of the EDA process
7. a digital spreadsheet of the raw water quality data used to establish baseline groundwater quality.

A checklist outlining the key information to include in any baseline groundwater quality reporting is included in Appendix C. Any data gaps should be clearly documented, along with a commentary of the impacts the data gaps may have on the establishment of baseline groundwater quality. Any actions intended to be undertaken to address the data gaps in the future (if required) should also be documented. As discussed in sections 3.2 and 3.3, this may include the establishment of an interim baseline groundwater quality while additional groundwater quality data is being collected. Where this is the case, the baseline groundwater quality should clearly be reported as interim.

It is important to note that the EPA expects that the reported baseline groundwater quality will be updated periodically (or as required) as new data becomes available[120]. As a minimum, it is expected that the baseline groundwater quality will be reviewed and updated (if required) every five years. For EPA licence holders, this review and update (if required) can simply be incorporated into a periodical review and update of the GMMP. This may include the addition of newly identified POCs[121], particularly for activities that operate for an extended period of time. For activities subject to groundwater monitoring requirements under different legislation, it is recommended that advice is sought from the relevant authority regarding the process to review and update (if required) the baseline groundwater quality.

Where an interim baseline has been established and used to develop site–specific assessment criteria while additional baseline groundwater quality data is being collected, the reported baseline groundwater quality data set should be updated at least annually to incorporate any new data and update the site–specific assessment criteria (where required)[122].

---

[119] Or other document used to formally report the established baseline groundwater quality data set.

[120] For example, new baseline groundwater quality data (from existing or new monitoring wells) or a change in climate conditions.

[121] For example, PFAS or other pollutants of emerging concern.

[122] Refer to Guideline 6 – Groundwater quality assessment reporting (to be published).

References

ALS 2022, *Ion Balance: Data Quality Validation for Metals & Anions in Waters*, ALS Limited. Accessed from https://www.alsglobal.com/en/Search#sort=relevancy.

ANZG 2018, *Australian and New Zealand Guidelines for Fresh and Marine Water Quality*, Australian and New Zealand Governments and Australian State and Territory Governments, Canberra, ACT, Australia, accessed from www.waterquality.gov.au/anz–guidelines.

Barcelona MJ, Lettenmaier DP & Schock MR 1989a, ‘Network design factors for assessing temporal variability in ground–water quality’, *Environmental Monitoring and Assessment* 12: 149–179.

Barcelona MJ, Wehrmann HA, Schock MR, Sievers ME & Karny JR 1989b, *Sampling Frequency for Ground–Water Quality Monitoring*, for the Illinois Department of Energy and Natural Resources, US Environmental Protection Agency, Las Vegas.

Bondu R, Humez P, Mayer B, Chaste E, Naumenko-Dèzes MO, Cloutier V, Rosa E & Kloppmann W 2022, ‘Estimating natural background concentrations for dissolved constituents in groundwater: A methodological review and case studies for geogenic fluoride’, *Journal of Geochemical Exploration* 233: 106906.

Bondu R, Kloppmann W, Naumenko–Dèzes MO, Humez P & Mayer B 2021, ‘Potential Impacts of Shale Gas Development on Inorganic Groundwater Chemistry: Implications for Environmental Baseline Assessment in Shallow Aquifers’, *Environmental Science and Technology* 55: 9657–9671.

Chai Y, Xiao C, Li M, Liang X 2020, ‘Hydrogeochemical Characteristics and Groundwater Quality Evaluation Based on Multivariate Statistical Analysis’, *Water* 12: 2792

DES 2021, *Using monitoring data to assess groundwater quality and potential environmental impacts*, *Version 2*, Department of Environment and Science, Queensland Government, Brisbane.

DMP & DW 2016, *Guideline for groundwater monitoring in the onshore petroleum geothermal industry*, Department of Mines and Petroleum & Department of Water, Government of Western Australia, Perth.

DPI 2014, *Groundwater Monitoring and Modelling Plans – Information for prospective mining and petroleum exploration activities*, Office of Water, Department of Primary Industries, NSW Government, Sydney.

DWR 2016, *Best Management Practices for the Sustainable Management of Groundwater: Hydrogeologic Conceptual Model*, Department of Water Resources, State of California, Sacramento.

Gemson WJ, Weaver TR & Heemink BL 2019, ‘Key considerations that can make or break a closure-focused groundwater-monitoring program’, in AB Fourie & M Tibbett (eds), *Mine Closure 2019: Proceedings of the 13th International Conference on Mine Closure*, Australian Centre for Geomechanics, Perth, pp 1367–1376.

Goudey R 2007, ‘Do statistical inferences allowing three alternative decisions give better feedback for environmentally precautionary decision–making’, *Journal of Environmental Management* 85: 338–344.

Guggenmos M, Daughney C, Jackson B & Morgenstern U 2011, ‘Regional–scale identification of groundwater–surface water interaction using hydrochemistry and multivariate statistical methods, Wairarapa Valley, New Zealand’, *Hydrology and Earth System Sciences* 15: 3383–3398.

Guler C, Thyne GD, McCray JE & Turner AK 2002, ‘Evaluation of graphical and multivariate statistical methods for classification of water chemistry data’, *Hydrogeology Journal* 10: 455 – 474.

Helsel D, Hirsch R, Ryberg K, Archfield S & Gilroy E 2020, *Statistical Methods in Water Resources: US Geological Survey Techniques and Methods, Book 4, Chapter 3.*, US Department of the Interior & US Geological Survey, Virginia.

Hill Labs 2023, *Anion-Cation Balances*, Version 7, Hill Labs, 13213v7_technical–note–anion–cation–balances.pdf (hill–labs.co.nz).

IDEQ 2014, *Statistical Guidance for Determining Background Ground Water Quality and Degradation*, Idaho Department of Environmental Quality, Boise.

ITRC 2013, *Groundwater Statistics and Monitoring Compliance, Statistical Tools for the Project Life Cycle*, GSMC–1, Interstate Technology & Regulatory Council, Groundwater Statistics and Monitoring Compliance Team, Washington DC.

Landscape South Australia 2023, Environmental and Cultural Flows, accessed from https://www.landscape.sa.gov.au/ny/projects/environmental-and-cultural-flows.

Mehdi M, Hakim DJ, Amina A, Nourredine G 2023, 'Multivariate Statistical Analysis of Groundwater Quality of Hassi R'mel, Algeria', *Journal of Ecological Engineering* 24(5): 22 – 31.

McCullough CD 2016, 'Key mine closure lessons still to be learned', in AB Fourie & M Tibbett (eds), *Mine Closure 2016: Proceedings of the 11th International Conference on Mine Closure*, Australian Centre for Geomechanics, Perth, pp 325–338.

NHMRC 2008, *Guidelines for Managing Risks in Recreational Water*, National Health and Medical Research Council, Commonwealth of Australia, Canberra.

NHMRC & NRMMC 2011, *Australian Drinking Water Guidelines Paper 6, National Water Quality Management Strategy*, National Health and Medical Research Council & National Resource Management Ministerial Council, Commonwealth of Australia, Canberra.

NT EPA 2017, *Northern Territory Contaminated Land Guideline*, Northern Territory Environment Protection Authority, Darwin.

NUDLC 2020, *Minimum Construction Requirements for Water Bores in Australia: Edition 4*, National Uniform Drillers Licensing Committee, Australian Drilling Industry Association, Wangara.

Peach D, Taylor A 2023, 'The development of a hydrogeological conceptual model of groundwater and surface water flows in the Silala River Basin', *WIREs Water* e1676.

SA EPA 2018, *Assessment of background concentrations*, South Australian Environment Protection Authority, Adelaide.

Shand P, Edmunds WM, Lawrence AR, Smedley PL & Burke S 2007, *The natural (baseline) quality of groundwater in England and Wales*, British Geological Survey Research Report No. RR/07/06.

Stefania GA, Zanotti C, Bonomi T, Fumagalli L, & Rotiroti M 2018, 'Determination of trigger levels for groundwater quality in landfills located in historically human-impacted areas', *Waste Management* 75: 400–406.

Sweeney J, O'Farrell C, Crosbie J & Osgerby B 2022, 'A risk-based framework for assessing non-metalliferous contamination prior to closure at MMG Rosebery Mine, Tasmania', in AB Fourie, M Tibbett, G

Boggs (eds), *Mine Closure 2022: 15th International Conference on Mine Closure, Australian Centre for Geomechanics*, Perth, pp 281–294.

US EPA 2000, *Practical Methods for Data Analysis. EPA QA/G–9 – QA00 Update*, United States Environmental Protection Agency, Washington DC.

US EPA 2006a, *Data Quality Assessment: A Reviewer's Guide – EPA QA/G–9R*, United States Environmental Protection Agency, Washington DC.

US EPA 2006b, *Data Quality Assessment: Statistical Methods for Practitioners – EPA QA/G–9S,* United States Environmental Protection Agency, Washington DC.

Weaver T, Kovacsy D, Windle D, Gemson W, Hausmann N, Shibata M, Staggs D & Iles M 2022, 'Deriving background concentrations of contaminants of potential concern in groundwater: an example from the Ranger uranium mine, Australia', in AB Fourie, M Tibbett, G Boggs (eds), *Mine Closure 2022: 15th International Conference on Mine Closure, Australian Centre for Geomechanics*, Perth, pp 691–704.

Appendix A Example of hydrogeological conceptualisations

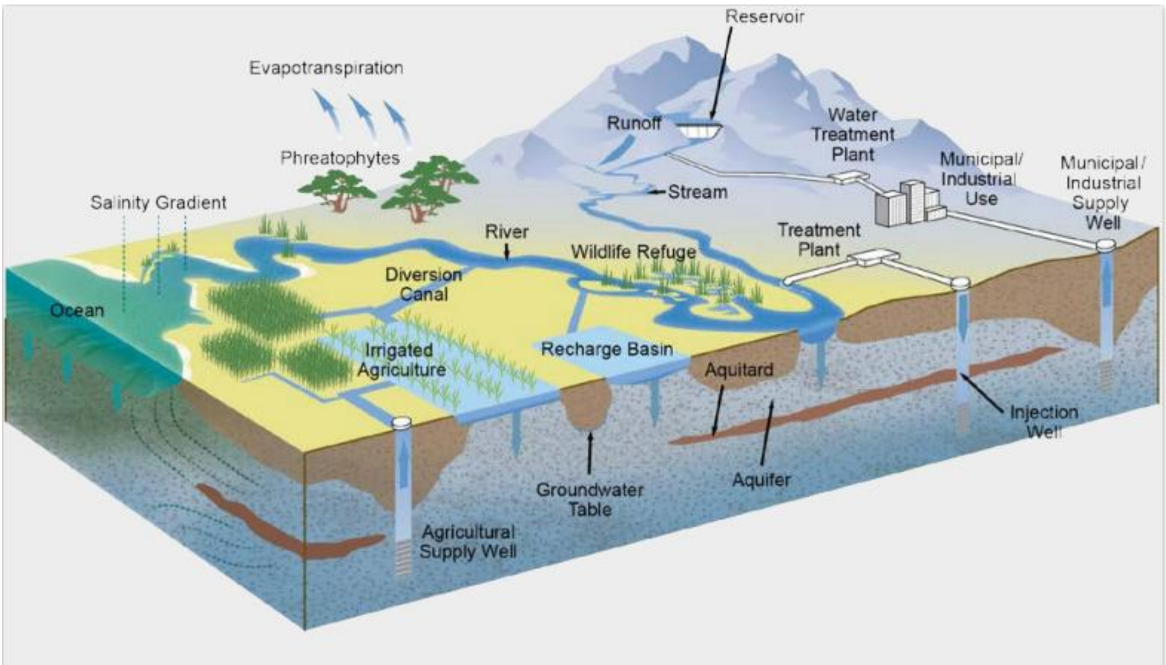


Figure 14 Example 3-D graphic conceptualisation of a groundwater system. Reprinted from DWR 2016, Best Management Practices for the Sustainable Management of Groundwater: Hydrogeologic Conceptual Model, Department of Water Resources, State of California, Sacramento.

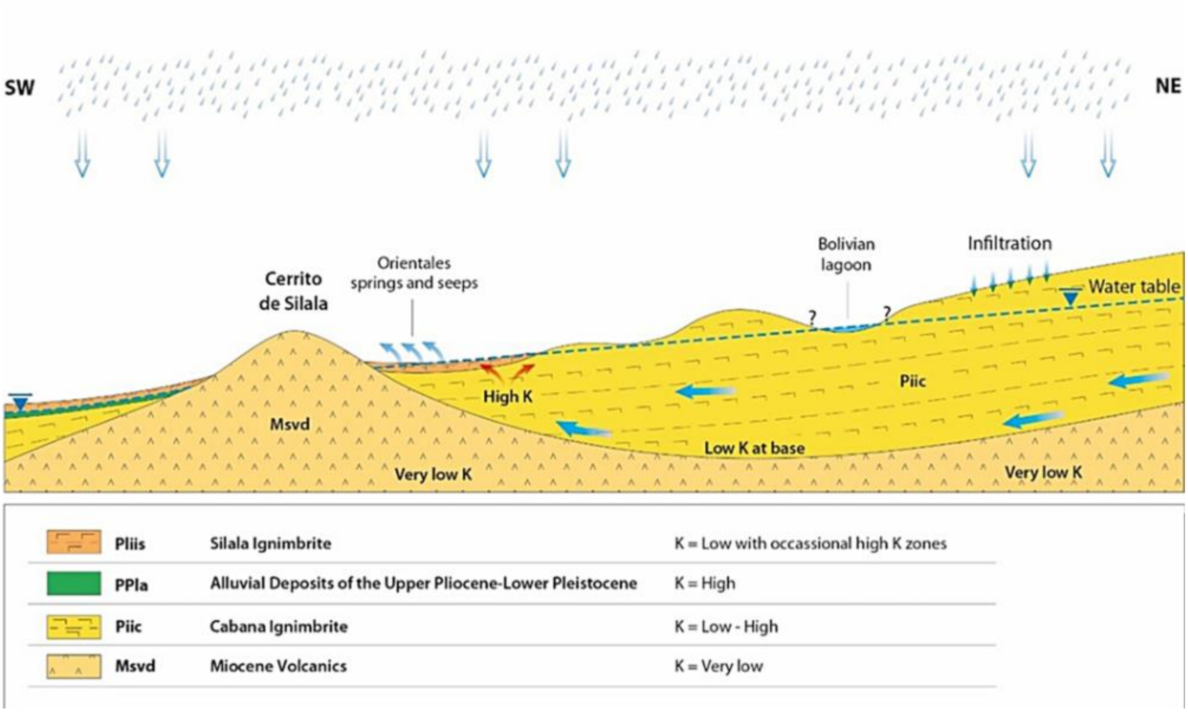


Figure 15 Example cross-section conceptualisation of a groundwater system. Reprinted from Peach D, Taylor A 2023, The development of a hydrogeological conceptual model of groundwater and surface water flows in the Silala River Basin, WIREs Water e1676.

Appendix B Example of baseline groundwater quality summary

Table 2 Example baseline groundwater quality summary. Target populations for each analyte are defined. Brief notes are included to justify the target populations defined and highlight key features of the overall baseline groundwater quality data set. A basic conceptual site layout and cross section is provided in Figure 16 to assist in understanding the tabular data.

TDS (mg/L)				
Target population	**A**	**B**	**C**	**D**
Monitoring wells	**GW 1–4, 7–9**	**GW 5 (Feb/Mar)**	**GW 5 (Aug/Sep)**	**GW 6**
# observations	140	11	10	21
% non-detect	0	0	0	0
Minimum non-detect	N/A	N/A	N/A	N/A
Maximum non-detect	N/A	N/A	N/A	N/A
Minimum	1,850	1,830	990	2,920
10th percentile	2,039	1,839	1,008	3,456
20th percentile	2,228	1,880	1,050	3,556
50th percentile (median)	2,600	1,970	1,145	3,885
80th percentile	2,962	2,066	1,232	4,496
90th percentile	3,380	2,098	1,247	4,638
Maximum	3,950	2,170	1,310	5,140

Notes:

- The conceptual understanding of the groundwater system was used to infer at least two spatially distinct target populations, but further analysis of the data using Piper plots identified three spatially distinct target populations (A, B/C, D).
- Time series plots were used to identify a seasonal trend in TDS concentrations for groundwater from GW5. This was formally verified through a Mann–Kendall test. The seasonal trend is likely due to recharge from the nearby freshwater creek during wetter months.

Ammonia (as N) (µg/L)			
Target population	**A**	**B**	**C**
Monitoring wells	**GW 1, 3**	**GW 2, 4–8**	**GW 9**
# observations	27	78	13
% non-detect	0	37	0
Minimum non-detect	N/A	0.01	N/A
Maximum non-detect	N/A	0.1	N/A
Minimum	3.1	0.012	24
10th percentile	4.6	0.01	30
20th percentile	5.1	0.01	33
50th percentile (median)	6.8	0.062	43
80th percentile	8.0	0.23	52
90th percentile	8.8	0.59	58
Maximum	11	1.6	64

Notes:

- Box plots were used to identify three spatially distinct target populations (A, B, C).
- The wells included in target population A identify ammonia pollution. As GW1 is located up hydraulic gradient of the landfill, it is reasonable to infer that the impacts identified in groundwater from the wells in target population A are associated with the adjacent agricultural site.
- Groundwater from GW9 identifies more significant ammonia pollution (greater than that identified in the up hydraulic gradient well, GW1) which can be attributed to the unlined waste cells.

PCE (ug/L)		
Target population	**A**	**B**
Monitoring wells	**GW 1–8**	**GW 9**
# observations	104	13
% non–defect	100	0
Min non–defect	1	N/A
Max non–defect	1	N/A
Minimum	N/A	7
10th percentile	1	8
20th percentile	1	9
50th percentile (median)	1	11
80th percentile	1	14
90th percentile	1	16
Maximum	N/A	17

Notes:

- The summary statistics identified PCE pollution in groundwater from GW9 as a result of the unlined waste cells.

Chromium (soluble) (mg/L)					
Target population	**A**	**B**	**C**	**D**	**E**
Monitoring wells	**GW 1, 3**	**GW 2**	**GW 4, 5**	**GW 6, 7, 8**	**GW 9**
# observations	26	13	24	40	13
% non-detect	35	0	0	100	7.7
Minimum non-detect	0.001	N/A	N/A	0.001	0.001
Maximum non-detect	0.005	N/A	N/A	0.001	0.001
Minimum	0.0011	0.032	0.018	N/A	0.008
10th percentile	0.001	0.035	0.020	0.001	0.0086
20th percentile	0.001	0.037	0.021	0.001	0.011
50th percentile (median)	0.0012	0.044	0.025	0.001	0.014
80th percentile	0.0019	0.049	0.029	0.001	0.017
90th percentile	0.0037	0.051	0.031	0.001	0.019
Maximum	0.009	0.053	0.025	N/A	0.014

Notes:

- The conceptual understanding of the groundwater system was used to infer at least three spatially distinct target populations, but further analysis of the data using Box plots and Q–Q plots identified five spatially distinct target populations (A, B, C, D, E).
- Groundwater from GW2 identifies chromium pollution. As GW2 is located up hydraulic gradient of the landfill site, it is likely that the impacts identified are associated with the adjacent former timber treatment facility. The wells in target population C also identify chromium pollution. However, as those wells are located down hydraulic gradient from GW2, and the concentrations are less than those identified in groundwater from GW2, it is reasonable to infer that the chromium pollution identified in groundwater from wells in target population C is also associated with the former timber treatment facility. This is further validated through the identification of a downwards trend in chromium data from GW3, GW4 and GW5.
- Groundwater from GW9 identifies chromium pollution. A small amount of chromium pollution is also identified in groundwater from GW1, but the concentrations are generally an order of magnitude lower. As GW1 is located up hydraulic gradient from GW9 (and the landfill), it is reasonable to infer that a

Chromium (soluble) (mg/L)
majority of the chromium pollution identified in GW9 is associated with impacts from the unlined waste cells. • Two potential outliers in target population B were identified in the Box plots and verified as outliers using Dixon's test. Subsequent assessment of the sampling and analysis records identified that both data points represented total chromium rather than soluble chromium. Both data points were removed from the data set.

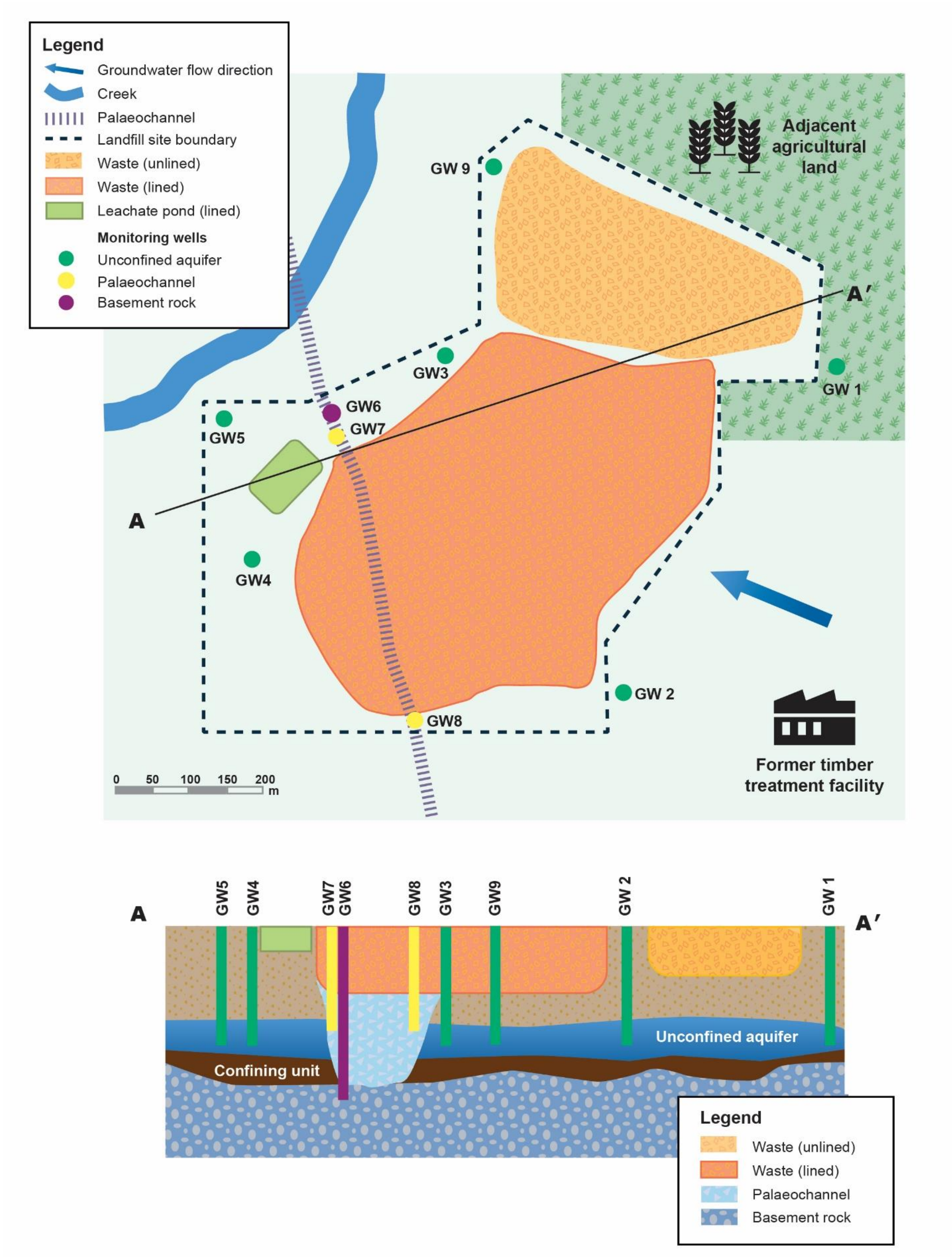


Figure 16 **Basic conceptual site layout and cross–section of landfill to assist understanding of data in Table 2.**

Appendix C Reporting checklist

Table 3 Checklist for reporting baseline groundwater quality

Description	Section	✓
Conceptual understanding of the groundwater system	**2**	
Conceptualisation of hydrostratigraphy (including hydraulic properties for each layer)	2	☐
Conceptualisation of groundwater flow (including direction and velocity)	2	☐
Conceptualisation of surface water interaction (including recharge and discharge)	2	☐
Identification of any spatial variability impacting groundwater movement	2.1	☐
Identification of any temporal variability impacting groundwater movement	2.1	☐
Key processes controlling groundwater quality, both natural and anthropogenic	2.2	☐
Identification of a clear target population(s)	2.3	☐
Data collection	**3**	
Location of all wells monitored, including depth, screened interval, and target geology	3.1	☐
Length, frequency and methodology of all sampling, including start and end dates	3.2	☐
List of all analytes sampled, with justification in the context of list of identified POCs	3.3	☐
Understanding the data	**4**	
Checking data quality	**4.1**	
Identification of any significant changes in sampling methodology	4.1.1	☐

Description	**Section**	✓
Identification of any inadequacies with sampling frequency (including correlated data)	4.1.2	☐
Identification of any changes in analytical methodology or reporting limits	4.1.3	☐
Identification and effective management of non-detect data	4.1.4	☐
Identification of any outliers, and subsequent removal (where justified)	4.1.5	☐
Calculating summary statistics	**4.2**	
Tabular description of summary statistics	4.2	☐
Examining graphical representations of data	**4.3**	
Identification and verification of data distribution for each data set (eg histograms, Q–Q plots)	4.3.1, 4.3.2	☐
Identification and verification of any outliers (eg Q–Q plots, box plots, Dixon's or Rosner's test)	4.1.5, 4.3.2, 4.3.4	☐
Identification and verification of any temporal trends (eg time series plots, Mann–Kendall test)	4.1.4, 4.3.3	☐
Identification of any spatial trends (eg box plots, Piper plots)	4.3.4, 4.3.5	☐
Re-calculating summary statistics	**4.4**	
Identification of any re-defined target population(s) (if required)	4.4	☐
Tabular description of updated summary statistics (if required)	4.4	☐
Reporting baseline groundwater quality	**5**	
Summary of baseline groundwater quality data set	5	☐

Description	Section	✓
Summary of any identified data gaps and proposed actions (eg additional data collection)	5	☐
Digital spreadsheet of raw water quality data	5	☐

A BETTER

ENVIRONMENT

FOR THE

HEALTH, WELLBEING

AND PROSPERITY

OF ALL SOUTH

AUSTRALIANS

South Australia

ENVIRONMENT PROTECTION AUTHORITY

GPO Box 2607 Adelaide SA 5001
Phone: 08 8204 2004 | Fax: 08 8124 4670
Email: epainfo@sa.gov.au
Web: epa.sa.gov.au