Photos: Lawn Hill Gorge, *Scaevola* sp. Spray Irrigation, Royal Spoonbills, Groundwater pump

Using monitoring data to assess groundwater quality and potential environmental impacts

Version 2

Queensland Government

Prepared by

Sonia Claus, Jason Dunlop and Ian Ramsay

Science and Technology Division
Department of Environment and Science
PO Box 5078
Brisbane QLD 4001

The Department of Environment and Science acknowledges Aboriginal peoples and Torres Strait Islander peoples as the Traditional Owners and custodians of the land. We recognise their connection to land, sea and community, and pay our respects to Elders past, present and emerging.
The department is committed to respecting, protecting and promoting human rights, and our obligations under the Human Rights Act 2019.

Citation

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

February 2021

Forward

The purpose of this document is to assist with the review of monitoring data to assess and manage anthropogenic activities that may impact on groundwater quality. The document provides guidance on the information required to assess groundwater quality and the approach used to define site-specific groundwater guidelines. It also discusses recommended compliance approaches for groundwater quality.

The advice provided in this document assumes that the data under review has been collected according to relevant sampling and analysis methods and is of an acceptable quality for such purpose. This document does not provide guidance on the assessment of groundwater quantity, bore construction, sampling procedure, monitoring design or QA/QC processes.

Table of Contents

1. Introduction

Groundwater is a valuable natural resource that has a range of environmental values including the provision of drinking water for humans and livestock, irrigation of crops, industrial uses, cultural and spiritual values, ecosystem values and provision of water flows to groundwater dependent ecosystems. In dryland areas, groundwater can be the only reliable source of water and can sustain water levels in river and wetland ecosystems during extended dry periods. Accordingly, it is necessary to ensure the protection of this valuable natural resource. In Queensland, the Queensland *Environmental Protection Act 1994* and the *Environmental Protection (Water and Wetland Biodiversity) Policy 2019* provides a framework for protecting groundwater quality.

Groundwater quality can be highly variable, both spatially and temporally (Australian Government 2013), more so than surface water quality. Groundwater quality can be influenced by local geology, residence time in the aquifer, groundwater chemistry and groundwater-rock interactions. Groundwater can have naturally elevated salinity concentrations, dissolved nutrients and metals. Because of the high variability in groundwater chemistry, in some cases, groundwater quality may not meet water quality guidelines set out for some relevant environmental values (EVs).

Assessing groundwater quality is generally based on a comparison of measured groundwater quality indicators against default guidelines that usually relate to the potential use of the water if extracted or if it is expressed as surface water. These default guidelines include the Queensland Water Quality Guidelines (QWQG) (DEHP 2013), and Australian and New Zealand Guidelines for Fresh and Marine Water Quality (ANZG 2018) for the protection of freshwater aquatic ecosystems, stock drinking water, irrigation of crops and the Australian Drinking Water Quality Guidelines (NHMRC, NRMMC 2011).

Water Quality Objectives (WQOs) are also available for some regional groundwater zones across the state. These WQOs describe the range of water quality observed within groups of monitoring bores (i.e. zones) with similar characteristics. The WQOs are progressively being determined for groundwater zones in accordance with the *Environmental Protection (Water and Wetland Biodiversity) Policy 2019*, after approval they are included in Schedule 1. For more information on currently scheduled WQOs for groundwater see https://environment.des.qld.gov.au/management/water/policy.

As groundwater quality can be variable, the use of default guidelines or WQOs as a method of assessing groundwater quality may not always be appropriate. Similarly, more traditional statistical assessment of the monitoring data (i.e. comparing observations at reference bores with test bores) may not be suitable given this variability and typically, a lack of available data. The use of inaccurate or unrepresentative guidelines may lead to a high likelihood of either a false exceedance. In the absence of dedicated assessment methodology for groundwater, an appraisal of groundwater quality monitoring data can be challenging. Where default guidelines are unsuitable, the approach typically adopted is to compare the measured value/values with a value derived from site-specific or locally relevant background, reference or baseline groundwater quality monitoring. Under the approaches described in the national and Queensland water quality guidelines, locally relevant WQOs may be applied in preference to default values where these are representative and meet minimum sampling and quality assurance requirements.

The purpose of this document is to outline a process to assess groundwater quality data and/or refine groundwater guidelines, where necessary, from site-specific data. Sections 3 and 4 provide a description of the information required for groundwater assessment. Section 5 describes the process for defining site-specific groundwater guidelines. Compliance approaches are discussed in Section 6.

The application of site-specific groundwater guidelines into environmental approvals is discussed in Section 6. This guideline presents a recommended compliance approach that has significant advantages over previously adopted approaches. It involves assessment of consecutive test samples and applying either one or two different limit (or trigger) types, called Limit A and Limit B in this guideline. The site-specific groundwater guidelines, compliance approach and frequency of sampling should all be considered when assessing compliance.

This guideline provides complimentary information to the procedure outlined in the Wastewater release to Queensland waters (DES, 2018), which should be applied in conjunction to assess environmental approval conditions for groundwater quality against the provisions of the *Environmental Protection Act 1994* (EP Act) and subordinate legislation.

2. Summary

This guideline outlines a process to review groundwater quality monitoring data, including (i) the information required to assess groundwater quality, (ii) approaches used to define site-specific groundwater guidelines and (iii) comparisons of measured values with default guidelines, WQOs, site-specific guidelines derived from locally relevant background, reference or baseline groundwater quality data.

The process to review groundwater quality monitoring data and the adoption of site-specific groundwater guidelines or an alternative compliance approach involves a number of stages:

- Identify EVs for groundwater and relevant default guidelines and WQOs (Section 3)
- Describe site and bore characteristics (Section 4)
- Analyse groundwater quality monitoring data (Section 4):

  Step 1 – Determine summary statistics for each monitoring bore separately for all indicators using all available data.

  Step 2 – Compare the composition of each bore.

  Step 3 – Graph and interpret data for each bore. Outliers, trends and peaks in the data should be identified and investigated.

  Step 4 – Adjust dataset and recalculate percentiles (if required).
- Identifying site-specific guidelines for groundwater quality, if required (Section 5)

  Step 1 – Are there sufficient good quality monitoring data and bores to calculate statistically robust site-specific groundwater guideline?

  Step 2 – Compare 20th and 80th percentiles of monitoring data to relevant default guidelines.

  Step 3 – Determine site-specific guidelines for groundwater quality.
- Determine an appropriate compliance approach (Section 6).
- Evaluate site-specific groundwater guidelines, triggers, limits and compliance approach (Section 6).

If the site-specific circumstances and data analyses (from steps above) suggest that default guidelines are not suitable and site-specific guideline values are required, all information required to identify appropriate guidelines and assess potential impacts to groundwater quality should be described and provided to the relevant department. The list of information typically needed includes:

- Identify EVs for groundwater and relevant default guidelines, WQOs and existing environmental approval conditions including any groundwater-related limits or triggers (if applicable).
- Describe site and bore characteristics.
- Minimum groundwater quality monitoring data requirements satisfied. Check sampling techniques and data integrity including QA/QC of the data.
- Table of descriptive statistics provided for each bore (date range, count, mean, minimum, maximum, median, 5^{th} (for pH), 20^{th}, 80^{th} and 95^{th} percentiles).
- Box plots for all bores for each indicator and time series plots of all available groundwater quality data.
- Investigation of increasing trends and peaks in the water quality data.
- Distribution of major anions and cations in each bore (e.g. piper diagram).
- If required, site-specific groundwater guidelines and compliance approach investigated.
- Testing and evaluation of proposed compliance approach.

3. Understanding the System

Any groundwater impact assessment requires a good understanding of what groundwater is, where it occurs, how it interacts with surface waters, and what environmental values it supports. The term ‘groundwater’ refers to water that seeps into the ground and accumulates in the pores and cracks of the saturated zone of the earth’s crust. Groundwater can occur in the *saturated zone* (e.g. aquifers and aquitards) where all available spaces are filled with water, and in the *unsaturated zone*, the area between the land surface and the water table (upper surface of the saturated zone of an unconfined aquifer), where there are pockets of air that contain some water (Centre for Groundwater Studies 2001). For further details on groundwater properties see Appendix 1.

3.1 Aquifers and aquitards

Aquifers are the zones within the underlying geology of an area that are saturated with groundwater (Figure 1). In some cases aquifers can provide suitable habitat for stygofauna and troglofauna (specialised groundwater fauna). In contrast, an aquitard is a zone within the earth that restricts the flow of groundwater from one aquifer to another. Aquitards comprise layers of either clay or non-porous rock with low hydraulic conductivity. These layers permit some vertical recharge; however, the movement of water is limited (or retarded).

Aquifers can be confined or unconfined. Confined aquifers are aquifers that are overlain by a confining layer or aquitard, often made up of clay (Figure 1). Aquifers can also be categorised, according to the physical properties of the rock, as unconsolidated or surficial (e.g. sand), sedimentary (e.g. sandstone) and fractured rock (e.g. fractures in sandstone or granite). Unconsolidated aquifers commonly occur in alluvium deposits (loose, unconsolidated soil or sediment which has been eroded or reshaped by water) and exist mostly within 100 m of the surface. Sedimentary aquifers are confined aquifers that are bounded by an impermeable (aquiclude) base and low permeable aquitard above (Figure 1). Fractured rock aquifers exist where groundwater fills voids that have been formed by fractures, fissures and faults.

3.2 Groundwater flow

To assess and manage anthropogenic activities that may impact on groundwater an understanding of at least the following characteristics describing groundwater flow is required:

- Hydrogeology and geochemistry
- Hydraulic properties
- Geophysics (surface and subsurface)
- Flow modelling and water balance studies
- Estimation of aquifer variability (Centre for Groundwater Studies 2001).

For further details on the groundwater hydrogeological and hydraulic properties see Appendix 1. Other information needs to be sourced if guidance is required on determining the hydrogeology or hydraulic properties of a site.

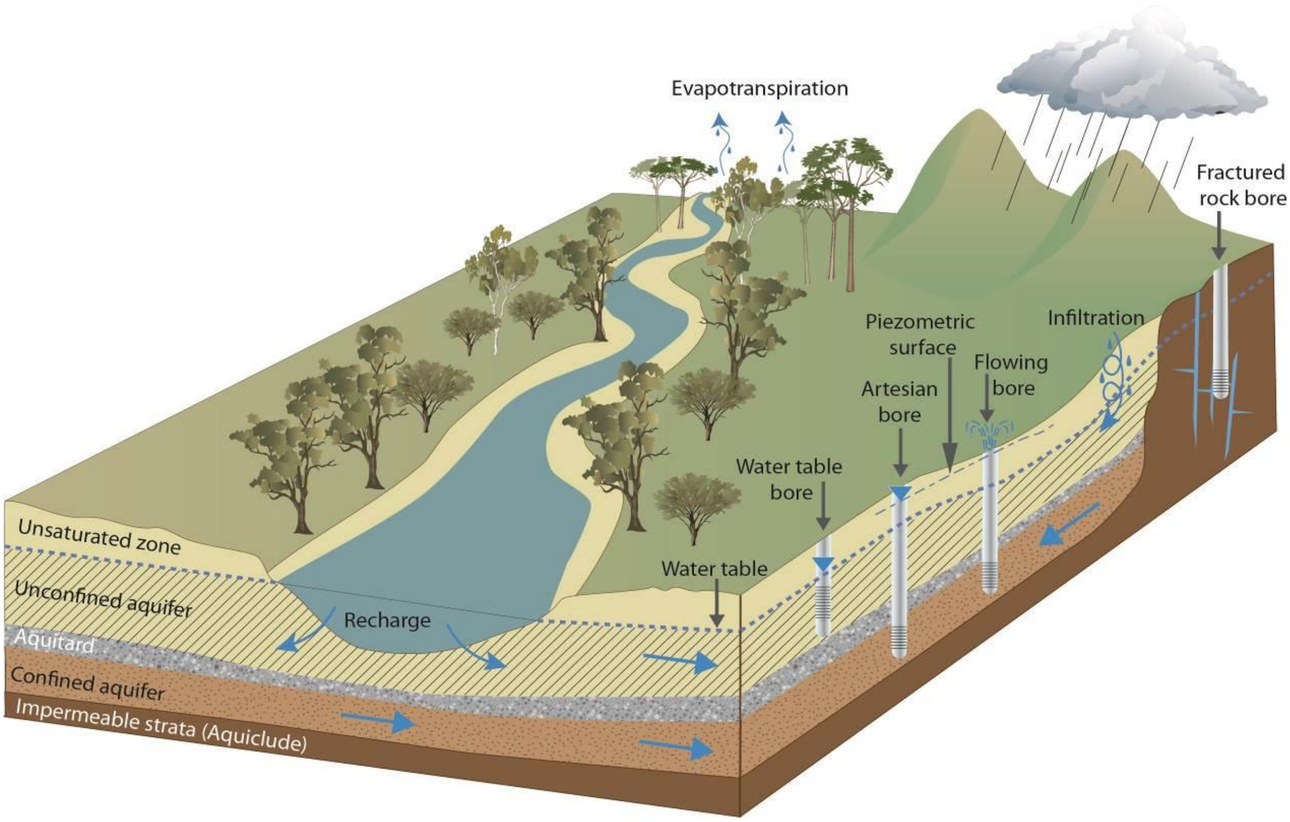


Figure 1: Conceptual model of groundwater processes and properties

3.3 Groundwater–surface water interaction

Groundwater-surface water interaction refers to the direction and magnitude of flow between water resources located above and below ground. A watercourse or wetland may lose or gain water to, or from, an aquifer. A surface water system can be connected to groundwater ephemerally, intermittently, seasonally or permanently (Wetland*Info* 2014). Along the length of its course, a river may receive groundwater from an aquifer, lose water to it, or both depending on the season and stage of river flows (Figure 2). A watercourse or wetland can either be a:

- Gaining Stream - receiving inflow of groundwater
- Losing Stream - losing water to the groundwater system by leakage to the aquifer.
- Streams that do both - gaining in some parts and losing in others, or perhaps alternating between gaining and losing depending on periodic changes in relative stream and groundwater levels.

Assessing groundwater-surface water interactions is complex and difficult, but important in determining water quality. Poor groundwater quality has the potential to impact both surface water quality and groundwater dependent ecosystems (GDEs). A GDE is an ecosystem that requires access to groundwater to maintain communities of plants and animals, ecological processes and ecosystem services. Springs are important waterbodies that may support GDEs that depend on the surface expression of groundwater. Guidelines which provide advice on assessing GDEs include *Groundwater-dependent ecosystems Toolbox* (Richardson et al. 20111) and *Information Guidelines Explanatory Note: Assessing groundwater-dependent ecosystems* (Doody et al. 2019).

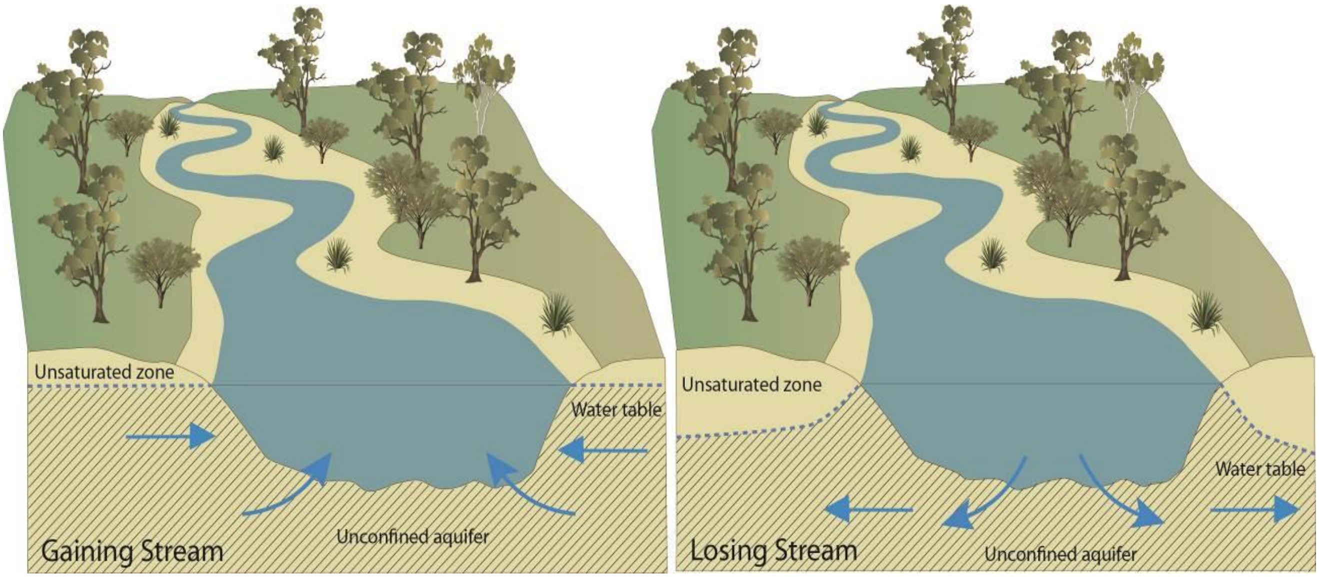


Figure 2: Conceptual model of a gaining and losing stream

3.4 Environmental values

In Queensland, the *Environment Protection Act 1994* and subordinate legislation including the Environmental Protection Regulation 2019 and Environmental Protection (Water and Wetland Biodiversity) Policy 2019 provide the legislative basis for protecting groundwater quality.

Based on the intent of the *Environment Protection Act 1994* and Environmental Protection Regulation 2019, groundwater quality is to be protected. The groundwater Performance Outcomes as stated in the Environmental Protection Regulation 2019 are:

- 1 Both of the following apply—
  - (a) there will be no direct or indirect release of contaminants to groundwater from the operation of the activity;
  - (b) there will be no actual or potential adverse effect on groundwater from the operation of the activity.
- 2 The activity will be managed to prevent or minimise adverse effects on groundwater or any associated surface ecological systems.

To protect the intrinsic environmental value of groundwater, the groundwater quality should be maintained within the range of natural quality variations established through baseline characterisation to ensure that no adverse effect on groundwater quality occurs from the operation of the activity.

Under the Environmental Protection (Water and Wetland Biodiversity) Policy 2019 the environmental values (EVs) of groundwater must be enhanced or protected. Environmental values (EVs) are the particular values or uses of the environment that are conducive to public benefit, welfare, safety or health and that require protection from the effects of pollution, waste discharges and deposits (NEPC, 1999). Several different EVs may be relevant for a particular water body. The desirable EVs should be those that the local community wish to protect and enjoy now and in the future. For several regions in Queensland, the range of EVs and associated water quality objectives (WQOs) for both surface and groundwater have been determined and scheduled as part of the *Environmental Protection (Water and Wetland Biodiversity) Policy 2019*. Further EVs and

WQOs for additional regions of Queensland will be determined and scheduled as they become available.

Additionally, EVs of potentially affected groundwater on or adjacent to the area of interest and water quality objectives to protect or enhance these values must be identified. The main nationally recognised EVs or uses of groundwater are:

Ecosystem values

- ecosystem protection (flora, fauna and habitat).

Human use values

- agricultural use (irrigation and stock watering)
- recreational use
- drinking water supply
- industrial use
- cultural values.

Each of these EVs requires its own specific set of default guidelines because the acceptable guidelines to maintain one type of EV may not be acceptable to maintain another EV. In addition, the ANZG 2018 toxicity based triggers for the protection of surface and groundwater aquatic ecosystems are applied to metals that are bioavailable to aquatic organisms, therefore this is the dissolved (filtered) fraction. Human use guidelines (e.g. stock watering) should be applied to total concentrations of metals, as the water is ingested. Water quality default guidelines give recommended values for indicators and are designed to ensure that EVs of waters are protected.

The ANZG 2018, as well as the QWQG and the *Environmental Protection (Water and Wetland Biodiversity) Policy 2019* recognise that some aquatic ecosystems have more intrinsic ecological value than others, and provide a framework for determining the level of protection that should be allocated to aquatic ecosystems. The Guidelines define level of protection as ‘the degree of protection afforded to a water body based on its ecosystem condition (current or desired health status of an ecosystem relative to the degree of human disturbance)’. For each level of protection, slightly different default guidelines may apply.

Where multiple EVs (with associated water quality guidelines) have been identified for an area of interest, surface or groundwater quality should be managed to comply with the most conservative water quality guideline so that the most sensitive EV is protected.

For further details regarding environmental values, level of protection and guidelines, see Appendix 2.

Example – Identifying EVs for groundwater

The following provides an example for identifying relevant EVs, level of protection and guidelines. This example will be referred to elsewhere in the document.

An interaction between groundwater and surface water was found between the shallow aquifer and the adjacent creek. Stygofauna was also detected in the groundwater bores. Therefore, the protection of groundwater aquatic ecosystems and surface water aquatic ecosystems was an EV that was to be protected.

It was also recognised that the aquatic ecosystems may have been adversely affected to a relatively small but measurable degree by human activity; therefore, the level of protection was classified as a slightly to moderately disturbed system under ANZG 2018.

Using the precautionary principle, the toxicant default guidelines for surface water are often applied to groundwater. The scheduled WQOs for Shallow (< 30m) Zone 24 Lower Nogoa groundwaters for pH and EC and the default toxicant guidelines for dissolved arsenic, copper, manganese and zinc were relevant in this case (Table 1).

Table 1. Relevant default guidelines and WQOs for the above example

Indicator	Default guidelines and WQOs
pH	7.9 - 8.2
Electrical Conductivity (µS/cm)	6,908
Dissolved arsenic (mg/L)	0.013
Dissolved copper (mg/L)	0.0014
Dissolved manganese (mg/L)	1.9
Dissolved zinc (mg/L)	0.008

4. Background Information and Data Analysis

Differences exist in the presence, accumulation and dispersion of contaminants between surface and groundwater systems. In contrast to surface waters, the flow of contaminants into and through an aquifer may take years or decades. The average residence time of contaminants in sedimentary aquifers might be in the order of centuries, whereas the residence time for surface water can be measured in hours or days for intermittent releases.

Once contaminants enter the groundwater they tend to form plumes outward from the source. The shape and size of the plume will depend on the local geology, the groundwater flow, the type and concentration of the contaminant, the duration of waste disposal and human modifications to the groundwater system (Todd and May 2005). This can mean that the cost of cleaning up groundwater, once polluted, is often extremely high and technically complex. Therefore, where possible the prevention of groundwater contamination should be the management intent for all activities.

Poor groundwater quality has the potential to impact environmental values. For example, groundwater can be used directly for supply to potable water treatment plants, stock watering and crop irrigation and poor water quality can affect its use for these purposes. In addition, if an aquifer becomes polluted, it is possible that polluted water may be transported to one or more discharge sites where there is surface expression. Such a consideration is important when the discharge site is a surface water body that is used for drinking water, stock watering or has aquatic ecosystem values.

To assess groundwater quality it is essential that the groundwater system and groundwater quality characteristics are adequately described and understood. Such information is required to assess current groundwater quality and the potential future risks to the groundwater. The information required would typically include:

- information that describes groundwater hydrogeology and hydraulics within the potentially impacted aquifers.
- water quality characteristics, including pH, electrical conductivity and the major cation and anion composition of the groundwater within the potentially impacted aquifers. For metals, both the dissolved and total concentrations should be measured.

The suggested minimum information requirements to assess groundwater quality is discussed in more detail below for (i) site and bore characteristics and (ii) water quality monitoring and data analysis.

The advice provided in this document assumes that the data under review has been collected according to relevant sampling and analysis methods and is of an acceptable quality for such purpose. Appropriate bore construction, sampling procedure, monitoring design, laboratory analysis methodology and QA/QC processes are essential in providing good quality data. Guidelines which provide advice on these aspects include *Bore assessment guideline under section 413 of the Water Act 2000* (DEHP 2016: EM1178), *Groundwater and Sampling Analysis – A Field Guide* (Sundaram et al. 2009) and the *Queensland Monitoring and Sampling Manual* (DES 2018).

4.1 Site and Bore Characteristics

Site plans and conceptual models of the surface water and groundwater aquifers are required to facilitate an understanding of the overall system and interactions that are subject to assessment. This will provide the basics for a conceptual understanding of the environmental values and the associated hazards and risks to water quality.

The conceptual model can be a simple box diagram that illustrates the components and linkages in the system, or a graphical representation of the system. Whatever model is used, it should present the factors that are perceived to be driving the changes in the system and the consequences of changes to these factors. For further details on how to develop conceptual models of groundwater systems see Barnett *et al*. (2012) and DSITI (2015).

A description of the characteristics of a site should include; the hydrogeological setting, the movement of groundwater; potential contaminants and sources, the interactions between groundwater and the surface, where groundwater discharges to the surface and the potential pathways of contaminant transport.

These characteristics can be identified on a scaled site plan that, ideally, includes:

- The location and coordinates of all monitoring bores overlayed on a map of the site with classification of the bores, such as compliance bore.
- Screening depth of each monitoring bore and a description of the geological unit and aquifer.
- A map showing topographical contours at suitable increments, shown with respect to Australian Height Datum.

- A description of hydrogeological features of the site and aquifers that includes soil and rock types (including porosity, permeability) and stratigraphy (including faulting and facture propensity).
- Identify the location of potential surface – groundwater interaction and sources of contaminants, such as the location of mine affected water dams, brine ponds, seepage collection drains and locations where there is groundwater expression to the surface (e.g. seeps and springs).
- Identify the direction(s) of surface water runoff and drainage lines that pass through or are near the site and any surface waters potentially impacted by the activity (including rivers, creeks, lakes, wetlands or drainage lines) that are within or adjacent to the site.
- The location and depth to groundwater (including perched aquifers or water tables), the depth to water level (potentiometric surface) on the site and the groundwater system boundaries.
- An indication of the movement (including direction and rate of flow) of groundwater across the site.
- Identify and describe any geological barriers that are overlying and underlying aquifers. The hydraulic connectivity within, and between, aquifers may also be important.
- Any existing (registered and unregistered) or proposed water bores or groundwater monitoring wells within the site or on land adjacent to the site, including bore log information.
- Identify the location of current and historic activities that may adversely impact water quality such as pits, dams, waste rock dumps, tailing storage facilities, releases to groundwater and spills.
- The location of waste storage, processing, treatment, and disposal locations. Include details for both raw and treated wastes and details of the relevant storage facilities. Plans must show any proposed point source discharges to waters from waste management processes onsite.
- Identify any environmentally and culturally sensitive places and describe the environmental values within or adjacent to the site.
- The location of underground ecosystems or groundwater dependant ecosystems associated with groundwater, details of those ecosystems and their interactions with the groundwater.

There is also a need to describe any historic events that may have affected groundwater at each bore across the life of the site to determine whether the bores have been contaminated and are highly disturbed as a consequence.

4.2 Water quality monitoring data analysis

The water quality characteristics of the groundwater at the site and in the surrounding area are important in assessing current groundwater quality and risks to the quality of the groundwater into the future.

Ideally, to protect the intrinsic environmental value of groundwater, groundwater quality should be maintained within its natural range of variability. In order to demonstrate that background groundwater quality has not been affected by an activity, there is a need to compare data collected before and after the activity commenced. This requires that detailed baseline assessments are undertaken to accurately represent the water quality in the aquifer prior to the activity commencing. The baseline assessments should also establish natural groundwater quality and variability in order to confidently attribute any potential contamination to a new development or activity. The length of the baseline monitoring and the frequency of sampling must be sufficient to establish confidence in the natural variability of groundwater quality in the project area (e.g. seasonal variability (wet and dry season) may be important for some aquifers and not others).

There is limited guidance in Australia on the minimum number of samples required to assess groundwater quality. The U.S. Interstate Technology & Regulatory Council (ITRC 2013) identified that most guidelines on sample size for groundwater tests recommend at least eight to ten background measurements when constructing limits and roughly the same number of compliance point measurements when calculating trend tests or confidence intervals. The ANZG 2018 recommend a minimum of two years of monthly baseline data to establish natural variability for assessing surface water. Longer periods of baseline monitoring are likely to be required to assess natural variability in groundwater quality as groundwater sampling is typically done less frequently than surface water samples, as groundwater typically moves slower than surface water, and this should be assessed on a case by case basis.

As a minimum, it is suggested that groundwater quality be monitored quarterly, however, in some cases where water quality may change over short time frames, more frequent monitoring may be required.

In order to increase the number of samples over a 12 month period, it may also be appropriate to sample multiple bores that represent the same aquifer, provided that it can be demonstrated that they have similar direction and rate of groundwater flow, geology, soil types and ionic composition. It is then possible to combine the data (from multiples bores) to calculate more robust descriptive statistics (e.g. 20th, 50th and 80th percentiles).

The proportions of the major cations and anions within different monitoring bores can provide an indication to the degree of connectivity between groundwater bores. The major cations and anions are influenced by the different lithologies, mineral suites, soil types and recharge source within the groundwater catchment. The ionic composition of the groundwater can also be influenced by mining activities, including tailing storage facilities, waste rock dumps and storage dams. Eight major ion species make up 95% of inorganic ions in groundwater (Centre for Groundwater Studies 2001). The major cations include sodium, potassium, calcium and magnesium and the major anions include chloride, sulphate, bicarbonate and carbonate. In addition, nitrate can also be proportionally abundant.

Piper diagrams and time series plots can also provide a rapid means of identifying whether there is a difference (or similarity) between bores. If a difference between bores is suspected, a test of statistical difference (e.g. Student's *t*-test or Mann-Whitney *U* Test) could be performed to determine the significance of the difference.

The groundwater quality indicators to be monitored should be determined based on the identification of contaminants of potential concern that are associated with the activity, and the EVs and associated WQOs for the groundwater aquifers and relevant surface waters. Regardless of the activity, major anions and cations must be monitored for all bores in order to characterise the groundwater. Such data allows the ionic signature to be compared between bores, which will assist in identifying whether bores are chemically comparable.

The recommended steps involved in the analysis of groundwater quality monitoring data are:

> **Step 1** - Determine summary statistics for each monitoring bore for all indicators using all available data.
>
> **Step 2** - Compare the ionic composition of each bore.

Step 3 - Graph and interpret data for each bore.

Step 4 - Adjust dataset and recalculate percentiles (if required).

All available groundwater quality data for all monitoring bores including the dates of sampling, number of samples, and range of indicators should be presented in addition to summary statistics (e.g. 5^{th}, 20^{th}, 50^{th}, 80^{th} and 95^{th} percentile, minimum and maximum at each monitoring bore). Outliers and trends in the data should be identified. An outlier is a measured data point that is distant from other data points (see Appendix 3 for details). There are a range of statistical techniques that may be suited to identify an outlier. As an example, an outlier can be identified as a data point which is greater than four standard deviations from the mean (U.S. EPA 2009). An outlier may be due to variability in the measurement or it may indicate measurement error. This data should be investigated further with regard to QA/QC and environmental conditions at the time of sampling.

When assessing compliance against environmental conditions, outliers should be treated with caution and must **not** be removed from the analysis as they may indicate potential impacts. However, when using a dataset to calculate percentiles to determine local site-specific groundwater guidelines, outliers should be removed.

Graphs provide a powerful evaluation tool by visually summarising data characteristics. If major cation and anion data are available, a piper diagram (or an equivalent means of representing ion composition) should also be provided. The data from each monitoring bore should be plotted as time series plots and box and whisker plots. Time series plots are an excellent tool for examining the behaviour of one or more indicators over time, as they display each and every data point and provide an initial indication of temporal dependence (U.S. EPA 2009). The time series data should also be analysed for trends. If a trend, either upward or downward, is detected then the groundwater may be impacted. A Mann-Kendall Test could be used to identify the presence of a significantly increasing or decreasing trend at a compliance bore or any trend in background data sets.

Box plots provide a graphical summary of data distribution and give an indication of spatial variability across multiple bore locations by presenting the central tendency, dispersion and unequal variances in the data at each bore. The box part of the box plot can represent the median, 20^{th} and 80^{th} percentile or quartiles (25^{th} and 75^{th} percentiles), and the whiskers can represent either the minimum and maximum for each bore or the 1.5 times the Inter Quartile Range. A separate box should be presented for each bore for a given water quality indicator and each box plot should also show the relevant default guideline and WQOs. See Appendix 3 for further details regarding non-detects, outliers and graphs.

For further details regarding site characteristics, groundwater quality monitoring and data analysis see "*Groundwater Statistics and Monitoring Compliance*" (ITRC 2013).

Example - Data analysis

This example outlines a typical assessment of groundwater quality data based on real data that accompanied applications for amendments to Environmental Authorities.

Step 1 - Determine summary statistics for each monitoring bore for all indicators using all available data.

Summary statistics for pH, electrical conductivity (EC) and dissolved arsenic, copper, manganese and zinc are provided in Table 2.

Table 2. Example summary statistics at six monitoring bores

	Monitoring Bores					
	Bore 1	**Bore 2**	**Bore 3**	**Bore 4**	**Bore 5**	**Bore 6**
pH						
Count (n)	28	27	34	28	34	33
% < LOR	0	0	0	0	0	0
Min	7.66	7.38	7.39	7.48	7.64	7.66
5th percentile	7.91	7.56	7.57	7.56	7.71	7.82
20th percentile	7.96	7.76	7.77	7.76	7.8	7.98
50th percentile (Median)	8.09	7.85	7.94	7.96	7.95	8.16
80th percentile	8.22	8.02	8.2	8.35	8.11	8.31
95th percentile	8.55	8.27	8.37	8.89	8.56	8.45
Max	8.64	8.46	8.53	9.08	8.98	8.69
Electrical Conductivity (µS/cm)						
Count (n)	28	27	34	28	34	33
% < LOR	0	0	0	0	0	0
Min	3,940	3,600	4,900	1,440	2,260	930
20th percentile	6,056	4,058	6,724	1,534	2,596	1,334
50th percentile (Median)	6,700	4,180	7,060	1,645	3,045	1,380
80th percentile	8,040	4,274	11,596	2,528	3,516	1,594
95th percentile	8,396	4,331	18,341	2,845	4,018	1,846
Max	8,470	4,860	18,790	3,190	4,230	2,000
Dissolved arsenic* (mg/L)						
Count (n)	12	27	34	27	34	31
% < LOR	0	0	0	4	0	5
Min	0.002	0.002	0.004	0.001	0.001	0.001
20th percentile	0.003	0.002	0.006	0.001	0.002	0.008
50th percentile (Median)	0.005	0.002	0.008	0.003	0.003	0.024
80th percentile	0.006	0.003	0.01	0.003	0.004	0.028
95th percentile	0.006	0.007	0.01	0.004	0.004	0.033
Max	0.006	0.008	0.012	0.006	0.005	0.036

	Monitoring Bores					
	Bore 1	Bore 2	Bore 3	Bore 4	Bore 5	Bore 6
Dissolved copper (mg/L)						
Count (n)	6	18	15	18	14	27
% < LOR	54	33	57	36	60	25
Min	0.001	0.001	0.001	0.001	0.001	0.001
20th percentile	0.001	0.001	0.001	0.001	0.001	0.001
50th percentile (Median)	0.002	0.002	0.002	0.002	0.001	0.004
80th percentile	0.002	0.004	0.003	0.004	0.002	0.005
95th percentile	0.002	0.004	0.003	0.005	0.002	0.010
Max	0.002	0.005	0.004	0.006	0.002	0.017
Dissolved manganese (mg/L)						
Count (n)	12	27	34	28	34	30
% < LOR	0	0	0	0	0	19
Min	2.5	0.001	1.34	0.033	0.09	0.001
20th percentile	2.94	0.002	1.78	0.078	1.41	0.008
50th percentile (Median)	3.36	0.004	2.25	0.46	2.44	0.156
80th percentile	3.71	0.033	4.04	0.642	3.24	0.383
95th percentile	3.93	0.258	7.73	1.14	3.85	0.53
Max	4.17	4.917	11.1	1.42	3.9	0.683
Dissolved zinc (mg/L)						
Count (n)	10	26	34	28	31	31
% < LOR	15	4	0	0	11	8
Min	0.005	0.005	0.007	0.005	0.005	0.006
20th percentile	0.006	0.007	0.012	0.009	0.008	0.065
50th percentile (Median)	0.007	0.013	0.018	0.014	0.014	0.1
80th percentile	0.009	0.018	0.022	0.025	0.019	0.164
95th percentile	0.012	0.026	0.03	0.556	0.025	0.225
Max	0.014	0.039	0.05	1.06	0.041	0.252

* The ANZG 2018 specify different toxicant trigger values for arsenic Arsenite As(III) and Arsenate As(V); when the speciation of arsenic is unknown, as in this case, the more conservative trigger value (Arsenate As(V)) should be used.

Step 2 - Compare the ionic composition of each bore.

Produce a plot that visualises the ionic chemistry of the groundwater sample (e.g. piper plot, see Appendix 3 for details). Determine whether the groundwater bores are from the same aquifer or can be grouped based on their ionic composition.

Step 3 - Graph and interpret the data for each bore.

An assessment of the data to determine if there are any increasing or decreasing trends, peaks and outliers is important to determine potential impacts and whether the data is used to derive site-specific groundwater guidelines. Based on the box plots and time series plots, outliers and trends in the data were identified.

In general, data with values increasing over time, as in the electrical conductivity example in Figure 3, should not be used to calculate site-specific groundwater guidelines. An increasing trend in the data may indicate contamination of the groundwater and should be investigated. In this example, there appears to be a strong trend in the data.

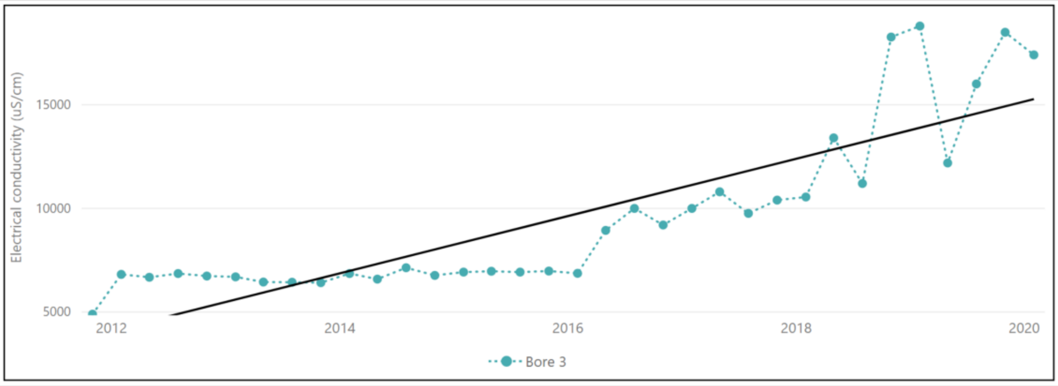


Figure 3: Electrical conductivity (µS/cm) data at monitoring Bore 3. Linear regression line included to illustrate increasing trend.

An outlier is a measured data point that is distant from other data points (See Appendix 3 for details), as in the example data presented in Figure 4. The concentration of dissolved metals should be less than the concentration of total metals. In the data presented in Figure 4, the concentration of dissolved metal *y* at the circled outlier data point was greater than four standard deviations from the mean and was greater than the concentration of total metal *y*; therefore, the outlier should be removed from the data set used to calculate site-specific groundwater guidelines.

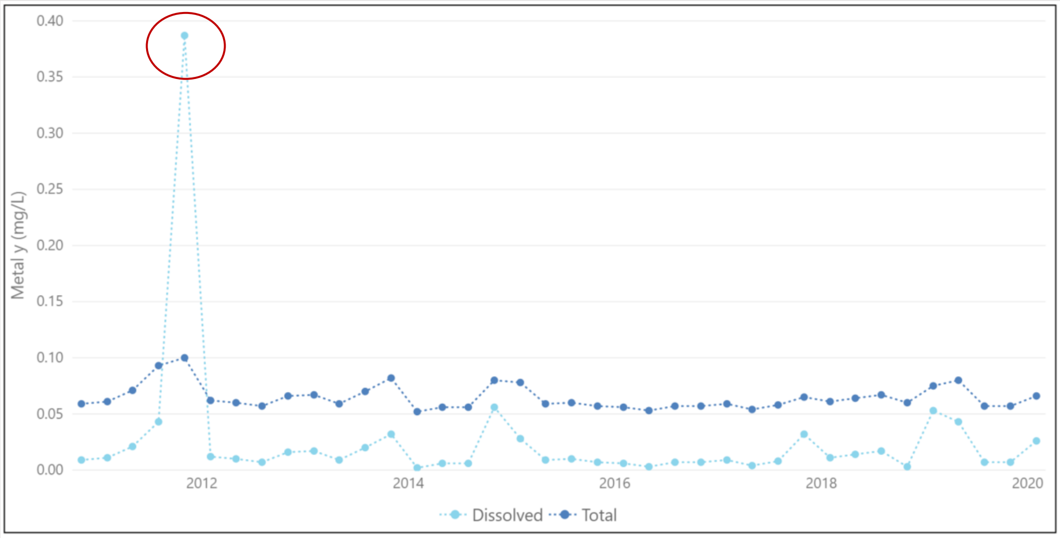


Figure 4: Dissolved and total Metal *y* concentration (mg/L) at monitoring Bore 5. Outlier in the data is circled.

A peak in data may be due to a potential contamination event and should be investigated, as in the example data presented in Figure 5. Such data points should not be used to calculate site-specific groundwater guidelines.

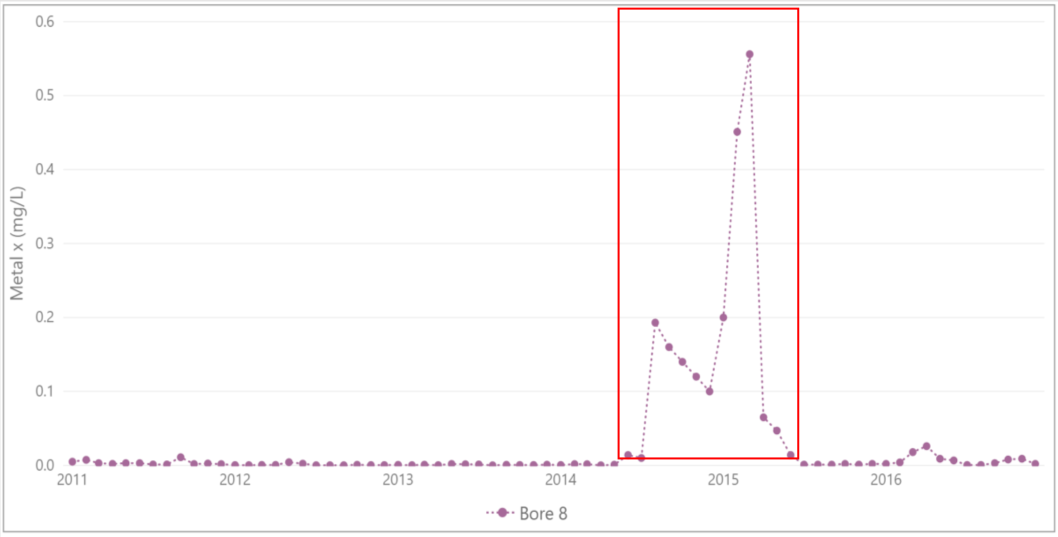


Figure 5: Metal *x* (mg/L) data at monitoring Bore 8.

Step 4 - Adjust dataset and recalculate percentiles (if required).

If the data is to be used to determine site-specific groundwater guidelines, trends, outliers and peaks identified in Step 3 should be removed from the dataset and percentiles should be recalculated.

The difference in an 80th percentile calculated with all data and a revised 80th percentile with the peak removed (data points inside the red box in Figure 5) is presented in Figure 6 for comparison.

Figure 6: Metal *x* (mg/L) data at monitoring Bore 8. The 80th percentile of all data (red line), revised 80th percentile (orange line) and default guideline (green line) are also presented.

5. Identifying Site-specific Guidelines for Groundwater Quality

The use of ANZG 2018 default guideline values for metals or toxicants often provides a conservative approach to protect surface and groundwater. It is suggested that these be adopted where possible. However, default guidelines should not be adopted as upper limits to which groundwater contaminant concentrations can be increased. Rather, if existing groundwater quality is below the water quality guidelines, then the groundwater should be managed within the range of natural water quality variability determined through baseline characterisation (Australian Government 2013).

It is also recognised that in some cases the concentration may be naturally greater than published default guideline for the protection of identified EVs, and it is necessary to identify alternative values (if there is sufficient data that meets quality standards). In such cases, the EVs still apply to the groundwater system; however, a site-specific guideline may be applied that is greater than the default guideline.

The toxicity and bioavailability of some metals (e.g. copper and zinc) are strongly influenced by water quality conditions such as hardness, pH or dissolved organic carbon. The default guidelines for these metals should be adjusted to the site-specific conditions at the time of sampling, where adjustment factor or formula are provided in the guidelines. Where relevant site-specific conditions such as hardness have not been determined, the un-adjusted default guidelines provide conservative values.

The identification of site-specific groundwater guidelines can be complicated by the need to account for underlying geology and influence of historic mining activities on groundwater. Site-specific groundwater guidelines may be defined to consider water quality issues specific to the groundwater system in question. Under the ANZG framework, if local reference values exceed default guidelines, then (for naturally occurring toxicants) site-specific values can be developed. In those situations, reference bore water quality is considered to be a suitable baseline or benchmark for assessment and management of bores with similar characteristics. Most commonly, reference condition refers to bores that are subject to minimal/limited disturbance and are not impacted by anthropogenic activity. The criteria adopted for minimally disturbed surface water reference sites are shown in Table 4.4.1 of the QWQG (DEHP 2013) and these are also applicable to groundwater.

The process used to derive site-specific guidelines, as outlined in Table 4.4.4 of the QWQGs (DEHP 2013), is based on the level of protection that has been attributed to the groundwater system. For high ecological value waters there should be no change to the natural attributes of the system. To derive local guidelines for high ecological value waters the 20^{th}, 50^{th} and 80^{th} percentiles are required, for one to two reference bores. For slightly to moderately disturbed waters the 80^{th} percentiles (and 20^{th} percentile for indicators such as pH and dissolved oxygen) of reference bore values are used. For highly disturbed systems a less stringent guideline can be derived at a local level based on a) a less stringent percentile, e.g. 10^{th}/90^{th} or b) reference data from more impacted but still acceptable reference bores.

The reference condition concept can be applied to disturbed systems. There are some regions (e.g. South East Queensland), and some aquifers, where it may be difficult to find any bores that would meet reference site criteria. In this situation it may be necessary to use lesser quality or best available groundwater quality data.

Where the data contains results that are less than the limit of reporting (LoR) (also referred to as Reporting Limit (RL) or non-detects) a simple substitution of one-half the LoR can be used when the dataset contains a relatively small proportion of non-detects (i.e. no more than 10-15%, for the specific purpose of deriving site-specific guidelines). If there are a large number of non-detects then these values should be removed from the dataset before site-specific guidelines are calculated as the summary statistics will be biased. Site-specific guidelines that are based on a factor greater than the LoR such as 3 or 5 times the LoR are **NOT** recommended, as LoR are based on laboratory methods and not the environmental or health risks associated with a toxicant. If there are less than eight data points available, that are greater than the LoR, it is suggested that the default guideline or WQO be applied until sufficient data is collected to calculate a site-specific value.

The steps to derive site-specific guidelines for groundwater quality for slightly to moderately disturbed bores that fail to meet some reference site criteria, but are not heavily impacted, are described below. These are based on the ANZG 2018 and QWQG 2009 (DEHP 2013).

Step 1 – Are there sufficient good quality monitoring data and bores to calculate a statistically robust site-specific groundwater guideline?

Determine if there are sufficient monitoring bores and data points. It is recommended that for estimates of 20th and 80th percentiles a minimum of 18 samples over at least 12 and preferably 24 months. However, percentile estimates based on eight samples can be used to derive guidelines. See QWQG (DEHP 2013) for further information on minimum data requirements.

Quality control checks on the raw data should be done to determine if there are outliers or trends in the data. Outliers should be identified, investigated and removed from the data set used to calculate percentiles (See Section 3). Trends should be documented in Step 2.

Assess the proportion of LoRs and where there are a low proportion of LoRs determine whether a simple substitution of one-half the LoRs can be used or whether the LoRs should be removed from the dataset. If there are less than eight data points available that are greater than the LoR, it is suggested that the default guideline or WQO be applied until sufficient data is collected to calculate a site specific value.

Step 2 – Compare 20th and 80th percentiles of monitoring data to relevant default guidelines.

Calculate 20th and 80th percentiles for each bore and compare with default guidelines or regional WQOs (e.g. the regional objectives defined under Schedule 1 of the *Environmental Protection (Water and Wetland Biodiversity) Policy 2019*).

Some physio-chemical indicators (e.g. dissolved oxygen and pH) have an acceptable guideline range where an upper and lower guideline value applies. The 20th and 80th percentile is relevant for these indicators.

Ideally, there should be three or more monitoring bores for each aquifer and the 20th and 80th percentile values for each bore should be calculated and then compared with each other. If the 20th and 80th percentiles from all bores from the same aquifer are consistent, then combined 20th and 80th percentiles can be calculated.

Prepare time series and box plots for monitoring bores of key indicators and overlay the relevant default guidelines and regional WQOs on the plots and assess whether:

- typical groundwater quality in the area is different from the default guidelines and WQOs for some indicators
- data is generally consistent between bores.

Step 3 – Determine site-specific guidelines for groundwater quality.

The 80th percentiles from all bores from the same aquifer should form the basis of the site-specific groundwater guidelines. The 20th percentiles from all bores from the same aquifer are relevant to indicators where a lower guideline value applies.

Determine whether the site-specific groundwater guidelines are more appropriate than existing default guidelines and WQOs. Below are four scenarios that may arise when determining whether default or site-specific groundwater guidelines are most appropriate for the area.

- Default guidelines and WQOs are available and the 80th percentile from site-specific bores are not substantially different from the default guideline or WQO. In this case it is recommended that the default values are adopted.
- Default guidelines and WQOs are available and the 80th percentile values from site-specific bores are significantly less than the default guideline or WQO. In this situation, for indicators where the site-specific values are significantly better than default values, then the site-specific values should be adopted. However, for toxicants, if the local groundwater quality is less than the default guideline (ANZG 2018), then the toxicant default guideline value may still be applicable.
- Default guidelines and WQOs are not available. In this situation, the 80th percentile value from the local bores (even where slightly impacted in some circumstances, which would need to be agreed by the regulator) may be the best available option for deriving a site-specific guideline. It is therefore, recommended that the site-specific values be adopted.
- Default guidelines and WQOs are available and the 80th percentile values from site-specific bores exceed the default guidelines. If the site-specific values better reflect the local conditions, then site-specific values would typically be adopted.

The below example applies the above methodology to determine site-specific groundwater guidelines for water quality. Once site-specific groundwater guidelines have been determined then an appropriate compliance approach should be considered. This is discussed in Section 5.

Example – Deriving Site-specific Groundwater Guidelines

Following on from the examples in Section 2 and 3, we are now applying the above methodology to determine appropriate site-specific groundwater guidelines.

Step 1 – Are there sufficient good quality monitoring data and bores to calculate statistically robust guideline values?

For those sampled indicators/bores with less than eight samples available or a large proportion of LoRs, the toxicant default guideline for aquatic ecosystem is often adopted. The use of toxicant default guideline for aquatic ecosystem protection (ANZG 2018) provide a conservative approach to protect groundwater quality.

For this example, greater than eight samples were available for the majority of bores and further assessment was performed on the data. The monitoring duration was different between bores and no pre-activity water quality data were available.

Outliers and trends in the data were identified. Outliers were identified, investigated and removed from the data set used to calculate percentiles (See Section 3). Trends are documented in Step 2.

Step 2 – Compare 20th and 80th percentiles of monitoring data to relevant default guidelines or Scheduled WQO.

The scheduled WQOs for Shallow (< 30m) Zone 24 Lower Nogoa groundwaters for pH and EC and the toxicant default guidelines for dissolved arsenic, copper, manganese and zinc were relevant in this case.

In this example, site-specific hardness was not measured and the toxicant default guidelines were therefore not adjusted.

- pH – the 20th percentile for Bore 2. Bore 3, Bore 4 and Bore 5 are less than the WQO 20th percentile and therefore outside the range of the WQO. The 80th percentile for Bore 1, Bore 4 and Bore 6 is greater than the WQO 80th percentile and therefore, outside the range of the WQO.
- Electrical conductivity – the 80th percentile of EC at Bore 1 and Bore 3 was greater than WQO 80th percentile and the 80th percentile values for Bore 2, Bore 4, Bore 5 and Bore 6 are significantly less than the WQO. An increasing trend in EC at Bore 1, Bore 3, Bore 4 and Bore 6 is observed.
- Arsenic – the species of arsenic present should be identified. However, when the species of arsenic is unknown, as in this case, the more conservative, Arsenate As(V), should be used. An increasing trend in the concentration of dissolved arsenic is observed at Bore 6. This may suggest potential contamination and requires further investigation. Data from this bore should not be included in calculations of site-specific values for dissolved arsenic. All measurements from the other monitoring bores are below the toxicant default guideline.
- Copper – the 80th percentile of dissolved copper at all monitoring bores was greater than the toxicant default guideline. There was also no obvious trend in the concentration of dissolved copper over time at the monitoring bores. This may suggest that the concentration of dissolved copper is naturally greater than the toxicant default guideline.
- Manganese – there is a decreasing trend in the concentration of dissolved manganese at Bore 3. This may require further investigation. The 80th percentile of dissolved manganese at Bore 2, Bore 4 and Bore 6 is less than the toxicant default guideline, however, the 80th percentile at Bore 1, Bore 3 and Bore 5 was greater than the toxicant default guideline.
- Zinc – two outliers were observed at Bore 4 in 2014. These values were removed from the analysis and the summary statistics recalculated. A peak in the concentration of dissolved zinc between 2008 and 2011 is observed at Bore 6 and the concentration has not returned to the pre 2008 level. This may suggest potential contamination and requires further investigation. The 80th percentile of dissolved zinc at other monitoring bores was greater than the toxicant default guideline. There was also no obvious trend in the concentration of dissolved zinc over time at monitoring bores, except Bore 6. This may suggest that the concentration of dissolved zinc is naturally greater than the toxicant default guideline.

Adjusted summary statistics and exceedances of the default guideline values are provided in Table 3.

Table 3. Example adjusted summary statistics and exceedances of default guideline values at six monitoring bores

		Monitoring Bores					
	Default Guideline or WQO	Bore 1	Bore 2	Bore 3	Bore 4	Bore 5	Bore 6
pH	**7.9 – 8.2***						
Count (n)		28	27	34	28	34	33
% < LOR		0	0	0	0	0	0
Min		7.66	7.38	7.39	7.48	7.64	7.66
5^{th} percentile		7.91	7.56	7.57	7.56	7.71	7.82
20^{th} percentile		7.96	7.76	7.77	7.76	7.8	7.98
50^{th} percentile (Median)		8.09	7.85	7.94	7.96	7.95	8.16
80^{th} percentile		8.22	8.02	8.2	8.35	8.11	8.31
95^{th} percentile		8.55	8.27	8.37	8.89	8.56	8.45
Max		8.64	8.46	8.53	9.08	8.98	8.69
Electrical Conductivity (µS/cm)	**6,908***						
Count (n)		28	27	34	28	34	33
% < LOR		0	0	0	0	0	0
Min		3,940	3,600	4,900	1,440	2,260	930
20^{th} percentile		6,056	4,058	6,724	1,534	2,596	1,334
50^{th} percentile (Median)		6,700	4,180	7,060	1,645	3,045	1,380
80^{th} percentile		8,040	4,274	11,596	2,528	3,516	1,594
95^{th} percentile		8,396	4,331	18,341	2,845	4,018	1,846
Max		8,470	4,860	18,790	3,190	4,230	2,000
Dissolved arsenic* (mg/L)	**0.013**						
Count (n)		12	27	34	27	34	31
% < LOR		0	0	0	4	0	5
Min		0.002	0.002	0.004	0.001	0.001	0.001
20^{th} percentile		0.003	0.002	0.006	0.001	0.002	0.008
50^{th} percentile (Median)		0.005	0.002	0.008	0.003	0.003	0.024
80^{th} percentile		0.006	0.003	0.01	0.003	0.004	0.028
95^{th} percentile		0.006	0.007	0.01	0.004	0.004	0.033
Max		0.006	0.008	0.012	0.006	0.005	0.036
Dissolved copper (mg/L)	**0.0014**						
Count (n)		6	18	15	18	14	27
% < LOR		54	33	57	36	60	25
Min		0.001	0.001	0.001	0.001	0.001	0.001
20^{th} percentile		0.001	0.001	0.001	0.001	0.001	0.001
50^{th} percentile (Median)		0.002	0.002	0.002	0.002	0.001	0.004
80^{th} percentile		0.002	0.004	0.003	0.004	0.002	0.005
95^{th} percentile		0.002	0.004	0.003	0.005	0.002	0.010

		Monitoring Bores					
	Default Guideline or WQO	Bore 1	Bore 2	Bore 3	Bore 4	Bore 5	Bore 6
Max		0.002	0.005	0.004	0.006	0.002	0.017
Dissolved manganese (mg/L)	**1.9**						
Count (n)		12	27	34	28	34	30
% < LOR		0	0	0	0	0	19
Min		2.5	0.001	1.34	0.033	0.09	0.001
20th percentile		2.94	0.002	1.78	0.078	1.41	0.008
50th percentile (Median)		3.36	0.004	2.25	0.46	2.44	0.156
80th percentile		3.71	0.033	4.04	0.642	3.24	0.383
95th percentile		3.93	0.258	7.73	1.14	3.85	0.53
Max		4.17	4.917	11.1	1.42	3.9	0.683
Dissolved zinc (mg/L)	**0.008**						
Count (n)		10	26	34	28	31	31
% < LOR		15	4	0	0	11	8
Min		0.005	0.005	0.007	0.005	0.005	0.006
20th percentile		0.006	0.007	0.012	0.009	0.008	0.065
50th percentile (Median)		0.007	0.013	0.018	0.014	0.014	0.1
80th percentile		0.009	0.018	0.022	0.022	0.019	0.164
95th percentile		0.012	0.026	0.03	0.032	0.025	0.225
Max		0.014	0.039	0.05	0.134	0.041	0.252

Red highlight - 80th percentile exceeds guideline value. Red text – summary statistics recalculated

* 20th and 80th percentile of scheduled WQO for Shallow (< 30m) Zone 24 Lower Nogoa groundwaters

Step 3 – Determine site-specific groundwater guidelines.

The proposed site-specific groundwater guidelines are outlined in Table 4 using the approach explained in this section. The comments in the table outline the dataset used to derive the proposed guidelines. The proposed site-specific groundwater guidelines are also presented on the box plots and time series plots for comparison (see Figure 7).

It is often found that there are different trends in the concentrations of different indicators at different bores. Bores with increasing trends need to be investigated before site-specific guidelines can be applied.

Table 4. Proposed site-specific groundwater guidelines for this example

Indicator	Default guideline or WQO	Site-specific guideline	Comment
pH	7.9 - 8.2	7.8 - 8.2	The 80th percentile from site-specific bores is outside the WQO range, but not substantially different from the WQO. The 20/80th percentile of combined data for all bores is 7.8 - 8.2.
Electrical Conductivity (µS/cm)	6,908	-	The 80th percentile values for Bore 2, Bore 4, Bore 5 and Bore 6 are significantly less than the WQO. The 80th percentile values for Bore 1 and Bore 3 exceed the WQO and the site-specific values better reflect the local conditions. Individual bores required site-specific values to be applied. Increasing trend at Bore 1, Bore 3, Bore 4 and Bore 6 requires investigation.
Electrical Conductivity (µS/cm) – Bore 2	6,908	4,274	The 80th percentile values for Bore 2 was significantly less than the WQO. The 80th percentile of site-specific data for Bore 2 was adopted.
Electrical Conductivity (µS/cm) – Bore 5	6,908	3,516	The 80th percentile values for Bore 5 was significantly less than the WQO. The 80th percentile of site-specific data for Bore 5 was adopted.
Dissolved arsenic (mg/L)	0.013	0.013	The 80th percentile from site-specific bores is less than, but not substantially different from the default guideline. Toxicant default guideline value adopted. Increasing trend at Bore 6 requires investigation.
Dissolved copper (mg/L)	0.0014	0.004	The 80th percentile values from site-specific bores exceed the default guideline and the site-specific values better reflect the local conditions. The 80th percentile of combined data for all bores adopted. Bore 6 may require investigation.
Dissolved manganese (mg/L)	1.9	1.9	The 80th percentile for Bore 2, Bore 4 and Bore 6 is less than but not substantially different from the default guideline. Toxicant default guideline value adopted for Bore 2, Bore 4 and Bore 6.
Dissolved manganese (mg/L)	1.9	3.6	The 80th percentile values for Bore 1, Bore 3 and Bore 5 exceed the default guideline and the site-specific values better reflect the local conditions. Site-specific guideline for Bore 1, Bore 3 and Bore 5, based on the 80th percentile of combined data for these bores.
Dissolved zinc (mg/L)	0.008	0.02	The 80th percentile values from site-specific bores exceed the default guidelines and the site-specific values better reflect the local conditions. The 80th percentile of combined data for all bores, except Bore 6 adopted. Increase in zinc concentration at Bore 6 requires investigation.

Figure 7. Example box and time series plots.

For box plots, boxes represent 1st and 3rd quartile, black dot represents the mean, the black line represents the median (50th percentile) and whiskers represent min and max. The orange line represents the site-specific groundwater guideline and the green line represents the scheduled WQO or default guideline value.

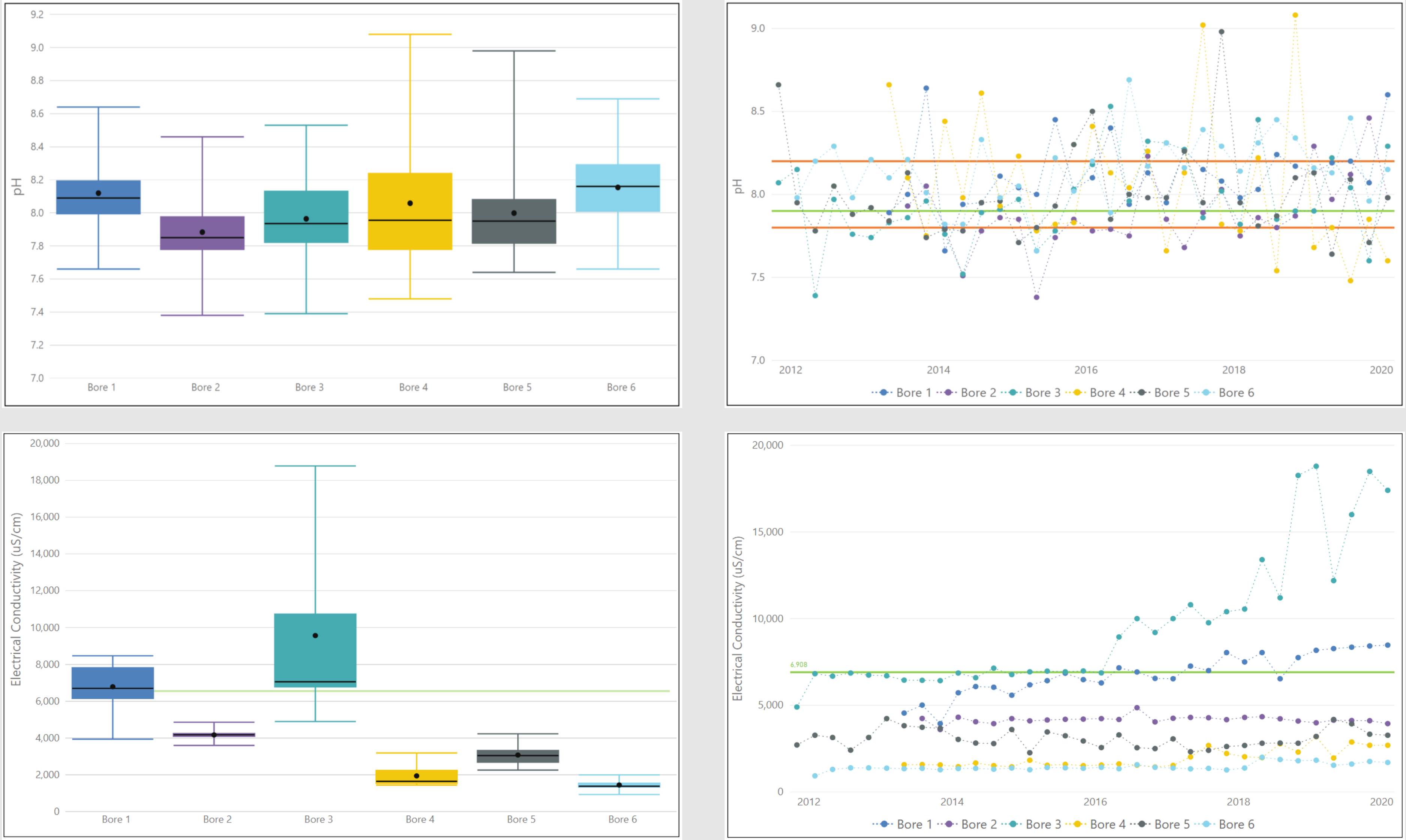


Figure 7 cont. Example box and time series plots.

For box plots, boxes represent 1st and 3rd quartile, black dot represents the mean, the black line represents the median (50th percentile) and whiskers represent min and max. The orange line represents the site-specific groundwater guideline and the green line represents the default guideline value.

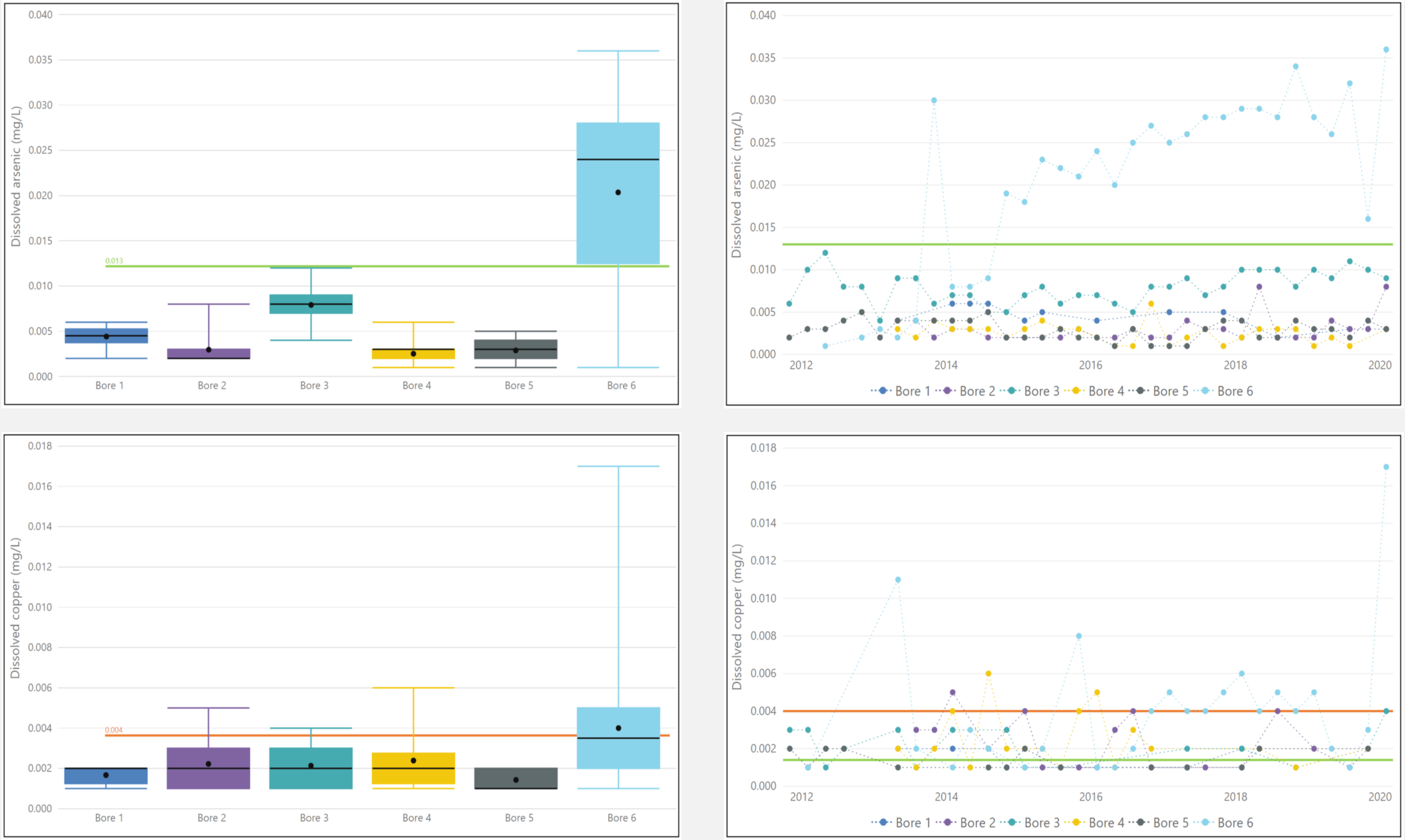
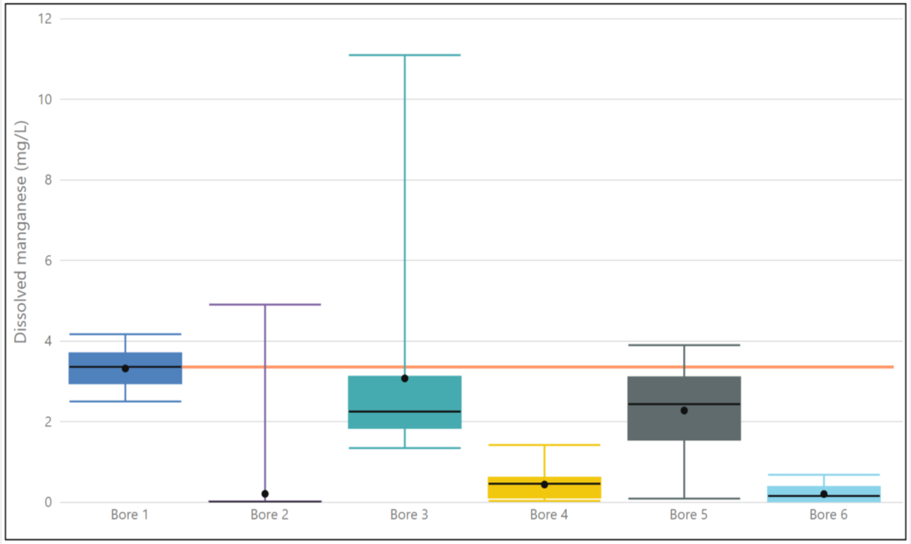
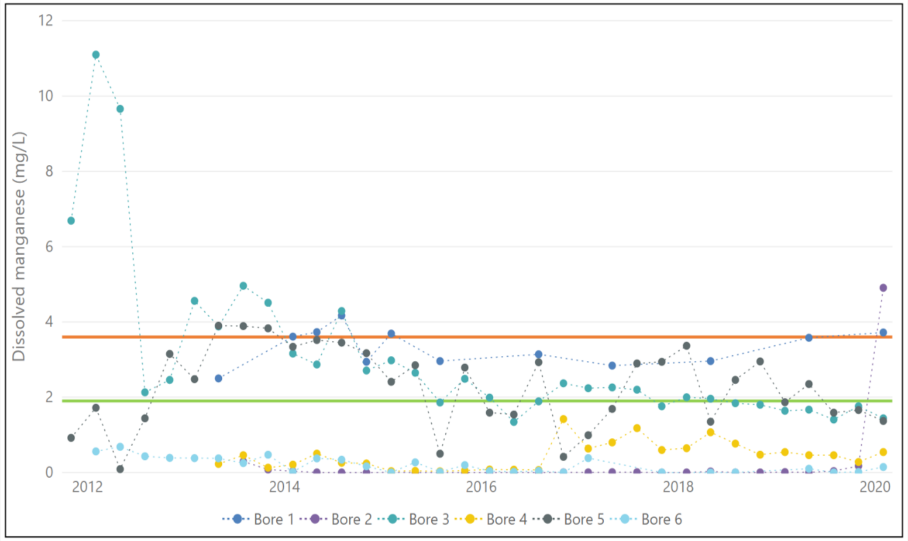
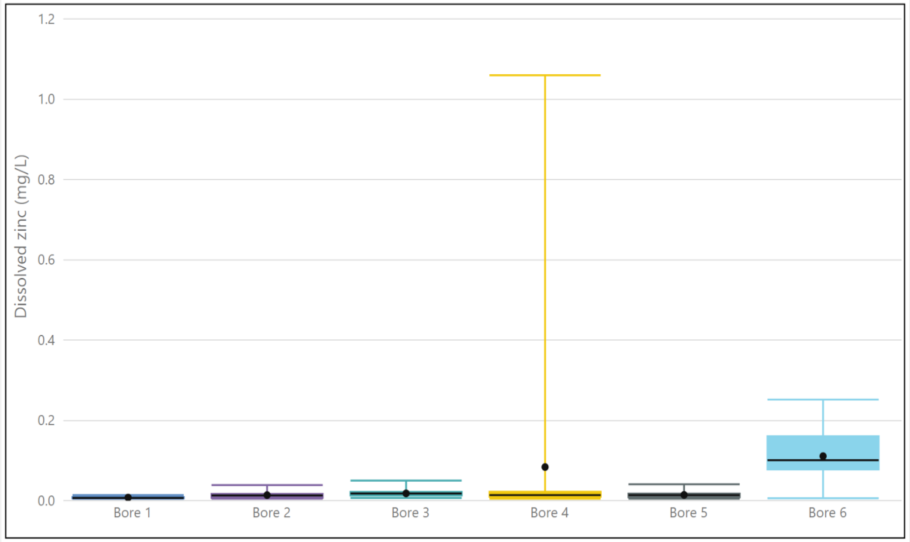


Figure 7 cont. Example box and time series plots

For box plots, boxes represent 1st and 3rd quartile, black dot represents the mean, the black line represents the median (50th percentile) and whiskers represent min and max. The orange line represents the site-specific groundwater guideline and the green line represents the default guideline value.







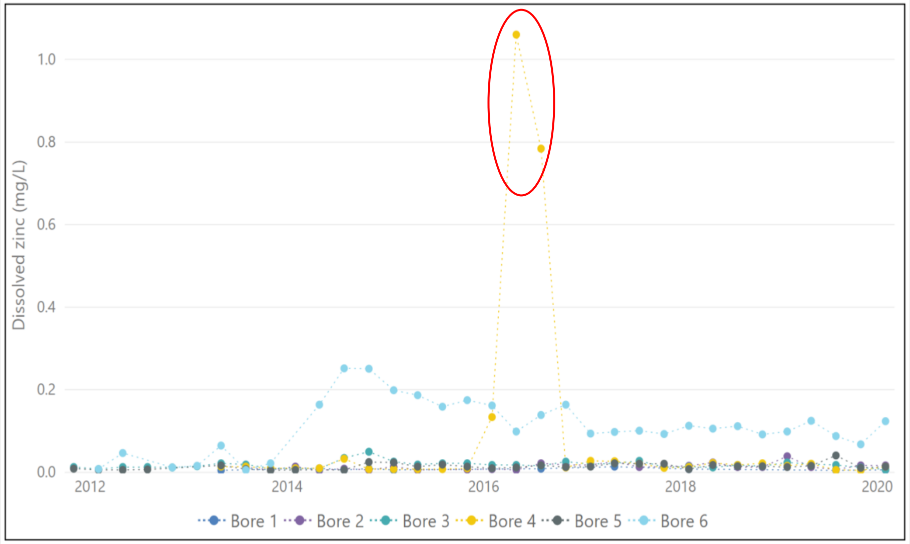


Figure 7 cont. Example box and time series plots

For box plots, boxes represent 1st and 3rd quartile, black dot represents the mean, the black line represents the median (50th percentile) and whiskers represent min and max. The orange line represents the site-specific groundwater guideline and the green line represents the default guideline value.

6. Compliance approaches

An important consideration in the compliance approach is the selection and location of compliance bores. Each compliance bore should be:

- fit for purpose;
- downgradient of potential sources;
- monitor appropriate aquifers and flow pathways associated with potential sources;
- detect emerging issues related to mining activities in a timely manner.

A review of the groundwater monitoring network should be undertaken to determine whether all potential sources, pathways and receptors are adequately monitored.

Site-specific guidelines and WQOs are often used as a basis for deriving water quality triggers or limits specified in environmental approvals used to regulate Environmentally Relevant Activities in Queensland. Trigger values are typically the numerical criteria that, if exceeded, provide an alert of a change that warrants further investigation for the contaminant being measured. Conversely, limit values are typically the highest allowable concentration of a contaminant that is permitted and should not be exceeded.

The choice of whether a trigger or limit is used for an indicator at a particular bore is often determined by the purpose of the groundwater quality monitoring. For example, where monitoring aims to detect the potential migration of contaminants, it would be appropriate to apply limits to contaminants of concern at bores located in proximity to potential contaminant sources. Indicators used to assist with interpretation of data such as hardness or major ions would not typically require limits to be applied for them. A combination of triggers and limits can be applied at one or more bores.

Trigger values should be fit for purpose and conservative enough such that, when applied, they provide an early warning of potential impacts to groundwater. Applying triggers that are set too high may mean that they are not sensitive enough to identify emerging contamination issues. The time to detect and respond to an impact should be considered when setting a trigger to ensure reporting and remedial action can be undertaken before groundwater is contaminated. Another important consideration is monitoring frequency, which may vary depending on the contaminant of concern and aquifer properties (e.g. alluvial aquifers may need to be sampled more frequently than deep aquifers as they are more likely to change over shorter time periods). Conversely, if trigger values are set too low, natural variability may be mistaken for contamination events and result in unnecessary reporting and investigation. However, a trigger value may be set with a higher likelihood of exceeding compared to a limit given they generally do not result in compliance action.

When setting limit values, it is important to consider the location of relevant compliance bores and the particular compliance approach that will be used to ensure the limits are suitable (i.e. they are sufficiently conservative to ensure environmental impact does not occur but do not result in false non-compliances).

A number of compliance approaches have been applied to groundwater quality in environmental approvals in the past, including, but not limited to:

- Comparing single test results to default guidelines, WQO, site-specific guidelines or reference bore monitoring, typically for toxicant indicators;
- Comparing median (50th percentile) of test results to WQOs, site-specific guidelines or reference bore monitoring, typically for physico-chemical indicators;

- Applying control charting using means and standard deviations (see Appendix 4 for more information).

There are a number of challenges associated with each of these compliance approaches. The approach to assess compliance when using a toxicant default guideline is to compare this to a single test value. This may over-estimate the risk as the advice in the ANZG is that the toxicant default guidelines are applied to a 95th percentile of sample results. The default approach to assess compliance when a site-specific value based on an 80th percentile is used, is to compare the median of test data to the site-specific value. Determining the number of samples required and the time period for the calculation of percentiles can be difficult and is often not well defined in the environmental approval conditions, making application of the compliance approach difficult. In addition, background levels in groundwater can naturally be greater than default guidelines.

Unfortunately, it is also often difficult to identify appropriate reference bores. This makes comparison between test and reference bores often not feasible. If reference bores are to be used in the compliance approach, it needs to be determined whether the selected reference bores are appropriate. Reference bores should be in the same aquifer as compliance bores, be upgradient of potential sources of contamination, have a similar ionic composition and concentration of indicators before the activity has commenced. If the ionic composition analyses suggest that the reference bores do not match the compliance bore ionic composition, then an alternative approach for assessing impacts to groundwater quality may be appropriate.

For the reasons explained above, it is recommended that “interpretation” bores, rather than reference bores, be included in monitoring networks. Interpretation bores should be selected such that they are not likely to be impacted by mining activities. As a result, they should represent localised natural variability or trends in water quality and can be used for comparative purposes to help interpret the water quality results from compliance bores. No triggers or limits are required for interpretation bores.

In addition to compliance and interpretation bores, other monitoring bores may also be included in the compliance approach to manage onsite operations and determine the effectiveness of management actions.

Control charts have been used as a compliance approach in relation to groundwater regulation. Control charts typically utilise the mean as the basis of trend analysis, along with upper and lower control limits, often defined as standard deviations (σ) from the mean, with control limits of mean + σ, mean + 2σ and mean + 3σ. This approach relies on a substantial number of samples to adequately represent the true population and the data being normally distributed, which is rarely the case for environmental data, and results in poor regulation and unnecessary reporting and investigations.

All possible effort should be made to avoid false non-compliances. When there is significant variability, uncertainty or likelihood of errors in the water quality data, it is generally not recommended that a limit be applied to a single test sample result. For example, if an 80th percentile value is applied as the limit, it is likely that 20% of individual measurements will be greater than the limit (or lower than the 20th percentile), which results in a high likelihood of non-compliance due to natural variability.

Given the significant limitations of the approaches discussed above, a new approach is recommended for assessing compliance of groundwater quality and is detailed in the following section.

6.1 Recommended Compliance Approach

The compliance approaches recommended in this guideline for groundwater quality assessment are based on comparing a number of consecutive sample tests at compliance bores to a limit based on percentile calculations. This approach is aimed at reducing the probability of a false positive, or what is referred to as a false non-compliance in regards to regulation while ensure approaches are sufficiently sensitive to detect potential impacts. The approach is based on combining ANZECC 2000 methodology, adapted control charting approaches, and statistics around applying one or two limit types (or trigger).

Suggested applications of this approach are -

- Single Limit (95th percentile) – 3 consecutive test samples exceed the limit. If a toxicant default guideline (ANZG 2018) is adopted, this can be applied as the limit.
- Limit A (80th percentile) and Limit B (95th percentile) – 5 consecutive test samples exceed the Limit A and 3 consecutive test samples exceed the Limit B. If toxicant default guideline (ANZG 2018) is adopted, it should be applied as a Limit B not Limit A.

Limit (and trigger) values are derived from default guidelines or site-specific data following the steps described in Section 4 and 5 for data analysis and deriving site-specific groundwater guidelines. In general, the first limit (also referred to as Limit A) is the 80th percentile of site-specific groundwater data. Limit A is applied to 5 consecutive test samples. The second limit is the 95th percentile (also referred to as Limit B) of site-specific groundwater data and is applied to 3 consecutive test samples. Each limit is a set value for the period of the compliance and would only be changed if the initial site-specific data used to derive the values was inadequate.

If a toxicant default guideline was adopted in Section 5 Step 3, and is used as a limit, in line with the precautionary principle it should be applied as a Limit B not a Limit A. Where a default reference-based guideline is adopted as a limit, such as for pH, this would ideally be applied as Limit A. However, site-specific data is likely to be more relevant and may be used.

Once the limits (and triggers) are determined they are then tested against the corresponding number of consecutive samples of test samples. The theoretical probability of false exceedances can be determined as shown in Table 5 for the two limit types. Limit A, which is based on an 80th percentile value and 5 consecutive samples, has a probability of false compliance of 0.032%. Similarly, Limit B, which is based on a 95th percentile and 3 consecutive samples, has a probably of false compliance of 0.0125%. The probability of a false exceedance in both cases is very low (< 0.05%) which is deemed acceptable for the purposes of compliance testing.

Both limits provide a statistically robust method for detecting change. Given the requirement to test five consecutive samples, Limit A is seen as a method for detecting gradual change over the medium term (i.e. 1 to 2 years). Limit B**,** which is tested over three consecutive samples, will potentially provide detection of change over a shorter term, particularly if the change is significant.

The period of time required to detect change for each limit will depend on sampling frequency. Based on the approach outlined above, the period of time to detect change could range from 3 months through to a year. In some circumstances (for example, where there are high ecological values), the sampling frequency and the number of consecutive samples applied in this approach may be varied depending on site-specific conditions.

Table 5. The theoretical probability (as a percentage) of false exceedances when comparing consecutive test site sample to Limit A and Limit B values.

Number of consecutive exceedances	1	2	3	4	5
Time (months)[1]	0	3	6	9	12
Limit A - 80th Percentile[2]	20%	4%	0.8%	0.16%	**0.032%**
Limit B - 95th Percentile[2]	5%	0.25%	**0.0125%**		

[1] assumes quarterly sampling; [2] the probability of a false exceedance is less than the theoretical probability percentage provided.

In general, any single sample exceedance should be checked for sampling or laboratory error and may include resampling. It is recommended that the requirement to resample is included in the QA/QC procedures and may be specified in the EA conditions in some circumstances.

Example – Limit A and Limit B approach

The following limits and compliance approach were derived for the examples discussed in Sections 3, 4 and 5.

As identified in Section 5, increasing trends were observed at some bores. These bores may be impacted and require investigation. If limits are required for these bores, separate limits can be applied to ensure no increase in the concentration of contaminants.

The proposed site-specific groundwater guidelines are outlined in Table 4. For each indicator, the data analysis and site-specific groundwater guidelines in Section 3 and 4 were reviewed and the proposed Limit A and Limit B are outlined in Table 6.

- pH – the 20th percentile for Bore 2. Bore 3, Bore 4 and Bore 5 was less than the WQO 20th percentile and the 80th percentile for Bore 1, Bore 4 and Bore 6 was greater than the WQO 80th percentile. The 80th percentile from site-specific bores is outside the WQO range, but not substantially different from the WQO. For limits, site-specific 20th/80th and 5th/95th percentiles which better reflect the background condition were used.
- Electrical conductivity – the 80th percentile of EC at Bore 1 and Bore 3 was greater than WQO 80th percentile and the 80th percentile values for Bore 2, Bore 4, Bore 5 and Bore 6 are significantly less than the WQO. An increasing trend in EC at Bore 1, Bore 3, Bore 4 and Bore 6 is observed. Site-specific 80th and 95th percentiles for each bore were calculated and applied as Limit A and Limit B.
- Arsenic – An increasing trend in the concentration of dissolved arsenic was observed at Bore 6. All measurements from the other monitoring bores were below the toxicant default guideline. Therefore, the toxicant default guideline is applied as a Limit B for all bores except Bore 6. Site-specific values are applied for Bore 6.
- Copper – the 80th percentile of dissolved copper at all monitoring bores was greater than the toxicant default guideline. Site-specific 80th and 95th percentiles were calculated and applied as Limit A and Limit B. Separate limits were applied to Bore 6.

- Manganese – The 80th percentile of dissolved manganese at Bore 2, Bore 4 and Bore 6 is less than the toxicant default guideline, however, the 80th percentile at Bore 1, Bore 3 and Bore 5 was greater than the toxicant default guideline. Therefore, the toxicant default guideline is applied as a Limit B for Bore 2, Bore 4 and Bore 6 and site-specific values are applied for Bore 1, Bore 3 and Bore 5.
- Zinc – The 80th percentile of dissolved zinc at other monitoring bores was greater than the toxicant default guideline. This may suggest that the concentration of dissolved zinc is naturally greater than the toxicant default guideline. Site-specific 80th and 95th percentiles were calculated and applied as Limit A and Limit B. Separate limits were applied to Bore 6.

In this example, the groundwater quality must comply with the limits specified in Table 6, for each bore and indicators listed.

Limit A relates to any five (5) consecutive sampling occasions while Limit B relates to any three (3) consecutive sampling occasions. Note that consecutive sampling occasions means any number of sampling results obtained sequentially regardless of frequency.

Table 6. Groundwater quality limits for this example

Indicator	Limit type	Limit A	Limit B	Comment
pH	Range	7.8 - 8.2	7.7 - 8.5	Site-specific 80th and 95th percentiles of combined data for all bores used
Electrical Conductivity (µS/cm) Bore 1	Maximum	8,040	8,396	Site-specific 80th and 95th percentiles for Bore 1 was used
Electrical Conductivity (µS/cm) Bore 2	Maximum	4,274	4,331	Site-specific 80th and 95th percentiles for Bore 2 was used
Electrical Conductivity (µS/cm) Bore 3	Maximum	11,596	18,341	Site-specific 80th and 95th percentiles for Bore 3 was used
Electrical Conductivity (µS/cm) Bore 4	Maximum	2,528	2,845	Site-specific 80th and 95th percentiles for Bore 4 was used
Electrical Conductivity (µS/cm) Bore 5	Maximum	3,516	4,018	Site-specific 80th and 95th percentiles for Bore 5 was used
Electrical Conductivity (µS/cm) Bore 6	Maximum	1,594	1,846	Site-specific 80th and 95th percentiles for Bore 6 was used
Dissolved arsenic (mg/L) All except Bore 6	Maximum		0.013	Default toxicant guideline used for Limit B
Dissolved arsenic (mg/L) Bore 6	Maximum	0.023	0.033	Site-specific 80th and 95th percentiles for Bore 6 was used
Dissolved copper (mg/L) All except Bore 6	Maximum	0.003	0.004	Site-specific 80th and 95th percentiles of combined data for all bores, except Bore 6 was used
Dissolved copper (mg/L) Bore 6	Maximum	0.005	0.008	Site-specific 80th and 95th percentiles for Bore 6 was used
Dissolved manganese (mg/L) Bore 2, Bore 4 and Bore 6	Maximum		1.9	Default toxicant guideline used for Limit B

Indicator	Limit type	Limit A	Limit B	Comment
Dissolved manganese (mg/L) Bore 1, Bore 3 and Bore 5	Maximum	3.6	4.6	Site-specific 80th and 95th percentiles of combined data for Bores 1, Bore 3 and Bore 5 was used
Dissolved zinc (mg/L) All except Bore 6	Maximum	0.02	0.03	Site-specific 80th and 95th percentiles of combined data for all bores, except Bore 6 was used
Dissolved zinc (mg/L) Bore 6	Maximum	0.16	0.23	Site-specific 80th and 95th percentiles for Bore 6 was used

6.2 Further consideration for compliance assessment

Importantly, the site-specific groundwater guidelines, triggers, limits and compliance approach should be assessed to determine if they meet their intended purpose. This should be done on a bore by bore basis. Applying site-specific groundwater guidelines and a compliance approach without an evaluation can lead to significant compliance issues. Triggers that are set too high may not be sensitive enough to identify current or future contamination. The time to detect an impact should be considered to ensure compliance action is possible before groundwater is contaminated. Conversely, if triggers and limits are set too low, natural variability may be mistaken for contamination events. Where compliance includes assessment of consecutive samples, this should be considered when assessing if the site-specific guideline values are appropriate. The number of occasions where proposed trigger or limit values are exceeded for all available historical data (using the suggested number of consecutive samples), should be included in the evaluation.

Typically, where ongoing groundwater quality results (e.g. >8 samples) from compliance bores do not comply with the existing triggers or limits in an environmental approval, then these exceedances should be investigated. If an investigation finds that the existing triggers or limits do not reflect the natural groundwater quality, then an alternative approach, as proposed in this guideline in section 5, may be required. However, where contamination is found to be present as a result of the activity, prescribed limits should only be changed following appropriate investigation and implementation of mitigation and remedial measures to prevent further impacts. A change in approach may also be required if the groundwater quality of water collected from compliance bores is contaminated by historic (pre-approval) activities.

7. Conclusion

Detailed assessment is required to consider the adoption of site-specific groundwater guidelines and the compliance approach recommended in this guideline. The minimum information requirements to undertake such an assessment includes a description of site and bore characteristics, descriptive statistics of monitoring and quality assurance data, a representation of major ion composition (e.g. piper plot), time series analyses and box plots for all relevant water quality indicators at all monitoring bores and an evaluation of the proposed compliance approach. The list of information typically required includes:

- Identify EVs for groundwater and relevant default guidelines, WQOs and existing environmental approval conditions including any groundwater-related limits or triggers (if applicable).
- Describe site and bore characteristics.
- Minimum groundwater quality monitoring data requirements satisfied. Check sampling techniques and data integrity including QA/QC of the data.
- Table of descriptive statistics provided for each bore (date range, count, mean, minimum, maximum, median, 5th (for pH), 20th, 80th and 95th percentiles).
- Box plots for all bores for each indicator and time series plots of all available groundwater quality data.
- Investigation of increasing trends and peaks in the water quality data.
- Distribution of major anions and cations in each bore (e.g. piper diagram).
- If required, site-specific groundwater guidelines and compliance approach investigated.
- Testing and evaluation of proposed compliance approach.

8. References

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.

www.waterquality.gov.au/anz-guidelines

ANZECC & ARMCANZ (2000). Australian and New Zealand Guidelines for Fresh and Marine Water Quality – Volume 1: The Guidelines. Australian and New Zealand Conservation Council, Agriculture Resource Management Council of Australia and New Zealand (ANZECC & ARMCANZ). Commonwealth of Australia, Canberra.

https://www.waterquality.gov.au/sites/default/files/documents/anzecc-armcanz-2000-guidelines-vol1.pdf

Australian Government (2013). Guidelines for Groundwater Quality Protection in Australia - 2013. National Water Quality Management Strategy. Australian Government, Canberra.

https://www.environment.gov.au/system/files/resources/2c1e1e7c-e57b-4971-8b41-adc6b1b2eb73/files/nwqms-groundwater-quality-protection-guidelines.pdf

Barnett B, Townley LR, Post V, Evans RE, Hunt RJ, Peeters L, Richardson S, Werner AD, Knapton A and Boronkay A. (2012). Australian Groundwater Modelling Guidelines. Waterlines report, National Water Commission, Canberra.

Centre for Groundwater Studies (2001). Australian Groundwater School. Fundamentals of Groundwater Science, Technology and Management. Fifth Edition.

Cornwell, D, Brown R and McTigue, N. (2015) Controlling lead and copper rule water quality parameters. *Journal of the American Water Works Association*. 107(2): E86 – E96.

http://dx.doi.org/10.5942/jawwa.2015.107.0011

DES (2018). Monitoring and Sampling Manual: Environmental Protection (Water) Policy. Brisbane: Department of Environment and Science Government.

https://environment.des.qld.gov.au/water/monitoring/sampling-manual/pdf/monitoring-sampling-manual-2018.pdf

DES (2018). Technical guideline - Licensing - Wastewater release to Queensland waters (ESR/2015/1654 Version 2.01). Department of Environment and Science (DES), Queensland Government, Brisbane.

https://environment.des.qld.gov.au/assets/documents/regulation/pr-gl-wastewater-to-waters.pdf

DEHP (2013). Queensland Water Quality Guidelines, Version 3, ISBN 978-0-9806986-0-2. Department of Environment and Heritage Protection (DEHP), Queensland Government, Brisbane.

DEHP (2013). Baseline assessment guideline. EM1088. Queensland Government, Brisbane

https://www.ehp.qld.gov.au/assets/documents/regulation/rs-gl-baseline-assessments.pdf

DEHP (2016). Bore assessment guideline under section 413 of the Water Act 2000 (EM1178).

http://www.ehp.qld.gov.au/assets/documents/regulation/rs-gl-bore-assessment.pdf

Doody TM, Hancock PJ, Pritchard JL (2019). Information Guidelines Explanatory Note: Assessing groundwater-dependent ecosystems. Report prepared for the Independent Expert Scientific Committee on Coal Seam Gas and Large Coal Mining Development through the Department of the Environment and Energy, Commonwealth of Australia. 2019.

http://www.iesc.environment.gov.au/system/files/resources/422b5f66-dfba-4e89-adda-b169fe408fe1/files/information-guidelines-explanatory-note-assessing-groundwater-dependent-ecosystems.pdf

DSITI (2015). Groundwater Dependent Ecosystem Conceptual Modelling: Technical Specifications, Department of Science, Information Technology and Innovation (DSITI), Queensland Government, Brisbane.

Fensham, R J and R J Fairfax (2003). Spring wetlands of the Great Artesian Basin, Queensland, Australia. *Wetlands Ecology and Management*. Volume 11: 343-362.

Gibbons, RD (1999). Use of combined Shewhart-CUSUM control charts for ground water monitoring applications. *Groundwater* 37(5): 682-691.

Grant, E (1964). Statistical Quality Control (3rd ed.) McGraw-Hill, New York.

Howard, D (2003). The basics of statistical process control and process behaviour charting. A User's Guide to SPC. Published by Management-NewStyle, Chislehurst, Kent, BR7 5NB, England.

ITRC (Interstate Technology & Regulatory Council). (2013). Groundwater Statistics and Monitoring Compliance, Statistical Tools for the Project Life Cycle. GSMC-1. Washington, D.C.: Interstate Technology & Regulatory Council, Groundwater Statistics and Monitoring Compliance Team.

http://www.itrcweb.org/gsmc-1/.

Lee PH, Huang YH, Kuo TI, Wang CC. (2013). The effect of the individual chart with variable control limits on the river pollution monitoring. *Quality & Quantity* 47(4):1803–1812.

Mann, R. M. and Dunlop, J. E. (2015). Surface and groundwater quality triggers for the management of the Mount Morgan Mine. Water Assessment and Systems Team, DSITI prepared for DEHP and DNRM.

Maurer D, Mengel M, Robertson G, Gerlinger T, Lissner A. (1999). Statistical process control in sediment pollutant analysis. *Environmental Pollution* 104(1):21–29.

National Environment Protection Council. (2011). Schedule B1. Guideline on investigation levels for soil and groundwater.

http://www.nepc.gov.au/system/files/resources/93ae0e77-e697-e494-656f-afaaf9fb4277/files/schedule-b1-guideline-investigation-levels-soil-and-groundwater-sep10.pdf

National Environment Protection Council. (1999). Schedule B6. Guideline Guideline on Risk Based Assessment of Groundwater Contamination.

https://www.legislation.gov.au/Details/F2008B00713/2e805a15-e9eb-4321-8cd2-9024073c5888

NHMRC (2008). Guidelines for Managing Risks in Recreational Waters. National Health and Medical Research Council, Australian Government, Canberra

http://www.nhmrc.gov.au/_files_nhmrc/publications/attachments/eh38.pdf

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

https://www.nhmrc.gov.au/_files_nhmrc/file/publications/nhmrc_adwg_6_february_2016.pdf

Paroissin C, Penalva, L, Petrau, A and Verdier, G. (2016) New control chart for monitoring and classification of environmental data. *Environmetrics* 27: 182-193.

Richardson S, Irvine, E, Froend, R, Boon, P, Barber, S. and Bonneville, B. (2011a). Australian groundwater-dependent ecosystem toolbox part 1: assessment framework, Waterlines report, National Water Commission, Canberra

http://www.bom.gov.au/water/groundwater/gde/GDEToolbox_PartOne_Assessment-Framework.pdf

Richardson S, Irvine, E, Froend, R, Boon, P, Barber, S. and Bonneville, B. (2011b). Australian groundwater-dependent ecosystems toolbox part 2: assessment tools, Waterlines report, National Water Commission, Canberra

http://www.bom.gov.au/water/groundwater/gde/GDEToolbox_PartTwo_Assessment-Tools.pdf

Shaban, M (2014). Drainage water reuse: State of control and process capability evaluation. *Water, Air and Soil Pollution* 225:2168. doi:10.1007/s11270-014-2168-6.

Shewhart, W.A. (1931). Economic control of quality of manufactured product. New York: Van Nostrand, 1931. Reprinted by American Society for Quality, Milwaukee, 1980.

State of Queensland (2011). Conceptual model Case Study Series. Eulo springs super-group. Updated 2013.

http://wetlandinfo.ehp.qld.gov.au/resources/static/pdf/resources/tools/conceptual-model-case-studies/cs-eulo-supergroup-22-april-2013.pdf

Sundaram, B. Feitz, A. J. de Caritat, P. Plazinska, A Brodie, R. Coram, J and Ransley, T. (2009). Groundwater Sampling and Analysis – A Field Guide. GeoScience Australia, Record 2009/27.

http://www.ga.gov.au/webtemp/image_cache/GA15501.pdf

Todd, D K and Mays, L W (2005). Groundwater Hydrology. Third Edition. John Wiley and Sons, Inc.

U.S. Environmental Protection Agency (EPA) (2009). Statistical analysis of groundwater monitoring data at RCRA facilities. Unified Guidance. EPA 530/R-09-007. Office of Resource Conservation and Recovery.

Wetland*Info* (2014). Groundwater dependent ecosystems. Department of Environment and Heritage Protection, Queensland, viewed 14 April 2016,

http://wetlandinfo.ehp.qld.gov.au/wetlands/resources/training/15-gde-ecology/

9. Acronyms

ANZG	Australian and New Zealand Guidelines
ARMCANZ	Agriculture and Resource Management Council of Australia and New Zealand
ANZECC	Australian and New Zealand Environment Conservation Council
CUSUM	Cumulative sum
DEHP	Department of Environment and Heritage Protection
DES	Department of Environment and Science
DSITI	Department of Science, Information Technology and Innovation
EA	Environmental Authority
EPP Water	Environmental Protection (Water and Wetland Biodiversity) Policy 2019
ERA	Environmental Relevant Activity
EV	Environmental Value
GDE	Groundwater dependent ecosystem
ITRC	Interstate Technology & Regulatory Council
LoR	Limit of Reporting
QA/QC	Quality Assurance/Quality Control
QWQG	Queensland Water Quality Guideline
U.S. EPA	United States Environmental Protection Agency
WQO	Water Quality Objective

10. Definitions

Alluvium	Loose, unconsolidated soil or sediment which has been eroded or reshaped by water
Aquifer	Zones within the underlying geology of an area that are saturated with groundwater
Aquitard	Zone within the earth that restricts the flow of groundwater from one aquifer to another. Aquitards comprise layers of either clay or non-porous rock with low hydraulic conductivity.
Aquiclude	Zone within the earth that is impermeable
Aquatic ecosystems	Any watery environment from small to large, from pond to ocean, in which plants and animals interact with the chemical and physical features of the environment. The animals, plants and micro-organisms that live in water, and the physical and chemical environment and climatic regime in which they interact. It is predominantly the physical components (for example light, temperature, mixing, flow, and habitat) and chemical components (for example organic and inorganic carbon, oxygen, nutrients) of an ecosystem that determine what lives and breeds in it, and therefore the structure of the food web. Biological interactions (for example grazing and predation) can also play a part in structuring many aquatic ecosystems.
Baseline data	Collected before a development begins.
Box plot	Graphic of selected descriptive statistics at a monitoring point. Descriptive statistics represented in the plot can include mean, median, upper and lower quartiles, percentile data, and max/min.

Catchment	The total watershed draining into a river, creek, reservoir or other body of water. The limits of a given catchment are the heights of land (such as hills or mountains) separating it from neighbouring catchments. Catchments can be made up of smaller sub-catchments.
Contaminant	Defined in section 11 of the *Environmental Protection Act 1994*.
Control chart	Graphical plots of measurements over time with upper and lower control limits.
Control limit	Horizontal lines drawn on a statistical process control chart, usually at a distance of ±3 standard deviations of the plotted statistic from the statistic's mean.
Criteria	Criteria by which decisions will be made as a result of monitoring for potential impacts.
Default Guideline value	A guideline value recommended for generic application in the absence of a more specific guideline value (e.g. a site-specific guideline value) in the Australian and New Zealand Guidelines for Fresh and Marine Water Quality.
Environmental harm	As defined in the section 14 of the *Environmental Protection Act 1994*.
Environmental value (EV)	Particular values or uses of the environment that are important for a healthy ecosystem or for public benefit, welfare, safety or health and that require protection from the effects of pollution, waste discharges and deposits. Specified in the EPP Water (Part 3, Section 6), and includes the qualities of a water that make it suitable for supporting aquatic ecosystems and human water uses.
Environmentally relevant activity (ERA)	A resource activity and/or an activity prescribed under a regulation as an ERA (where a contaminant will or may be released into the environment when the activity is carried out and the release will or may cause environmental harm). Schedule 2 of the EP Regulation lists prescribed ERAs.
Factors	Represent causes of the variation seen in the data, e.g. position of the bore, time of year (seasonality), time (date of sample), etc. They are defined as independent experimental variables. Factors can be quantitative (e.g. depth of bore) or qualitative (e.g. location of the bore).
General environmental duty	As defined in Section 319 the *Environmental Protection Act* (1994).
Groundwater	Water stored underground in rock crevices and in the pores of geologic materials that make up the earth's crust; water that may supply springs and wells
Groundwater dependent ecosystem (GDE)	Ecosystems which require access to groundwater on a permanent or intermittent basis to meet all or some of their water requirements so as to maintain their communities of plants and animals, ecological processes and ecosystem services
Guideline (water quality) value	A measurable quantity (e.g. concentration) or condition of an indicator for a specific environmental value below which (or above which, in the case of stressors such as pH, dissolved oxygen and many biodiversity responses) there is considered to be a low risk of unacceptable effects occurring to that environmental value.
Inter-bore	Comparisons between two monitoring bores separated spatially
Intra-bore	Comparison of measurements over time at one monitoring bore
Limit	The maximum concentration of a contaminant which can be released to waters.
Mann-Kendall Test	The Mann-Kendall test is a nonparametric test for monotonic trends, such as concentrations that are either consistently increasing or decreasing over time.
Mean	Also known as the average, the arithmetic mean is the sum of a set of data divided by the number of data points in the set.

Median	Also known as 50th percentile of a set of data is the middle value. Fifty percent of the values in a dataset will be above and below the median.
Normal distribution	Symmetric distribution of data (bell-shaped curve), the most common distribution assumption in statistical analysis
Outliers	Outliers are extreme values of a data set. An observation that lies well outside the common spread of data may be an indication of some infrequent but important phenomenon, and as such should be included in the analysis. An outlier should only be removed if there are strong reasons to believe it is the result of erroneous procedure in sampling, laboratory analysis or reporting.
Percentiles	Division of a frequency distribution into one hundredths. A percentile is a measure of the spread of a group of data. For example, the 25th percentile of a set of data is a value below which 25 percent of observations are found.
Pooled	Groundwater samples from more than one sampling point.
Populations	For the purposes of this guideline, samples from two bores are from the same population if the bores receive water from the same aquifer and are subject to the same variations in water quality due to changes in seasonal rainfall, etc.
Queensland Water Quality Guidelines	The Queensland Water Quality Guidelines (2009), as amended, prepared by the former Department of Environment and Resource Management.
Significant	'Significant' in a statistical sense does not necessarily mean 'major'. A small difference between groups of data can be significant if it is sufficient for the statistical test to reject the *null hypothesis*.
Site-specific guideline value	A guideline value that is relevant to the specific location or conditions that are the focus of a given assessment or issue.
Standard deviation	Standard deviation is a measure of the variation or spread of a set of data. It is equal to the square root of the variance.
Time series plot	A graphic of data collected at regular time intervals, where measured values are indicated on one axis and time indicated on the other. This method is a typical exploratory data analysis technique to evaluate temporal, directional, or stationarity aspects of data.
Trigger values	These are the concentrations (or loads) of the key performance indicators measured for the ecosystem, below which there exists a low risk that adverse biological (ecological) effects will occur. They indicate a risk of impact if exceeded and should 'trigger' some action, either further ecosystem specific investigations or implementation of management/remedial actions.
Water	The whole or any part of surface water or groundwater, tidal or non-tidal, and including any river, stream, creek, lake, lagoon, swamp, wetland, spring, unconfined surface water, natural or artificial watercourse, dam, tidal waters (estuarine, coastal and marine waters to the limit of Queensland waters) and underground or artesian water.
Water quality indicator (for an EV)	Defined in the EPP Water as a property that can be measured or decided in a quantitative way. Examples of water quality indicators include physical indicators (for example temperature), chemical indicators (for example nitrogen, phosphorus, metals) and biological indicators (for example macroinvertebrates, seagrass and fish).

Water quality objective (WQO)	A numerical concentration limit or narrative statement that has been established to support and protect the designated uses of water at a specified site. It is based on scientific criteria or water quality guidelines but may be modified by other inputs such as social, cultural and economic constraints (EPP Water Part 4, Sections 12 and 13). WQOs are specified in the EPP Water (Part 4, Section 11).

11. Appendices

Appendix 1 – Understanding groundwater systems

Additional information regarding aquifers, groundwater flow and groundwater – surface water interactions is provided in this appendix.

Aquifers and aquitards

The process by which water moves into an aquifer is called recharge. Recharge commonly occurs when an aquifer is close to the surface and is defined as the amount of water that fills an aquifer over a given period of time. Groundwater can also be lost from an aquifer or discharged via flow to surface water features such as rivers, streams, springs, wetlands and oceans, as well as to other aquifers. Groundwater losses from aquifers also occur through evaporation from the water table, transpiration by vegetation, and groundwater pumping.

Aquifers can range in the level of confinement or the degree to which groundwater is bounded by the confining layer. An aquitard has the potential to offer some protection from surface contamination (Centre for Groundwater Studies 2001).

Aquifers may occur at various depths. A shallow aquifer is typically unconfined and does not have an aquitard between it and the surface. Aquifers closer to the surface are more likely to be used for water supply and irrigation and to be topped up by local rainfall.

Aquifers can also be categorised according to the physical properties of the rock, such as unconsolidated or surficial (e.g. sand), sedimentary (e.g. sandstone) or fractured rock (e.g. fractures in sandstone or granite). Unconsolidated aquifers commonly occur in alluvium (loose, unconsolidated soil or sediment that has been eroded or reshaped by water) and exist mostly within 100 m of the surface. They are highly productive aquifers with water stored in inter-granular spaces and can exhibit marked seasonal fluctuations in water levels in response to recharge.

Sedimentary aquifers are confined aquifers that are bounded by an impermeable base (aquiclude) and an aquitard of low permeability above. Flow paths in these aquifers may extend several thousand metres in depth and hundreds to thousands of kilometres in distance. Closed sedimentary aquifers (basins) in arid zones often have salinities greater than 1500 mg/L[1] (e.g. Great Artesian Basin) (Centre for Groundwater Studies 2001).

Fractured rock aquifers exist where groundwater fills voids that have been formed by fractures, fissures and faults. The depth to the water bearing zone is generally less than 200 m below the surface. The recharge in these aquifers is local and seasonal fluctuations in water level often occur. Fractured rock aquifers are heterogenous in nature and unevenly distributed within the rock formation. Estimating the rate of groundwater flow within these aquifers is difficult. Therefore, it can also be challenging to predict the potential for large drawdown or the rate of movement of contaminants within groundwater systems. In some systems, where there are high water velocities through fractures, there is the potential for contaminants to travel large distances very rapidly.

[1] measured as total dissolved solids

Groundwater flow

Hydrogeological processes, including porosity and storativity, can operate at a local and regional scale. Local processes can exhibit large temporal variability, whereas regional processes tend to be temporally stable. The hydraulic properties of an aquifer, including hydraulic conductivity, density of water, fluid viscosity and laminar flow, are important in determining the direction and speed of groundwater flow. These processes are further explained in Table 7. The fluid pressure, compressibility and surface tension may also be important within the groundwater hydraulics.

The water level within an aquifer can fluctuate because of a number of factors, including evapotranspiration, atmospheric pressure, rainfall, tides, earthquakes and land subsidence. Unconfined aquifers with a water table close to the ground surface may exhibit diurnal fluctuations. These diurnal fluctuations may be due to evaporation and/or transpiration (Todd and Mays 2005).

Table 7: Groundwater hydrogeological and hydraulic properties

Properties	Description
Porosity	The porosity is the percentage of water that can be stored in a saturated geologic medium. Secondary porosity can occur if fracturing occurs within the aquifer.
Storativity	The storativity of an aquifer gives a measure of the volume of water that is released or added to the storage per unit area. The measure is different between confined and unconfined aquifers.
Specific yield	The specific yield of an unconfined aquifer is the volume of water that will drain under gravity from a unit volume of the aquifer.
Fluid viscosity	The density and viscosity of water, as a function of temperature and dissolved salt content, is of utmost importance in the calculation of hydraulic head and water pressure. Fluid viscosity is important as it influences the rate of fluid movement through a porous media. The more viscous a fluid the greater the shear stress at any given velocity gradient. At a certain critical velocity, the viscosity is insufficient to dampen turbulent flow, and mixing within an aquifer will occur.
Laminar flow	Laminar flow is where fluid particles move along parallel paths in layers or laminae.
Hydraulic conductivity	A measure of hydraulic conductivity, describes the ease with which a fluid is transported through the media. Hydraulic conductivity can be: • homogeneous and isotropic which is independent of the position in the aquifer and the direction of flow. • heterogeneous and anisotropic, where the hydraulic conductivity varies in different directions both vertically and horizontally.
Hydraulic head	Within confined aquifers the hydraulic head lies above the base of the upper confining layer (aquitard or aquiclude). This is called the potentiometric or piezometric surface (Figure 1).
Water level	The upper surface of the zone of saturation of an unconfined aquifer is referred to as the water table, and it appears as the level at which water stands in a bore penetrating the aquifer (Figure 1).

Groundwater – surface water interaction

A watercourse or wetland may lose or gain water to, or from, an aquifer. Factors such as topography, geology and climate can change the direction and magnitude of these flows.

Streamflow originating from groundwater discharge is referred to as base flow. During periods of rainfall, streamflow may be derived primarily from surface runoff; however, during extended dry periods, all streamflow may be derived primarily from base flow (Todd and May 2005).

Streams and lakes have similar gaining and losing properties; however, lakes can also be flow through systems, with capture zones on their upgradient side and release zones on their downgradient side.

Groundwater dependent ecosystems (GDE) in riverine systems are those that require access to groundwater on a permanent or intermittent basis to meet all or some of their water requirements so as to maintain their communities of plants and animals, ecological processes and ecosystem services, particularly in low flow conditions.

Pumping from an aquifer near a river can dramatically change the amount of groundwater contribution to the baseflow of the watercourse. Conversely, if the groundwater is contaminated as a consequence of anthropogenic activities, groundwater discharge to waterways can have a negative effect on surface water quality. Understanding the extent of groundwater-surface water connectivity is crucial for the sustainable management of the overall water resource.

Springs

Groundwater dependant ecosystems (GDEs) can depend on sub-surface and surface expression of groundwater and can be categorised into three broad types (Australian Government 2013):

- Surface ecosystems that rely on groundwater discharge to rivers, wetlands and springs
- Surface ecosystems that access groundwater from below the water table, e.g. vegetation
- Subterranean aquatic ecosystems, which include stygofauna and troglofauna, in aquifers and caves.

Springs are waterbodies that may support GDEs that depend on the surface expression of groundwater. Changes to groundwater quality have the potential to impact on spring ecosystems. Springs are hydrogeological features by which groundwater discharges naturally to the land surface or cave from the sedimentary rock aquifers, typically along footslopes, faults or fractures. Upon reaching the surface, these underground features combine with surface features (e.g. topography, evaporation rate) and a range of springs result. There are different types of groundwater springs including small seeps, mound springs and vegetated springs (Figure 8). Small seeps are small springs with historically low flow, resulting in seeps with no distinct raised mound or vegetation.

Mound springs can form around vents (the spring's surface outlet) where subterranean pressure expresses water through cracks or faults. They are typically a raised mound with a central pool fringed by vegetation.

Mounds can form over time through a number of processes, as reported by Fensham and Fairfax (2003):

- Inorganic material is transported upwards and deposited at the spring vent
- Deposition of dissolved solids at the spring vent
- Swelling of surface clays due to the presence of water
- Accumulation of peat from the vegetation supported by the spring.

Some springs vent from a point, but do not form a mound. These springs provide suitable conditions for wetland vegetation to establish. Vegetated springs provide habitat, shelter, structure, shade, food and materials for a variety of animals; therefore, these springs tend to be more biodiverse than mound springs or unvegetated pools of open water (Wetland*Info* 2014).

Groundwater can also run over the land surface and evaporate at the margins of springs, or discharge in areas under lower pressure and evaporate on the surface. In arid environments, these processes result in salt accumulation and create a habitat where salt-loving plants are found (e.g. Eulo springs super-group) (State of Queensland 2011).

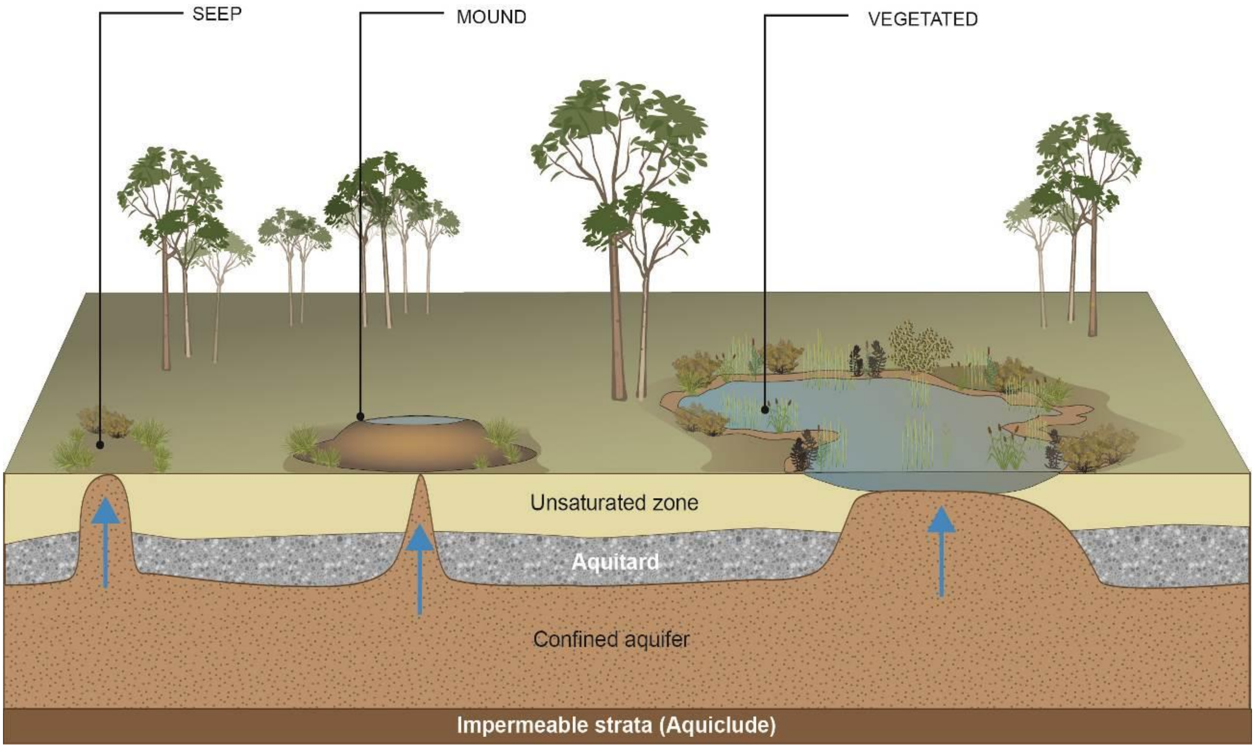


Figure 8: Conceptual model of spring types (seep, mound, vegetated)

Appendix 2 - Environmental Values and Level of Protection

The Environmental Values (EV) or groundwater uses that are recognised at a national level are provided in Figure 9.

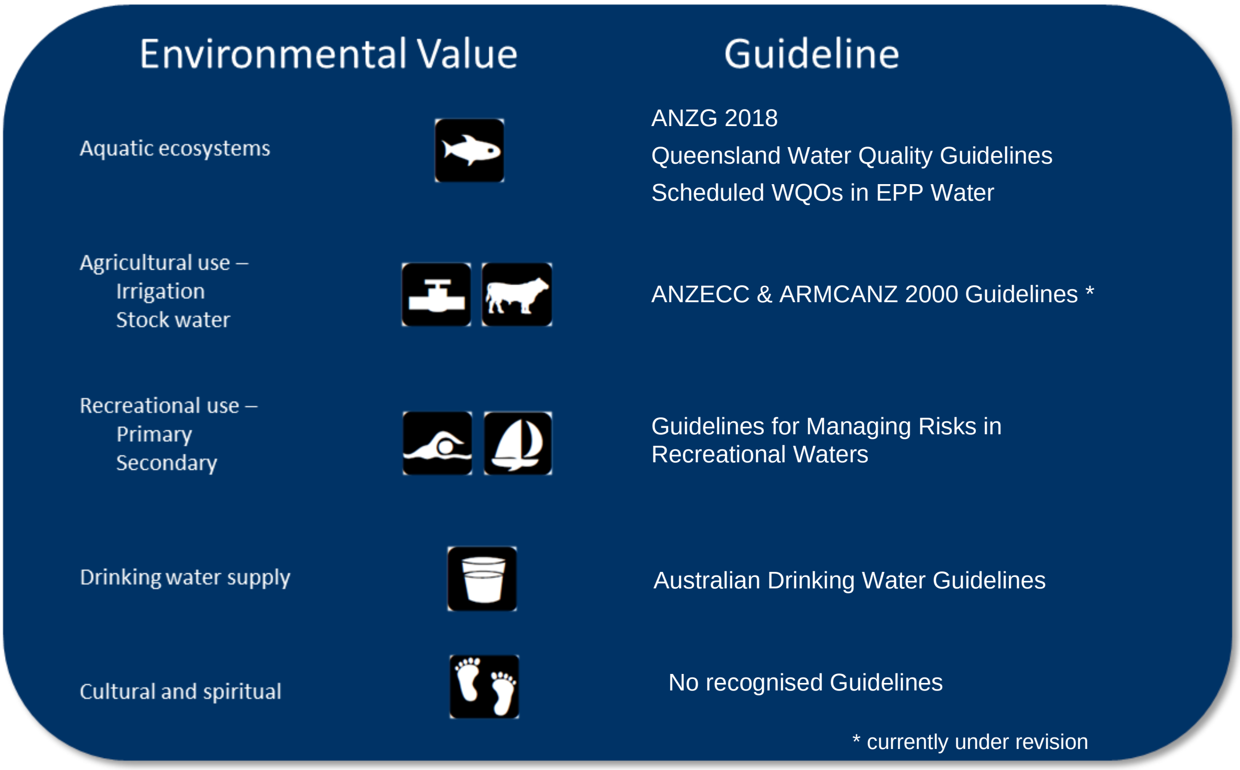


Figure 9: Environmental value and relevant guideline

Schedule B1 of the National Environment Protection (Assessment of Site Contamination) Measure made under the *National Environment Protection Council Act 1994* also provides investigation levels for soil and groundwater in the assessment of site contamination. The groundwater investigation level (GIL) is the concentration of a groundwater parameter at which further investigation is required (National Environment Protection Council 2011). The schedule includes Australian water quality guidelines/drinking water guidelines/guidelines for managing risk in recreational water criteria and site-specific derived criteria.

Regionally specific water quality guidelines are used in preference over national or state guidelines, with the Environmental Protection (Water and Wetland Biodiversity) Policy 2019 values taking precedence over the Queensland Water Quality Guideline (QWQG). In turn, the QWQG values take precedence over ANZG 2018 values. Where no QWQG values are available, users should default to ANZG 2018 or ANZECC & ARMCANZ 2000 Guidelines.

Under the ANZECC & ARMCANZ (2000) framework, if local reference/control values exceed the regional or national guidelines, then local guidelines may need to be derived. Groundwater quality guidelines may be locally defined to consider water quality issues specific to the groundwater system in question (See Section 4). The hierarchy for the application of groundwater quality guidelines is presented in Figure 10.

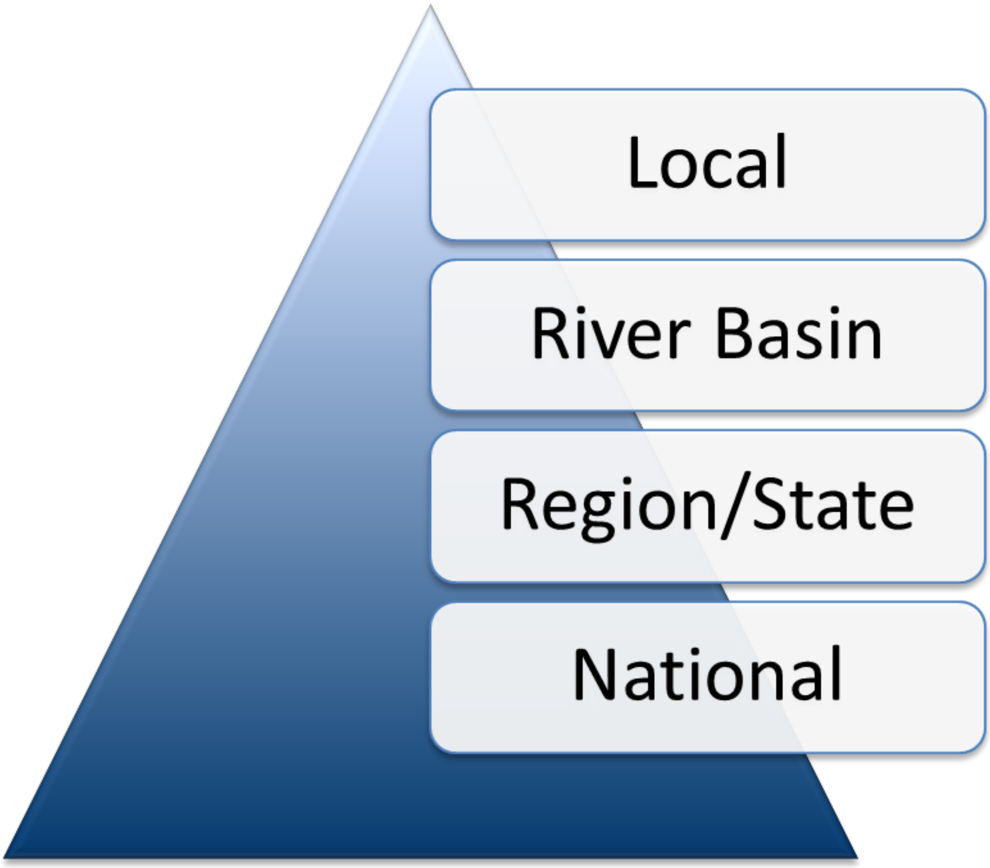


Figure 10: Hierarchy for the application of water quality guidelines.

Where groundwater discharges to a Queensland waterway, the water quality objectives for that waterway would apply. Guidelines for the protection of aquatic ecosystems are applicable for GDEs that rely on groundwater discharge to rivers or wetlands because the quality of the groundwater component can potentially alter the surface water quality at its discharge environment. There are currently no water quality guidelines for the protection of stygofauna (subterranean fauna) or GDEs that rely on the subsurface presence of groundwater (e.g. vegetation).

Guidelines should not be adopted as upper limits to which groundwater contaminant concentrations can be increased. Rather, if existing groundwater quality is below the water quality guidelines, then the groundwater should be managed within the range of natural water quality variability determined through baseline characterisation (Australian Government 2013).

The ANZG (2018) and QWQG (DEHP 2013) recognise three levels of protection for ecosystem condition, as defined below. For each level of protection, slightly different guidelines may apply.

1. High conservation/ecological value systems — effectively unmodified or other highly-valued ecosystems, typically occurring in national parks, conservation reserves or in remote and/or inaccessible locations.

2. Slightly to moderately disturbed systems — ecosystems in which aquatic biological diversity may have been adversely affected to a relatively small but measurable degree by human activity. The biological communities remain in a healthy condition and ecosystem integrity is largely retained. Typically, freshwater systems would have slightly to moderately cleared catchments and/or reasonably intact riparian vegetation; marine systems would have largely intact habitats and associated biological communities.

3. Highly disturbed systems. These are measurably degraded ecosystems of lower ecological value.

The *Environmental Protection (Water and Wetland Biodiversity) Policy 2019*, however, recognises four possible levels of ecosystem condition and corresponding management intent, namely high ecological value (effectively unmodified) systems; slightly disturbed, moderately disturbed, and highly disturbed systems (Table 8).

Table 8: Definition of aquatic ecosystem condition and management intent

EPP Water	QWQG	ANZECC & ARMCANZ	Definition	Management Intent
High Ecological Value (HEV)	HEV	HEV	Waters in which the biological integrity of the water is effectively unmodified or highly valued	Maintain existing water quality (20^{th}, 50^{th} and 80^{th} percentiles).
Slightly Disturbed (SD)			Waters that have the biological integrity of high ecological value waters with slightly modified physical or chemical indicators but effectively unmodified biological indicators	Achieve effectively unmodified water quality (20^{th}, 50^{th} and 80^{th} percentiles).
Moderately Disturbed (MD)	Slightly to Moderately Disturbed (SMD)	SMD	Waters in which the biological integrity of the water is adversely affected by human activity to a relatively small but measurable degree	Achieve Water Quality Objectives
Highly Disturbed (HD)	HD	HD	Waters that are significantly degraded by human activity and have lower ecological value than high ecological value waters or slightly or moderately disturbed waters	Progressively improve

Appendix 3 – Further Information on Data Analysis

Information below adapted from U.S. EPA 2009.

Non-detects simple substitution

Basic purpose: A simple adjustment for non-detects in a dataset, including results less than the limit of reporting (LoR). One-half the reporting limit (RL) is substituted for each non-detect to provide a numerical approximation to the unknown true concentration.

Underlying assumptions: The true non-detect concentration is assumed to lie somewhere between zero and the reporting limit. Furthermore, the probability of the true concentration being less than half the RL is about the same as the probability of it being greater than half the RL.

When to use: In general, simple substitution should be used when the dataset contains a relatively small proportion of non-detects, i.e. no more than 10-15%. Use with larger non-detect proportions can result in biased estimates, especially if most of the detected concentrations are recorded at low levels (*e.g.*, at or near RL).

Steps involved: 1) Determine the reporting limit; and 2) replace each non-detect with one-half RL as a numerical approximation.

Advantages/Disadvantages: Simple substitution of half the RL is the easiest adjustment available for non-detect data. However, it can lead to biased estimates of the mean and particularly the variance if employed when more than 10-15% of the data are non-detects.

Removing outliers

Basic purpose: Outliers are extreme values of a data set. An outlier should only be removed if there are strong reasons to believe it is the result of erroneous procedure in sampling, laboratory analysis or reporting. Outliers should not be removed *solely* on a statistical basis. The outlier tests can provide supportive information, but generally a reasonable rationale needs to be identified for removal of suspect outlier values (usually limited to background data). At the same time there must be some level of confidence that the data are representative of groundwater quality.

Underlying assumptions: Statistical tests assume that data without the suspected outliers are normally distributed.

When to use: Outliers can be screened using probability and box plots or using statistical outlier tests (Dixon's or Rosner's Test). Dixon's test is only recommended for sample sizes $n < 25$. Rosner's test is recommended when the sample size (n) is 20 or larger.

Steps involved: 1) Plot the data (box plot or times series plot); 2) Determine four standard deviations from the mean; 3) Identify any extreme values in the data set, this can be measurements that are greater than four standard deviations from the mean; and 4) Investigate these data points further.

It is suggested that only the most extreme observations (e.g. those that are four or more standard deviations from the mean) be excluded unless other good reasons can be established. When assessing compliance against environmental approval conditions, outliers should **not** be removed as they may indicate potential impacts. However, when using a dataset to calculate percentiles to determine site-specific guidelines, outliers should be removed.

Advantages/Disadvantages: Discarding observations is something that should not be undertaken lightly and needs to be fully justified.

Box and whisper plots

Basic purpose: Diagnostic and exploratory tool that provides a graphical summary of data distribution, giving a picture of central tendency and dispersion of data. Although not a formal statistical test, a side-by-side box plot of multiple datasets can be used as a rough indicator of either unequal variances or spatial variation (via unequal medians).

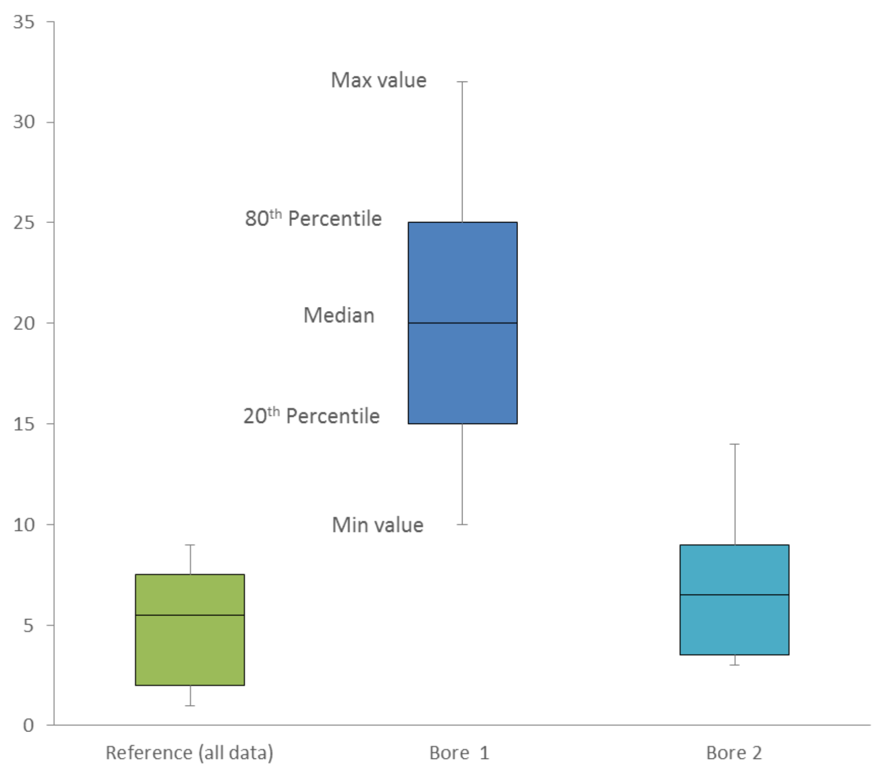
Underlying assumptions: When used to screen outliers, underlying population should be approximately symmetric.

When to use: Can be used as a quick screen in testing for unequal variances across multiple populations. Substantially different box lengths suggest possibly different population variances. It is useful as a rough indication of spatial variability across multiple bore locations.

Box plots can also be used to screen for outliers: values falling beyond the 'quartile whiskers' on the box plot are labelled as potential outliers.

Steps involved: 1) Compute the median, lower and upper quartiles (*i.e.* 25th and 75th percentiles) **or** other percentile (i.e. 20th and 80th) of each dataset; 2) Graph each set of summary statistics side-by-side on the same set of axes; 3) Compute the 'whiskers' by extending lines below and above the box by an amount equal to 1.5 times the interquartile range **or** the maximum and minimum data points.

Advantages/Disadvantages: The box plot is an excellent visual aid in diagnosing either unequal variances, the possible presence of spatial variability, or potential outliers. However, it is not a formal statistical test, and should generally be used in conjunction a statistical test.



Time series plots

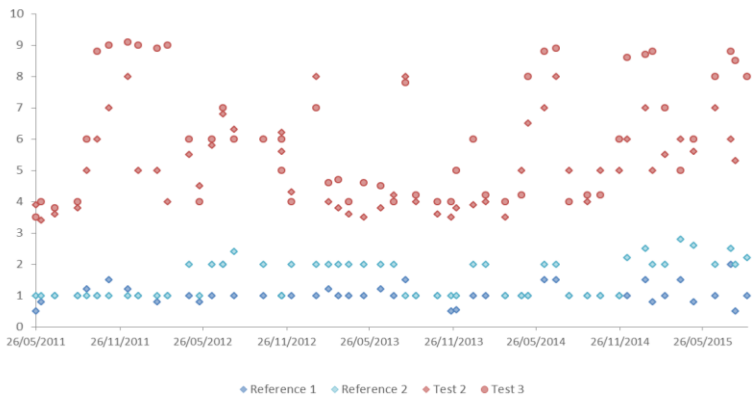
Basic purpose: Diagnostic and exploratory tool. It is a graphical technique to display changes in concentrations at one or more bores over a specified period of time or series of sampling events. Time series plots can be used to informally gauge the presence of temporal and/or spatial variability in a collection of distinct bores sampled during the same time frame.

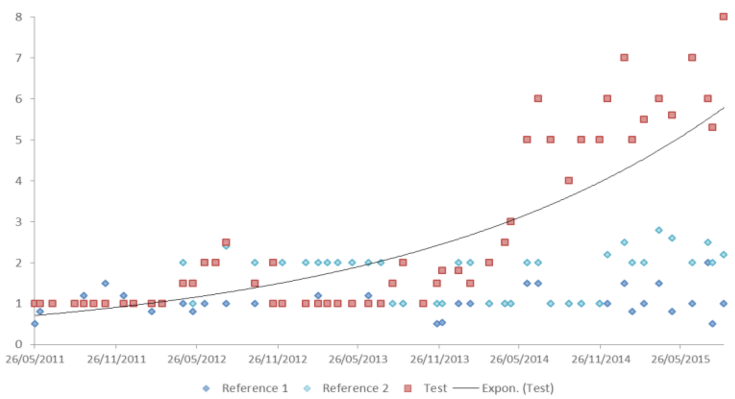
Underlying assumptions: None.

When to use: A time series plot can provide information on whether there exists time-related or temporal dependence in the data. When several bores are plotted together, such temporal dependence can be seen in parallel movement on the time series plot. That is, several bores will either exhibit the same pattern of up-and-down fluctuations over time, or display different temporal variation.

Steps involved: 1) Plot concentration against time or date of sampling for the sampling events that occurred during the specified time period at each bore on the same graph; 2) Make sure each bore is identified on the plot with a distinct symbol and/or connected line pattern; 3) To observe possible spatial variation, look for bores that are substantially separated from one another in concentration level; 4) To look for temporal dependence, look for bores that rise and fall together in roughly the same (parallel) pattern; 5) To ensure that artificial trends due to changing reporting limits are not reported, plot any non-detects with a distinct symbol, colour, and/or fill.

Advantages/Disadvantages: Time series plots are an excellent tool for examining the behaviour of one or more bores over time. Although, they do not offer the compact summary of distributional characteristics that box plots do, time series plots display each and every data point and provide an excellent initial indication of temporal dependence. Since temporal dependence affects the underlying variability in the data, its identification is important so adjustments can be made to the estimated standard deviation.





Piper diagrams

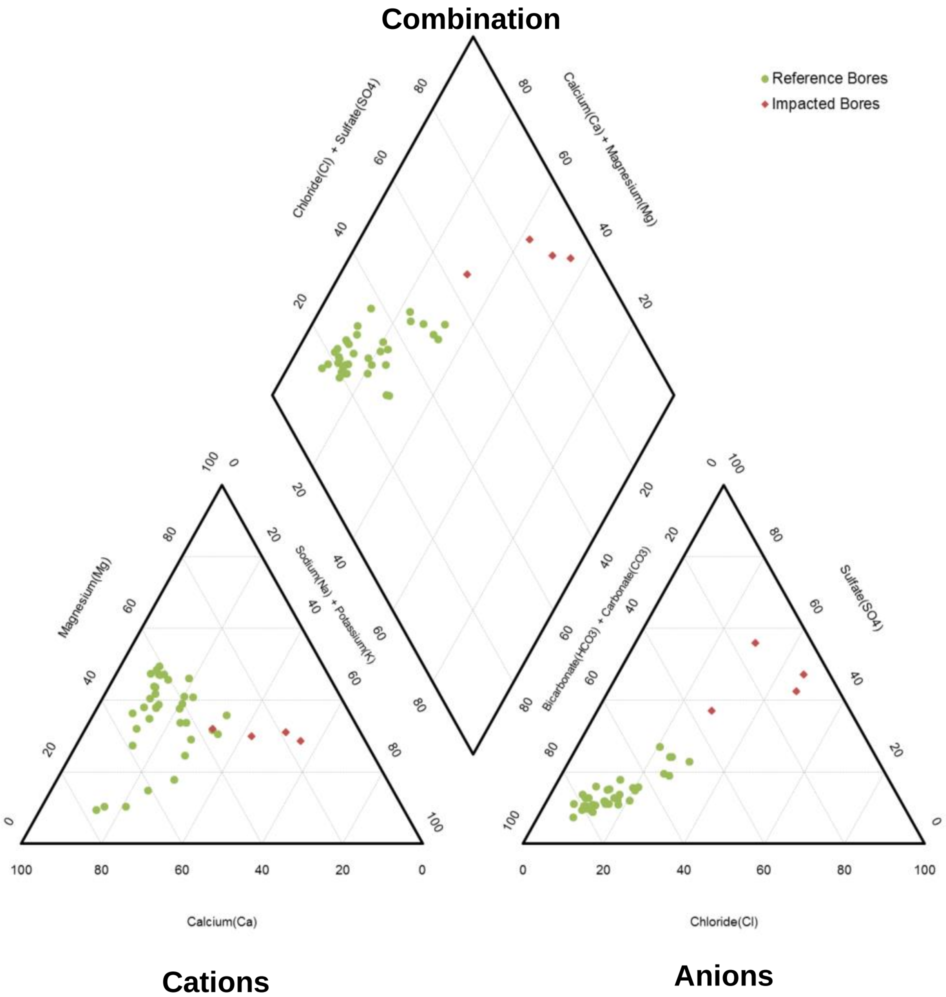
Basic purpose: Diagnostic and exploratory tool. A piper plot is a way of visualising the chemistry of a groundwater sample. Provides a convenient method to classify and compare water types based on the ionic composition of different groundwater samples.

Underlying assumptions: The similarity in water type between ground-water samples suggests that similar geochemical processes may be controlling major-ion chemistry in these aquifers and that the waters had the same or similar origins.

When to use: Piper diagrams are trilinear representations of the ions found in the water (cations, anions & combined properties) and are used in order to classify water types. The major cations include sodium, potassium, calcium and magnesium and the major anions include chloride, sulphate, bicarbonate and carbonate.

Steps involved: 1) determine the normalised concentrations of each of the three cations for a sample and plot them on the lower left ternary diagram; 2) determine the normalised concentrations of the three anions for a sample and plot them on the lower right ternary diagram; 3) following a line parallel to the outer axis of each ternary diagram, project each point in the ternary diagrams upward until they intersect with one another in the diamond plot.

Advantages/Disadvantages: Many water analyses can be plotted on the same diagram. The piper diagrams can be used to classify waters, with the top quadrant calcium sulfate waters (gypsum ground water and mine drainage), the left quadrant calcium bicarbonate waters (shallow fresh ground water), the right quadrant sodium chloride waters (marine and deep ancient ground water), and the bottom quadrant sodium bicarbonate waters (deep ground water influenced by ion exchange).



Shewhart-CUSUM Control Chart

Basic purpose: Method for detection monitoring. These are used to quantitatively and visually track concentrations at a given bore over time to determine whether they exceed a critical threshold (*i.e.*, control limit), thus implying a significant increase above background conditions.

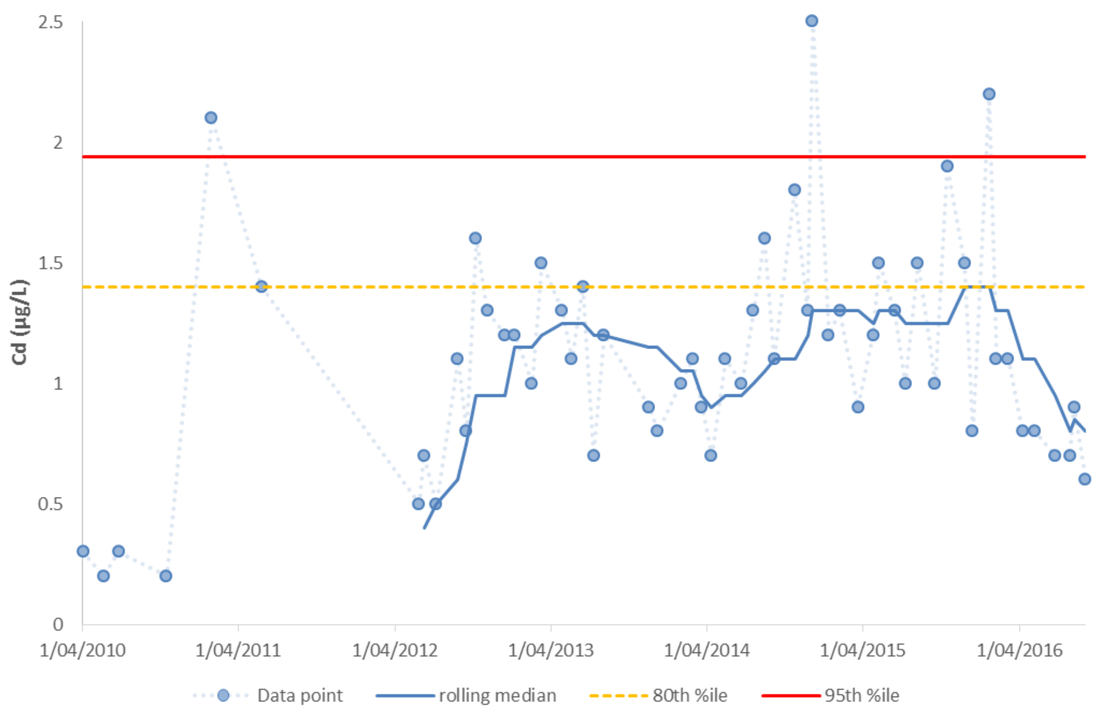
Hypothesis tested: H_0 — Data plotted on the control chart follow the same distribution as the background data used to compute the baseline chart parameters.

Underlying assumptions: Data used to construct the control chart must be approximately normal or normalized. There should be no discernible trend in the baseline data used to calculate the control limit.

When to use: Use control charts as an alternative to parametric prediction limits, when 1) there are enough uncontaminated baseline data to compute an accurate control limit, and 2) there are no trends in intra-bore background. Retesting can be incorporated into control charts by judicious choice of control limit. This may need to be estimated using Monte Carlo simulations.

Steps involved: 1) Compute the intra-bore baseline median/mean, standard deviation and percentiles; 2) calculate an appropriate control limit from these baseline parameters, the desired retesting strategy and number of bore-constituent pairs in the network; 3) construct the chart, plotting the control limit, the compliance point observations, and the cumulative sums [CUSUM] (rolling median); and 4) determine that the null hypothesis is violated when either an individual concentration measurement or the cumulative sum (rolling median) exceeds the control limit.

Advantages/Disadvantages: Unlike prediction limits, control charts offer an explicit visual tracking of compliance point values over time and provide a method to judge whether these concentrations have exceeded a critical threshold. The Shewhart portion of the chart is especially good at detecting sudden concentration increases, while the CUSUM portion is preferred for detecting slower, steady increases over time.



Appendix 4 – Superseded compliance approaches

Control Charting

Background and Suitability

Control charting is a graphical tool that can be used to track changes in measured data over time. Relative changes over time can be compared in order to determine if temporal changes are within the expected range of natural variability. A defining element of a control chart is the inclusion of 'control criteria lines' or 'control limits' that can be used to inform or trigger management actions. Control limits have generally been defined in the past as descriptive statistics such as standard deviations from the mean.

Control charting is derived from statistical process control within the manufacturing industry, which is used to measure changes of means and variance of the values of an indicator relative to notional 'action thresholds'. Known as Shewhart control charts, they were developed in the 1920s with a stated objective to separate out random variation and to allow the "*diagnosis and correction of production troubles and to bring improvement to product quality*" (Grant 1964). Control charts are plots of raw data or summary statistic from samples taken sequentially in time. Control charts typically utilise the mean as the basis of trend analysis, and upper and lower control limits. Control limits are set to determine whether or not a particular average is 'within acceptable limits' of random variation.

Control charts have been used in a number of instances to determine variability in water and sediment quality and identify trends (Lee *et al*. 2015, Maurer *et al*. 1999, Paroissin *et al*. 2016) and to evaluate water quality within water treatment plants and distribution networks (Shaban 2014; Cornwell *et al*. 2015). When applied to assess water treatment and distribution, control charts were used to allow the plant operator to properly evaluate whether the measured concentrations remain within an acceptable range, and if not, then management actions would be required to bring the indicator back within the control limits (Cornwell *et al*. 2015). Control charts should be seen as best practice for managing water quality on site not necessarily for use in compliance.

An advantage of control charting is that it allows point data to be viewed and assessed graphically over time. Trends and changes in concentrations can be easily seen, as the measures are consecutively plotted on the chart. Control charts can be useful for comparing data values with guidelines or trigger values and it is seen as an appropriate management tool to monitor changes in groundwater quality overtime (ANZECC & ARMCANZ 2000). In contrast, commonly used compliance and assessment approaches compare *point-in-time comparisons* with trigger levels and limits; however, these approaches don't permit consideration of long-term trends.

For the purpose of applying control charts to groundwater quality assessment, control limits can be developed from either inter-bore or intra-bore data. Inter-bore control limits are derived from hydraulically upgradient reference bore data, and potentially, data from other background bores. Intra-bore control limits are derived from historical measurements collected from the same compliance bore. These can also be considered as pre-disturbance baseline data. The type of comparison that is most appropriate often depends on the local hydrogeological conditions (ITRC 2013).

The U.S. EPA (2009) suggests that control charts may be used when there is sufficient uncontaminated baseline data to calculate reliable control limits, and therefore, no confounding trends in the intra-bore background data. The specific control chart recommended by the U.S. EPA (2009) for groundwater quality assessment is the combined Shewhart - cumulative sum (CUSUM) control chart. The Shewhart control chart is designed to detect an immediate increase over the background distribution for a particular contaminant, whereas the CUSUM control chart is designed to detect more gradual or cumulative increases over background (Gibbons 1999).

Where the control limits are derived from historical data of bores likely to have been impacted, then the control limits can only be viewed as ‘lines in the sand’. This is a reasonable approach if the management intent is to have “no worsening” or “gradual improvement over time” with regard to toxicants and physico-chemical indicators. When considering the use of control charts to support environmental approvals, the following should be considered:

- Providing justification for using a control charting approach
- Deriving triggers for control charts
- Evaluation of the proposed approach.

Approach

There are a number of approaches that could be adopted when deriving triggers for control charting, often called control limits, for a given site or location. The following is provided as a guide to determining and applying two methods. Generally, the second method is more likely to be applicable to groundwater quality monitoring given it does not require an assumption of normally distributed data.

The data analysis described above (Sections 3 and 4) should assist to firstly determine whether bores need to be considered separately or whether bores can be grouped and data pooled. Once this is determined, the characteristics of the available data should be considered, such as whether the data is normally distributed, whether there are outliers in the data set that should be removed and how non-detects will be dealt with.

A further important consideration is whether the entire dataset or a subset of the available data, such as pre-development, should be used to the calculate control limits. In addition, what algorithm should be used to calculate the control limits and how will the limits be applied to test data.

If a control charting approach is to be adopted, a control chart (see Figure 13 for example) for each bore or group of bores and each indicator with the proposed control limits or Tier triggers needs to be reviewed to determine if the proposed triggers and limits are appropriate.

The first approach commonly used relies on the data being normally distributed, which is rarely the case for environmental data. Here, control limits are generally based on testing individual observations against the mean (μ) and standard deviation (σ) from the mean. The mean and standard deviation for each indicator is determined from initial baseline monitoring data. Once the activity commences, monitoring data for each indicator are plotted against time, with control limits of mean + σ, mean + 2σ and mean + 3σ. The mean + σ is equivalent to the 68th percentile, mean + 2σ the 95th percentile and mean + 3σ the 99th percentile (Figure 11).

Control limits are set which constitute an alert when:

- one observation is above the mean + 3σ, or
- two consecutive observations are above the mean + 2 σ, or
- five successive observations are above the mean + σ.

If there is an extended period of no observations greater than the control limits (e.g. after 12 observations), then the mean and standard deviations could be recalculated to provide higher confidence in the control triggers and limits.

Control limits using standard deviation from the mean have been used in the past; however, the use of mean + 3 σ is approximately equal to the 99th percentile, which is at the extreme of the distribution. The degree of confidence around the calculation of the 99th is generally low and it may not be sufficiently conservative for a compliance approach.

In a recent study, Cornwell *et al*. (2015) suggested that upper and lower control limits based on standard deviation from the mean in control charts should not be used for a regulatory action. As an alternative, the use of a median and percentile statistics to define control limits could be applied as these do not require any assumptions about the data distributions, and can be used in the same manner. The benefit of using the median versus the mean is that the median is not affected by a highly skewed distribution or the inclusion of an outlier in the dataset (U.S. EPA 2009).

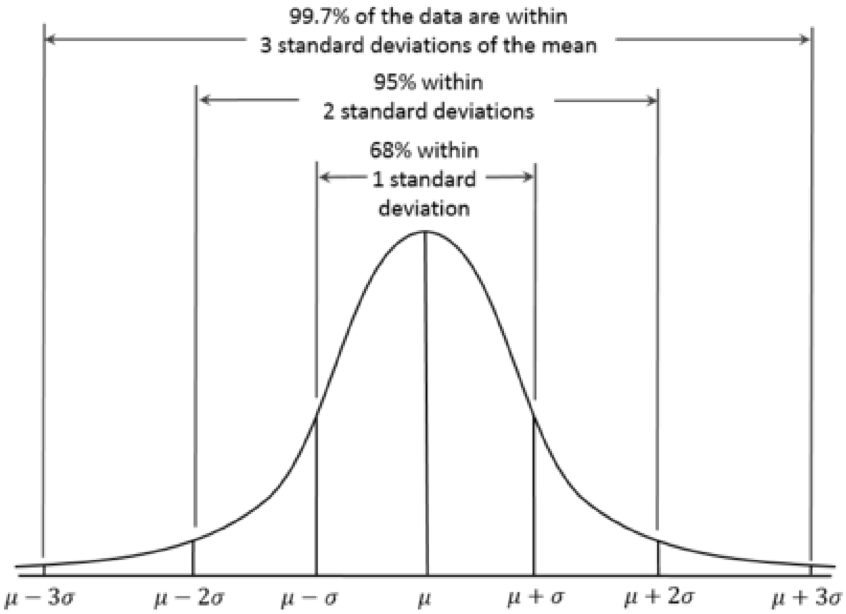


Figure 11: Normal distribution curve showing mean (μ), standard deviation (σ) and equivalent percentiles.

Example - Control chart triggers with percentiles

The percentile control chart example uses the 85th, 97th, 99th percentiles of the entire data set (*n* = 36) for two bores (Figure 12). These percentiles were calculated as equivalent to the standard deviations based on historic groundwater quality data for the bores. There is a difference between the limits derived using all available data at the two bores from the same area based on the variability within the data at each bore. The limits derived for Bore 2 are influenced by an increasing trend in the data and a potential outlier in 2009. This example highlights the need to determine trends, peaks and outliers in the data before control limits are determined.

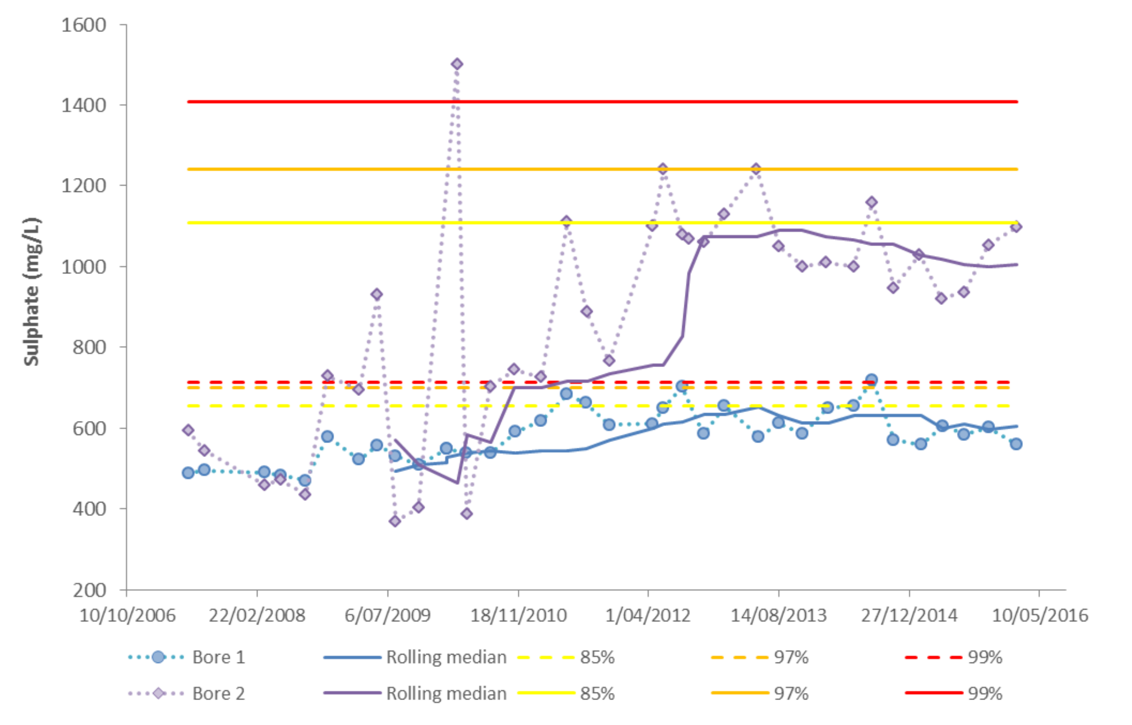


Figure 12: 85th, 97th and 99th percentile control triggers calculated using all available sulphate data at monitoring bore 1 and 2. The median is calculated as a rolling value from the most recent eight sample values.

Two-tiered nonparametric triggers

To provide a pragmatic approach to assessing compliance at a highly disturbed site, a two-tiered nonparametric trigger system approach was developed by Mann and Dunlop (2015) for the Mount Morgan Mine site – see Table 9 and Figure 10. This approach was developed as a practical method for monitoring the rate and magnitude of change, which improves on the approach discussed above. Two triggers were proposed as a compromise between the more traditional approach that employs a comparison of the median (50th percentile) with a trigger based on the 80th percentile of a historical reference data-set, and the control charting approach described below. **Trigger 1**, which uses the more traditional approach, is a statistically robust method for detecting gradual change over the medium term (i.e. 1 to 2 years). **Trigger 2** permits the detection of event related change over a relatively short term (i.e. several months depending on the sampling frequency).

Table 9: Proposed two tier triggers and assessment methodology.

	Trigger	Assessment approach for trigger	Comment
Tier 1 Trigger	20^{th} (pH) / 80^{th} percentile of a valid data-set.	Rolling median (50^{th} percentile) of between 8 and 12 consecutive samples. We recommend a median based on 12 months of data representing wet and dry seasons. Where triggers are split between wet and dry season data, there may be insufficient data to calculate a suitably reliable median, and it will be more appropriate to use data collected over a longer period of time.	Because 20% of all measurements will be higher than the 80^{th} percentile (or lower than the 20^{th} percentile), it is not appropriate to assess individual measurements against this trigger because of the probability of assigning a false positive. However, comparing a median calculated using a suitably large number of consecutive samples with the trigger (based on 20^{th} and 80^{th} percentiles) will reduce the probability of a false positive to a very low risk.
Tier 2 Trigger	5^{th} (pH) / 95^{th} percentile of a valid data-set.	Two or three consecutive individual exceedances.	Because 5% of all measurements will be higher than the 95^{th} percentile (or lower than the 5^{th} percentile), there remains a low probability of assigning a false positive. However, the probability of two or three consecutive false positives is very low (0.25% & 0.0125% respectively) unless the exceedances are real.

The two tiered trigger nonparametric approach proposed compares a rolling median against triggers that are calculated based on the long-term 80^{th} (Tier 1 trigger) and three consecutive individual exceedances of the long-term 95^{th} percentile (Tier 2 trigger). The triggers are static and not calculated on a rolling basis. Because there will be instances, where there are inadequate valid data points to calculate reliable percentile statistics, the ANZG values for protection of 80%, 90% or 95% species (Table 3.4.1) may be substituted as surrogate Tier 2 triggers in place of the 95^{th} percentile. The ANZG values would effectively only be assigned for those toxicants that never or only rarely exceed the limit of reporting.

Example – Two tiered trigger using nonparametric triggers

Two examples using the two tiered trigger are provided below. Note that the triggers proposed in these examples are more conservative than the 85th, 97th, 99th percentile in Figure 12.

An example application of the proposed two tiered control charting approach is provided in Figure 13 below. In this example the blue line indicates the rolling median (*n*=8) and the orange dashed and red lines represent the long-term 80th and 95th percentiles that have been used to calculate Tier 1 and 2 triggers, respectively. This example illustrates how control charting allows for the monitoring of water quality indicators over time. Where a control charting approach identifies long-term temporal changes in water quality, it is essential that there is ongoing review of the percentile-based control criteria (trigger levels) used to assess water quality.

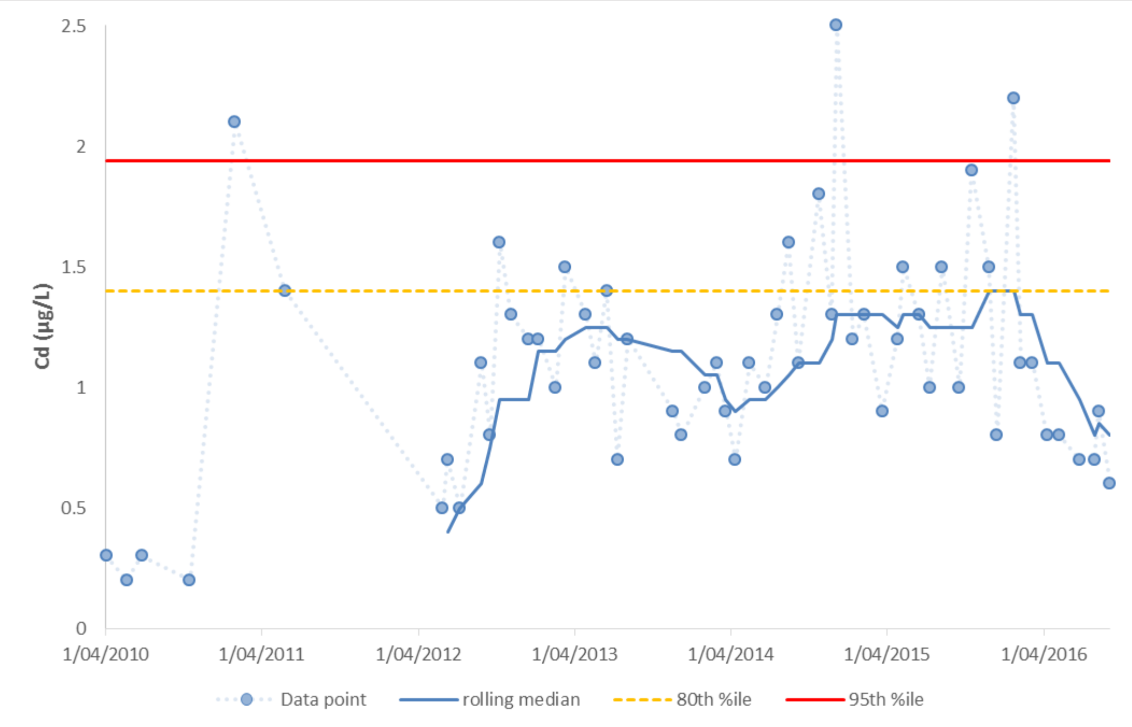


Figure 13: Example of using the 80th and 95th percentile triggers to assess total cadmium using a control chart at a monitoring bore where the rolling median is calculated from the most recent eight consecutive sample values.

The data in Figure 13 would at no point trigger a management response because the rolling median does not exceed the 80th percentile of the long-term data-set (Tier 1 trigger), and although there are individual exceedances of the 95th percentile (Tier 2 trigger) on three separate occasions, consecutive exceedances do not occur.

In the below example the 80th and 95th percentiles are based on pre-mining data (2008 – July 2011) (dashed lines) and the entire dataset (solid lines) (Figure 14). This example illustrates the dramatic differences that can be expected when triggers are derived using pre-mining data compared with triggers derived from data represented from post-mining samples at a contaminated site. In the example in Figure 14 there is an increasing trend in the indicator, decisions need to be made about which data are used to calculate triggers and whether the increasing trend in the indicator should be investigated.

The data used to define the triggers and limits can greatly influences when an exceedance occurs. This example highlights the need for the proponent to provide control charts with the proposed triggers and limits before a control charting approach can be adopted.

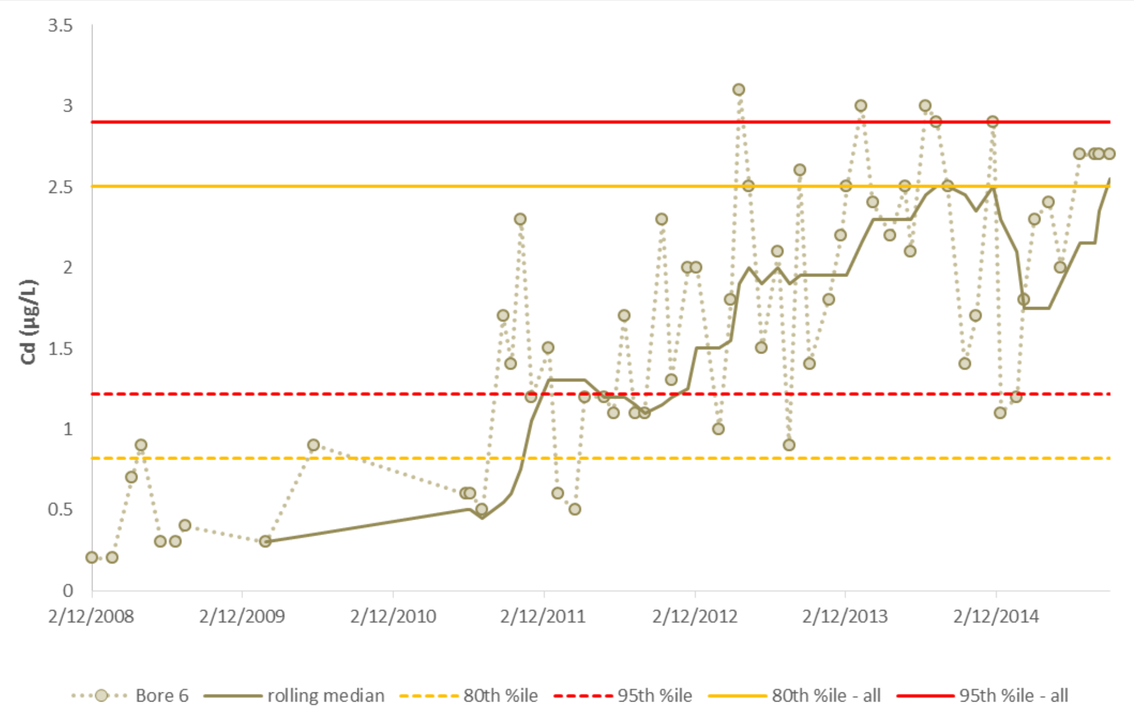


Figure 14: 80th and 95th percentile triggers calculated using pre-mining (dashed line) and the 80th and 95th percentile triggers calculated using all available total cadmium data (solid line) at a monitoring bore. The median is calculated as a rolling value from the most recent eight sample values.

In these examples, non-compliance is when:

- Rolling median of eight samples is greater than Tier 1 trigger (80th percentile of pre-mining all available data)
- Three (3) consecutive individual exceedances greater than Tier 2 trigger (95th percentile of a valid data-set).

An evaluation of the proposed approach should be done. In this case an evaluation should assess whether the rolling median of the most recent eight samples at each bore exceed the proposed Tier 1 trigger and a pass or fail assigned for each indicator. For the Tier 2 trigger an evaluation of whether there were three consecutive individual exceedances of the proposed trigger at each bore should be done and a pass or fail assigned for each indicator.