



An Australian Government Initiative



Australian & New Zealand

**GUIDELINES FOR
FRESH & MARINE
WATER QUALITY**

Guidance on the use of ecosystem receptor indicators for the assessment of water and sediment quality

July 2023

© Commonwealth of Australia 2023

Ownership of intellectual property rights

Unless otherwise noted, copyright (and any other intellectual property rights, if any) in this publication is owned by the Commonwealth of Australia (referred to as the Commonwealth).

Creative Commons licence

All material in this publication is licensed under a Creative Commons Attribution 4.0 Australia Licence, except for photographic images, logos and the Commonwealth Coat of Arms.



Creative Commons Attribution 4.0 Australia Licence is a standard form licence agreement that allows you to copy, distribute, transmit and adapt this publication provided you attribute the work. See the [summary of the licence terms](#) or the [full licence terms](#).

Inquiries about the licence and any use of this document should be emailed to copyright@dcceew.gov.au.

Cataloguing data

This publication (and any material sourced from it) should be attributed as: ANZG (2023) *Guidance on the use of ecosystem receptor indicators for the assessment of water and sediment quality*, Australian and New Zealand guidelines for fresh and marine water quality. CC BY 4.0. Australian and New Zealand Governments and Australian state and territory governments, Canberra, ACT, Australia.

This publication is available at <https://www.waterquality.gov.au/anz-guidelines/resources/key-concepts/indicators/ecosystem-receptor-indicators>.

Contact

Australian Government Department of Climate Change, Energy, the Environment and Water
GPO Box 858 Canberra ACT 2601
General enquiries: 1800 920 528
Email waterquality@dcceew.gov.au

Liability

The authors of this publication, all other entities associated with funding this publication or preparing and compiling this publication, and the publisher of this publication, and their employees and advisers, disclaim all liability, including liability for negligence and for any loss, damage, injury, expense or cost incurred by any person as a result of accessing, using or relying on any of the information or data in this publication to the maximum extent permitted by law.

Acknowledgements

This guidance document was prepared by Dr Ross Smith (Hydrobiology), Dr Leon Barmuta (University of Tasmania) and Dr Graeme Batley (CSIRO), and reviewed by Dr Chris Humphrey (Office of the Supervising Scientist, Department of Climate Change, Energy, the Environment, and Water) and contracted technical advisor Dr Rick van Dam (WQadvice). The document was also reviewed and approved by jurisdictional technical and policy oversight groups prior to being published.

New Zealand Government



Contents

Summary	iv
1 Introduction.....	1
2 Selecting indicators	3
3 Ecosystem receptor lines of evidence and their indicators	9
3.1 Biodiversity	9
3.1.1 Population-level indicators	9
3.1.2 Community-level indicators	10
3.1.3 Ecosystem function indicators	15
3.1.4 Additional guidance	17
3.2 Toxicity	17
3.2.1 Single-chemical toxicity testing	17
3.2.2 Direct toxicity assessment	19
3.2.3 Toxicity testing in sediments	26
3.2.4 Semi-field-based assessments of toxicity	26
3.3 Bioaccumulation	27
3.4 Biomarkers	29
4 Summary	31
Appendix A: Additional guidance for matching biodiversity indicators to problems	32
Glossary and acronyms.....	37
References	39

Figures

Figure 1 Weight-of-evidence process for assessing water and sediment quality.....	1
---	---

Tables

Table 1 Broad indicator types, suitable scenarios and examples for assessing biodiversity, conservation status and ecosystem-level responses	6
--	---

Appendix tables

Table A1 Matching biodiversity indicators to a problem	32
--	----

Summary

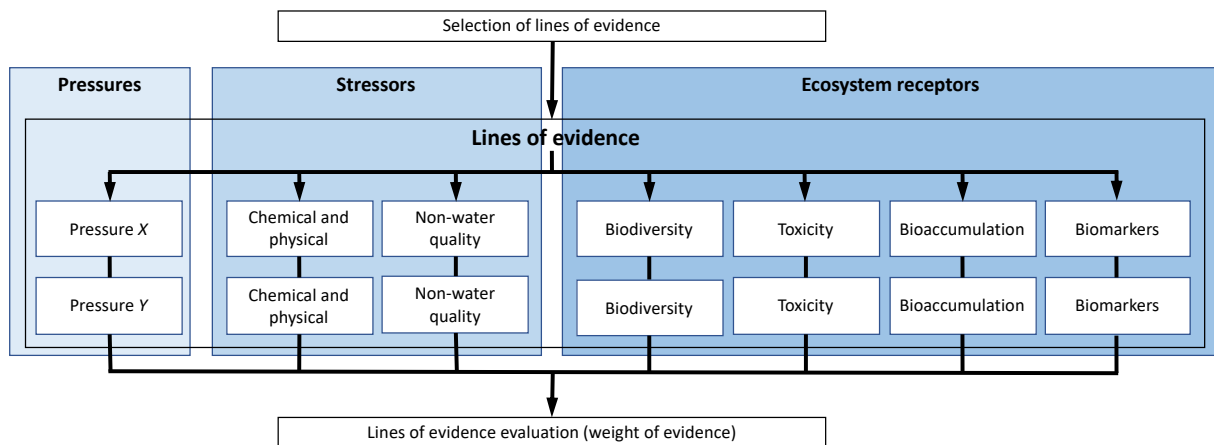
This guidance document provides updated information on the measurement of ecosystem receptor indicators required for the ecosystem receptor lines of evidence for the assessment of water and sediment quality in Australia and New Zealand. The guidance provided here is an extension of the ANZG (2018) guidance on indicators and their selection for the assessment of water and sediment quality, and is applicable to all water types. Hyperlinks to the relevant ANZG (2018) guidance are provided throughout this document.

This document describes 4 lines of evidence for ecosystem receptors: biodiversity, toxicity, bioaccumulation and biomarkers.

1 Introduction

ANZG (2018) provides guidance on how to assess water and sediment quality using the water quality management framework (WQMF) and an associated weight-of-evidence process, which includes the selection and assessment of indicators for pressures, stressors and ecosystem receptors. This guidance document is an extension of the ANZG (2018) guidance for the selection and assessment of ecosystem receptor indicators. The general guidance presented here is applicable to all water types, even if guidance or examples for specific water types are not presented. Additional specific guidance for certain water types (e.g. groundwaters, hypersaline waters) can be provided in the future.

Any living organism or natural living habitat that could be exposed to a stressor can be an ecosystem receptor. Ecosystem receptor indicators are measurable characteristics or components of specific ecosystem receptors that can be used to measure the response of an ecosystem receptor to a stressor. They are also referred to as ‘biological indicators’. Ecosystem receptor indicators, along with indicators for pressures and stressors, are recommended to be used as lines of evidence in a weight-of-evidence process to assess water and sediment quality (ANZG 2018). The shaded boxes in Figure 1 represent the 3 overarching causal pathway elements (pressures, stressors, ecosystem receptors) that ANZG (2018) recommends be used for selecting indicators for weight-of-evidence assessments.



Source: Adapted from ANZG (2018).

Figure 1 Weight-of-evidence process for assessing water and sediment quality

Each causal pathway element has one or more different lines of evidence that may be relevant to an investigation. For example, the ‘ecosystem receptors’ causal pathway element includes biodiversity, toxicity, bioaccumulation and biomarker lines of evidence, as depicted in Figure 1. Each line of evidence has indicator types. For example, the ‘biodiversity’ line of evidence could consider biotic assemblages, individual species (e.g. keystone species, threatened species) and ecosystem processes as indicator types. Each indicator type has a range of indicators that can potentially be used. For example, the ‘biotic assemblages’ indicator type could consider benthic macroinvertebrate communities. For each indicator, one or more parameters can be measured. For example, parameters for benthic macroinvertebrate communities could include community structure, total species abundance and species richness. More examples of the weight-of-evidence guidance in ANZG (2018) are provided in Table 1.

For simplicity, all the indicators that fall within the ecosystem receptors causal pathway element are collectively referred to as 'ecosystem receptor indicators'. Similarly, the ecosystem receptor indicators that fall within a line of evidence can be collectively considered as indicators of that line of evidence (e.g. biodiversity indicators for the biodiversity line of evidence).

2 Selecting indicators

It is important that selected ecosystem receptor indicators are relevant to the issue and its temporal and spatial scales. The suitability of indicators should become apparent as the current understanding of an issue (including the pressure, stressor and ecosystem receptor causal pathways) is articulated and an associated conceptual model is developed. For example, depending on the understanding of the water or sediment quality issue gained at [Step 1](#) and the associated management goals identified at [Step 2](#) of the ANZG (2018) WQMF, the roles of ecosystem receptor indicators could range from short-term, early-warning indicators to long-term indicators of the performance of management actions. Their spatial scope can range from site-specific issues to regional-scale assessments, depending on the scale of the issue being considered (see [7 typical uses of the WQMF](#)). Candidate indicators should be identified and selected in [Step 3](#) of the WQMF, taking into account the appropriate sampling designs and data analysis methods and constraints. Critically, ecosystem receptor indicators should be selected and interpreted as part of a weight-of-evidence process, as described in Section 1.

Rohr et al. (2016) provide an overview of the merits of different levels of organisation of ecosystem receptor indicators (sub-individual, individual, population, community, ecosystem) for use in ecological risk assessment in the USA. They conclude that there was no universal ideal level of biological organisation from which to select indicators. They also foreshadowed the imminent development of modelling and mathematical tools that would integrate 'bottom-up' indicators (e.g. biomarkers, ecotoxicity) and 'top-down' indicators (e.g. ecosystem processes). Additionally, frameworks such as the 'adverse outcome pathways' framework (Vinken 2013) can assist in framing conceptual models, identifying candidate indicators, and guiding modelling approaches to extrapolate observations of responses at lower levels of biological organisation to broader issues, such as ecosystem responses and services (Ankley et al. 2010, Madden et al. 2014, Forbes et al. 2017).

Generally, ecosystem receptor indicators are a fundamental consideration in all weight-of-evidence assessments of water and sediment quality because they provide diagnostic information and inference about the presence of an impact and its extent. Some 'keystone' biota can act as modifier indicators because they can form habitats, alter habitats, or facilitate or inhibit the responses of other indicators (e.g. populations of invasive plant species, irruptions of dominant urchins, loss of seagrass in marine systems). The other important role of ecosystem receptor indicators is as performance indicators. Specific biodiversity aspects of performance indicators are identified in [Step 2](#) of the WQMF as a management goal in themselves (e.g. protection of an endangered species or community).

ANZG (2018) distinguishes the roles of toxicity and biomarker indicators from those of biodiversity indicators. Toxicity and biomarker indicators also differ from each other. However, they typically provide either early-detection information or diagnostic information in a weight-of-evidence assessment (e.g. stressor presence, magnitude and type of effect). Often, for these (non-biodiversity) indicators, the measurement program objective is to detect any change (including trend) from a control or reference condition associated with stressors. Where toxicity or biomarker assessments are undertaken in the laboratory, they are less likely to be confounded by non-water-quality

stressors and other environmental influences than biodiversity assessments. For field-based studies, particularly those assessing biodiversity indicators for ecosystem-level response or biomarker indicators, non-water-quality stressors and other environmental influences contribute to inherent variability, and the indicators may respond to one or several stressors in a similar way. Further, because of the inherent variability, associated sampling designs have a limited capacity to detect and quantify change relative to an undisturbed or reference state.

The selection of ecosystem receptor indicators is highly dependent on the nature of the issue(s), including its specific pressures and stressors, the questions being asked, and the associated type of assessment required. Candidate indicators must be relevant and sensitive (relative to other potential indicators) to the pressures and stressors of concern, suitable for the type of assessment required, measurable in practice at the appropriate scale (e.g. catchment-level, site-specific), and relevant to the management goals. Candidate indicators should be considered and prioritised in a quality-of-evidence table as part of the weight-of-evidence process, based on the current understanding of (Step 1) and management goals for (Step 2) the issue. A quality-of-evidence table will help identify an appropriate suite of indicators that will result in an acceptable quality of evidence for the assessment. Guidance on selecting indicator types for pressures and stressors is in Table 1. Once indicators are selected, specific details about spatial scale, frequency, required limits of detection, and other details of sampling and measurement methods need to be considered. These are discussed in ANZG (2018).

The type of assessment required varies depending on the issue and on the questions being asked. Certain types of ecosystem receptor indicators will suit specific types of assessments. As a guide, some examples of indicators that are potentially suitable for 3 key types of assessment (early detection, broad-scale assessment and site-specific assessment) are listed as follows.

Early detection

- Bioaccumulation can be cost-effective where presence, absence or increase in tissue concentration is relevant to the issue or where bioaccumulation is relevant to community values other than aquatic ecosystems (e.g. concentrations of PFAS [per- and polyfluoroalkyl substances] in seafood). There is a vast literature on this topic – examples of relevant reviews include van der Oost et al. (2003) and DeForest et al. (2007).
- Biomarkers typically provide evidence of organism exposure to a stressor more readily than they measure effects resulting from that exposure (Hook et al. 2014).
- *In-situ* toxicity monitoring (e.g. Method 1A, Appendix 3, vol. 2 of ANZECC and ARMCANZ 2000, with revised protocol in Supervising Scientist 2011) can provide early detection of pulsed or continuous exposure to toxicants but can be expensive to implement.
- Genomics, including use of environmental DNA (eDNA) for analyses such as species assemblages and microbial processes, is a developing area and is currently mostly used as a research tool. However, successful applications for the derivation and assessment of sediment guideline values highlight the value of the methods (e.g. Harford et al. 2013, Chariton et al. 2014, Laroche et al. 2018).
- Rapid assessments (e.g. Method 3A(i), Appendix 3, vol. 2 of ANZECC and ARMCANZ 2000, the Australian River Assessment System [Chessman 1995]) may be relatively fast to mobilise

compared to other indicators but will typically be less sensitive and specific, particularly for subtle or point-source impacts. However, they may be useful for identifying locations that need more management or more detailed assessment using other indicators. They can also provide necessary cost and efficiency benefits for broad-scale (e.g. catchment-level) assessments (see below).

Table 1 Broad indicator types, suitable scenarios and examples for assessing biodiversity, conservation status and ecosystem-level responses

Indicator type	Scenarios suitable for indicator type	Examples
Sub-organismal	Exposure to a substance infers increased risk of impact or Physiological impairment of key taxa is a management concern or Reliable monitoring at individual or population level is difficult to achieve (e.g. remote area, highly variable flows, after a fish kill)	Bioaccumulation monitoring Biomarker assessment of potentially exposed species Effect biomarker measurement in a variety of collected organisms (e.g. plasma sorbitol dehydrogenase or stress-hormone levels at control and impact sites)
Individual organism	Collection of individuals is more feasible than population sampling or species is of high conservation value or Presence or absence of key (e.g. culturally significant) taxa is important or Reliable population assessment is difficult (e.g. highly variable environment)	Health measurement (e.g. condition index) of individuals of target taxa Visual census of freshwater fish species Deployment of caged bivalves for bioaccumulation monitoring
Population parameters	Population is the management target or Population is 'keystone' species or provides habitat for the target community	Threatened or socioeconomically important species Coverage by a species of seagrass
Community measures	Community is the management target or Community provides habitat or resources for rest of ecosystem or Community measure is a surrogate for (i.e. closely related to) diversity or ecosystem function	Structure of freshwater fish communities Vegetation structure in wetlands Macroinvertebrate community structure
Measures of ecosystem processes	Process is essential to functioning of the system or Links to structural attributes are well demonstrated	Measures of gross primary production and respiration in streams Measures of decomposition potential in sediments

Broad-scale assessments

- Remote sensing can be used to perform habitat mapping and analysis (e.g. seagrass density and extent, riparian vegetation extent and health, wetland macrophyte cover, vegetation type – including using these as proxies for biodiversity; see Method 5, Appendix 3, vol. 2, ANZECC and ARMCANZ 2000 for another example). Similar to rapid assessments for early detection, remote sensing may provide broad coverage in a cost-effective manner and may be useful for identifying areas that require further management focus or more detailed assessment using other indicators. Rapid advances in drone technology, sensor types and artificial intelligence for image classification mean that these technologies are now capable of providing detailed assessments of a broad range of indicators with high spatial resolution in near-real time (Gray et al. 2018).
- Chlorophyll measurement using remote sensing typically has strengths and weaknesses that are comparable to other remote-sensing approaches (e.g. habitat mapping) but it can provide very broad coverage quickly, particularly where satellite data are adequate for the assessment needed (i.e. where surface or near-surface spectral data are sufficient). See Ritchie et al. (2003), Gholizadeh et al. (2016) and Groom et al. (2019) for current status of remote sensing technologies for measuring chlorophyll *a*.
- Rapid assessments (e.g. the Australian River Assessment System [AUSRIVAS]) are relatively fast and cost-effective to mobilise at broader scales compared with other indicators but are typically less sensitive for point-source stressors. They may be useful for identifying locations that need more management attention or more detailed assessment using other indicators.
- Genomics, including use of eDNA for analyses such as species assemblages and microbial processes, is a developing area and is currently mostly used as a research tool. However, it has been successfully applied as a monitoring technique in Australia and internationally (e.g. Robson et al. 2016).
- Zooplankton surveys, such as the Australian Continuous Plankton Recorder Survey (Richardson et al. 2006), measure plankton communities as a guide to the health of Australia's oceans.

Site-specific assessments

- Many ecosystem receptor indicators are potentially suitable. Ideally, a number of indicators should be used for a weight-of-evidence assessment approach. Biodiversity indicators should focus on groups known to be sensitive to the stressors of interest (relative to other potential indicators). For toxicants and physical and chemical stressors, a starting point is to examine the relevant toxicant technical brief (or other such published information) to identify sensitive taxonomic groups (e.g. including the most sensitive taxa from the species sensitivity distribution [SSD] used to derive a default guideline value [DGV]). Similarly, biomarker, bioaccumulation and toxicity indicators should be responsive to the stressors of concern (again, being guided by the DGV SSD or other relevant information). In some cases, direct toxicity assessment (DTA) or the development of site-specific guideline values may be necessary.
- Rapid assessment approaches (e.g. AUSRIVAS) would not usually be recommended for this type of assessment.
- For approved wastewater discharges, there may be indicators included in the discharge permit that are required to be included in the assessment.

- For a known waste-discharge location, sampling-site designs should be selected to optimise inferences about indicator responses to the discharge (see [Considerations for ecosystem receptors](#)). For information on the analysis of data from these types of designs, see [Data analysis considerations for ecosystem receptors](#).

Further advice on selecting indicators is provided in Appendix A: Additional guidance for matching biodiversity indicators to problems. Note that the guidance in Appendix A: Additional guidance for matching biodiversity indicators to problems is only the starting point to a thorough review of the applicability of the indicator types being considered for the pressures and stressors of concern. Additional guidance, including cautionary advice and updated protocols, is provided in recent literature reviews and ANZG (2018) information on recommended protocols for each state or territory.

Most ecosystem receptor indicators do not have DGVs. Changes in ecosystem receptor indicator responses are inferred from comparison with data collected under an agreed study design. That is, for most ecosystem receptor indicators, guideline values will be negotiated in terms of the maximum acceptable deviation from control or reference condition or departures from expected trends, with the negotiations based on the management questions being addressed and the candidate monitoring designs (also see Chapter 10 of Downes et al. 2002).

3 Ecosystem receptor lines of evidence and their indicators

This section describes 4 lines of evidence for ecosystem receptors: biodiversity, toxicity, bioaccumulation and biomarkers.

3.1 Biodiversity

The assessment of system ecology has been related to principally structural components of biodiversity, such as whole plant or animal assemblages or population parameters, including the occurrence of particular taxa. ANZECC and ARMCANZ (2000) also include the measurement of ecosystem function in the biodiversity line of evidence. More recently, attention has focused on either quantifying the functional traits of plant or animal communities or measuring ecosystem traits such as metabolism. Typical indicator types for the biodiversity line of evidence are:

- species of high conservation value or species important to the integrity of ecosystems (e.g. marine turtles, dugongs, endemic invertebrates)
- communities of organisms, such as phytoplankton, macrophytes (also constituting important habitat), corals, zooplankton, macroinvertebrates (including stygofauna for groundwaters) and fishes
- important ecosystem processes (e.g. measures of gross primary production and respiration, functional traits).

Significant advances in the measurement of biodiversity indicators in recent years have substantially increased the amount of information that can be obtained rapidly and inexpensively.

Biodiversity management goals are often linked to detecting changes at the population, community or ecosystem level. These levels are described in Section 3.1.1, Section 3.1.2 and Section 3.1.3, respectively. There are 2 main uses for indicators at these 3 levels: detecting a response to a stressor and assessing biodiversity or conservation status. These uses often overlap.

The selection of indicators depends on the management question being addressed, but there are occasions when an indicator is too difficult or costly to monitor. In these circumstances, alternative indicators need to be monitored instead. For example, management may be focused on the conservation of a rare fish species but numbers may be too low to monitor the population reliably and quantitatively. Therefore, the monitoring program should include an alternative indicator, such as the aquatic vegetation that provides crucial habitat for the fish.

3.1.1 Population-level indicators

Population-level indicators include measures for a given taxon, such as density of a species of fish or percentage cover by macroalgae. Sometimes, individual taxa are included as indicators because of their intrinsic value to stakeholders (e.g. species of high conservation value, or that are economically or culturally significant), but inclusion of such species in a program should still pass tests of specificity and relevance to the water-quality issue at hand.

For broad-scale water-quality issues, which involve surveying many sites across a landscape, the presence or absence of key taxa in each site may be a relevant measure of their population status. While this can provide substantial cost and efficiency benefits, it will only detect impacts that are severe enough to eliminate entire taxonomic groups (typically families) of organisms. As a result, occupancy modelling (MacKenzie 2006) has become popular for broad-scale water-quality investigations. Instead of estimating abundance, the focus is on estimating the proportion of sample units that support (are 'occupied by') a given species (Noon et al. 2012). Mackenzie and Royle (2005) provide some practical guidelines for designing occupancy surveys, while Gwinn et al. (2016) assess the incomplete and variable detection of fish by different sampling methods. Concurrently, rapid improvements in technologies such as automated remote cameras, sound recordings, use of animal sign and genetic techniques (e.g. eDNA; Rees et al. 2014) have enabled multiple samplings over short periods and have increased the opportunities to reliably estimate detectability (Noon et al. 2012).

There continues to be progress in sampling designs and methods for estimating abundance and densities of populations, particularly for large vertebrate species, as well as in the development of remote-sensing methods for vascular plants and other large sedentary species (Simmonds and MacLennan 2005, Dafforn et al. 2016). The literature should be consulted during the early stages of program development to ensure that the most recent developments are considered and that appropriate specialists are engaged in design and data acquisition.

3.1.2 Community-level indicators

Community-level indicators include all techniques that involve analyses of organism assemblages. The indicators may consist of recording the presence or absence of multiple taxa (community composition) or of estimating the relative abundance of each taxon (community structure).

Traditionally, community-level indicators required the sampling of whole organisms and typically involved the identification and enumeration of individuals to finely resolved taxonomic levels. There is a substantial literature on these approaches for many different groups of organisms, and well-established and accepted approaches have been developed for each group which will vary in costs and efficiency depending on the groups used. More recently, the rapid growth of ecogenomic techniques based on next-generation sequencing (NGS) means that thousands of species (and higher taxonomic and genetic levels) can be identified across macrobiota, meiobiota and microbiota in a single sample, compared to dozens or hundreds of taxa using visual identification (see 'Ecogenomics' section). Ecogenomic techniques can be used to support or complement traditional approaches, or as a replacement when the traditional whole-organism approaches are difficult, inefficient or expensive to implement. Once the data are acquired, they can be analysed by condensing into univariate indices or scores, or by using multivariate techniques. The analytical methods for acquiring ecogenomic community data are developing rapidly, but the sampling designs and statistical analysis methods remain broadly similar to those used for traditional community-level indicators.

Community-level indicators using whole organisms

The community composition or structure of algae, invertebrates (particularly infauna and benthic macroinvertebrates), and – to a lesser extent – fish, vascular plants and other taxa have been used in both marine and freshwater systems to assess water and sediment quality (Downes et al. 2002).

Algal, microinvertebrate (meiofauna) or macroinvertebrate species usually require identification using optical microscopy, which can be costly, slow and labour-intensive. However, the improved

commercial supply of these services in some regions over the last decade has substantially reduced the cost and turn-around time of analysis. Similarly, sampling riparian plants, fish and amphibians quantitatively requires considerable skill, especially for small taxa or juvenile stages. Accordingly, well-documented procedures for quality assurance and quality control (QA/QC) are essential. Videography and sonar methods for fish and vascular plant surveys that use artificial intelligence for image classification represent new promising approaches to accurate, safe and cost-effective monitoring under development in parts of Australia (Jalal et al. 2020, Crook et al. 2021, Olsen et al. 2020).

Ecogenomics

The power of large ecogenomic databases based on NGS to identify differences between sampling locations or times is considerable. Baird and Hajibabaei (2012) outline these techniques, and there are many ‘proof-of-concept’ studies demonstrating their potential to distinguish between impacted and unimpacted situations (e.g. McNerney et al. 2016, Harvey et al. 2017).

NGS techniques can be divided into 2 broad categories (Chariton et al. 2016):

- amplicon sequencing, which amplifies known or targeted RNA or DNA sequences from specific gene regions to provide a list of taxa from a site (e.g. Carew et al. 2013)
- ‘shotgun’ sequencing, in which small random fragments of the entire genomic contents of a sample unit are sequenced to provide a more broadly based estimate of the biodiversity of a site.

While these techniques provide advantages, they are not without biases and disadvantages. Table 1 of Chariton et al. (2016) summarises comparisons between NGS and conventional techniques in biomonitoring. In general, taxa abundances are not routinely inferable from such datasets, although some taxon-specific amplicon methods show promise in particular applications (e.g. Jarman et al. 2006, McLellan et al. 2010, Legge 2012, de Vos et al. 2018; also see Golob et al. 2017).

In species-specific studies, amplicon sequencing is important to ensure that the gene regions being amplified are unique or diagnostic of the taxon being studied (Carew et al. 2013, MacDonald and Sarre 2017). Shotgun sequencing may provide rich datasets on both community structure and function, including for previously intractable components of the ecosystem (e.g. microbes). While very promising, the main drawbacks include the limited databases available for assembling partial genomes and the need for ‘deep sequencing’ of diverse communities (Chariton et al. 2016).

Use of DNA from environmental samples to assess biodiversity or the presence or absence of particular species (eDNA techniques) requires careful field protocols to avoid false positives or false negatives. Shaw et al. (2016) found some protocol-specific discrepancies between eDNA and conventional fyke-netting surveys for fish in drought-prone rivers in South Australia but considered that further development could render estimates of abundance feasible.

Analysing community data: univariate summaries

The 2 main classes of univariate summaries are diversity indices and biotic indices.

Diversity indices require the separation of each taxon in the sample or collection. Most indices also require estimates of the number of individuals, biomass or cover of each taxon. Separation is often at the species level but is sometimes at higher taxonomic levels. While the taxa need to be separated,

they may not necessarily need formal identification. It is essential to ensure consistency in taxa separation between different operators, sampling sites and dates as part of QA/QC procedures.

There are many diversity indices. Their early popularity rested on assumptions about the way disturbances would affect the species-abundance distribution in communities. However, empirical support for these suppositions is variable (Washington 1984). Accordingly, the absolute values of these indices have little meaning, and their use for assessing water quality relies on a strongly justified sampling design in which control or reference conditions are included.

While diversity indices have some intuitive appeal (e.g. many stakeholders would expect poor water quality to result in low diversity), they have the following drawbacks (Magurran 2013):

- Indices differ in their sensitivity to various aspects of the species-abundance distribution. These differences can hinder rather than enhance communication with stakeholders.
- They are often sensitive to the number of organisms collected, making comparisons between sites or times difficult.
- Information on the identity of taxa is lost in the computation of an index. For example, 2 sites may have similar values of a diversity index but vastly different species compositions.
- The statistical properties of many diversity indices are poorly known or are likely to violate the assumptions of commonly used univariate statistical techniques.
- Diversity indices are not suitably for use in systems with naturally low diversity, particularly areas with high seasonal variability in species abundances and relative abundances.

Even so, a diversity index may form part of a set of line-of-evidence indicators regarded as desirable by stakeholders (e.g. remediation might be expected to increase diversity of a given assemblage). It would be an acceptable indicator provided the behaviour of the index under the anticipated conditions and planned study design is well understood.

Biotic indices have a long history in biological assessments of water quality. They are usually used to assess the effects of contaminants, mostly in rivers. Many biotic indices are specific to a site and contaminant type (usually organic). Compared to diversity indices, biotic indices are more clearly related to the conditions that led to their development.

The calculation of biotic indices usually requires:

- total counts of individuals or taxa
- counts (or biomass measurements) of specific taxonomic groups (e.g. mayflies, stoneflies and caddisflies in cool, freshwater habitats)
- details of the responses of different taxa to contaminants
- division of taxa into groups with different traits (e.g. growth forms in plants, feeding or respiratory adaptations in invertebrates).

One example of a biotic index is SIGNAL 2 in Australia (Chessman 2003), originally developed to assess changes due to sewage pollution in waterways in New South Wales (e.g. Chessman 1995, Gowns et al. 1995) but subsequently updated and adapted to other regions. Another is the macroinvertebrate community indices, including quantitative and semi-quantitative applications, in

New Zealand (Stark 1993, Stark and Maxted 2007). Aylagas et al. (2017) propose a microbial index that could complement or supplement these indices. Cost savings are often achieved with algal and macroinvertebrate data by using coarser taxonomic resolution than species-level identifications. More elaborate schemes involve computing several indices ('metrics') from the community data, with each metric reflecting a different aspect of community structure or function, often with protocols for converting raw scores for the metrics into assessments of ecosystem condition (e.g. Kerans and Karr 1994, Fore et al. 1996, Miller et al. 1988, Harris and Silveira 1999). Once the metrics are developed and tested for a particular region, they may be applicable over wide geographic areas with minor modification (Resh et al. 1995). However, Resh and Jackson (1993) emphasise that the metrics and their protocols need to be calibrated for different regions of the country and, perhaps, for different types of stressors.

Biotic indices and metrics often have diagnostic value and are frequently used in broad-scale monitoring programs, such as *Australia state of the environment* reporting, and other applications where some measure of ecosystem status or health is needed (e.g. preliminaries to development approvals, baseline studies for greenfield developments). Some biotic indices or metrics can be used as response variables in monitoring designs and data analyses, although their sensitivities, specificity and statistical behaviour must be well understood beforehand.

Analysing community data: multivariate methods

Similarity indices and multivariate analyses

Similarity indices and multivariate analyses identify taxa and infer impacts based on changes in community structure relative to control or reference conditions. Traditionally, these analyses were based on computing a (dis)similarity measure on all pairwise combinations of sample units in a program (Legendre and Legendre 2012) and then undertaking randomisation tests (also called permutation tests) on the resulting (dis)similarity matrix using the standard general linear model designs (e.g. from the MBACI [multiple before–after control–impact] family) for univariate responses (e.g. Anderson 2001, Ferrier et al. 2002, Anderson and Robinson 2003, Anderson and Ter Braak 2003). Alternative approaches have been proposed based on generalised linear models (Warton et al. 2012, Niku et al. 2019) and Bayesian extensions to account for residual correlations between species (Hui et al. 2016). There is a burgeoning literature on novel approaches to model-based inferences using 'messy' community ecological data (Warton et al. 2015a, 2015b, Ovaskainen et al. 2017, Ter Braak 2019).

Multivariate community data based on similarities are usually displayed using clustering, ordination or a combination of both. Various techniques are readily available to display diagnostic information about which taxa are responsible for the observed patterns in these displays and to link community differences to patterns in environmental measures (Clarke 1993, Borcard et al. 2011, Legendre and Legendre 2012). [Additional information](#) on multivariate analyses can be found in ANZG (2018).

Sometimes the similarities serve as measures of β -diversity (or species turnover) (Anderson et al. 2006) and can be used as a response variable in conventional univariate analyses as for paired designs (e.g. MBACI designs) (Faith et al. 1995, Supervising Scientist 2013).

Predictive modelling

[Applications of predictive modelling](#), including regression and neural networks, to water-quality assessments are covered elsewhere in ANZG (2018) and are not discussed further. This discussion

focuses on an Australia-wide, rapid biological-assessment approach using predictive modelling, built upon multivariate precepts – AUSRIVAS. Macroinvertebrates and associated habitat (characterised by a standard set of physical and chemical variables that are largely unrelated to likely pollutants) are sampled from appropriate reference sites. The group of reference sites is then classified according to their biota to produce groups of sites containing similar fauna. A numerical analysis is used to identify the environmental attributes which describe each group of reference sites. The environmental attributes measured from a monitoring or test site are compared with those of the reference sites to determine which group or groups of reference sites it resembles most closely. The fauna of these corresponding reference sites are then compared with the test site. If the test site supports fewer taxa than are predicted by the reference sites, it is judged to be disturbed.

ANZECC and ARMCANZ (2000), and affirmed here, advise against the application of AUSRIVAS for those small-spatial-scale applications (i.e. a specific activity, development or point-source disturbance within a catchment) requiring strong inference and where sensitivity to the water-quality pressures and stressors of concern is required. AUSRIVAS was chiefly developed for *Australia state of the environment* reporting across broad spatial scales. The advice noted the limitations of a static reference condition and that presence–absence, family-level data are useful only for detecting impacts that are severe enough to eliminate these higher order groups. Even some applications at broader scales have recently been questioned (Cox et al. 2019, Chessman 2021).

ANZECC and ARMCANZ (2000) note that if AUSRIVAS is to be adopted at small spatial scales, it should be conducted in a design framework that has the necessary sample and site replication, as well as baseline data, to allow the study objectives to be met. Aspects of the rapid biological-assessment protocol can be adapted or modified so that the data gathered are amenable to both AUSRIVAS and quantitative assessment. The latter is described above ('Similarity indices and multivariate analyses') and elsewhere in ANZG (2018).

Ensuring quality of evidence in chosen biodiversity indicators

To ensure adequate inference arising from the study of any chosen biodiversity indicator, particularly for site-specific studies, the application should ensure key study design and analysis precepts are incorporated. These include adequate spatial and temporal replication for sampling and, where community data are gathered, additional analyses and outputs beyond single-number metrics such as observed-to-expected (O/E) ratios, biotic indices or similarity indices. A more comprehensive analysis and output suite for community data that compare 'exposed-site' data to those from an appropriate reference condition dataset(s) would include community summaries such as relative abundance and taxa number, descriptive and exploratory multivariate methods including ordination, multivariate hypothesis testing, and analyses that relate the community structure data to environmental variables and identify influential taxa that separate any multivariate structure in the data. Ensuring quality of evidence for site-specific assessments should be carefully considered within the context of the WQMF, where:

- The weight-of-evidence process (in [Step 3](#) of the WQMF) requires that the selected group of indicators is accompanied by an appropriate quality-of-evidence statement that reflects the quality of the indicators. Less robust indicators (and groups of indicators) would clearly be assigned a lower quality-of-evidence rating that would translate to a lower confidence in the conclusions of the assessment.

- The selected indicators and the water-quality objectives they would be compared with need to be approved with the stakeholders in [Step 5](#), taking the quality-of-evidence statements into account.
- [Step 7](#) of the WQMF provides the ability to add additional indicators if those assessed are considered to be insufficient or give a result that suggests more detailed investigations are required.

Local jurisdictions should be consulted if there is any uncertainty over the selection of rapid biological-assessment approaches for specific purposes.

3.1.3 Ecosystem function indicators

Ecosystem function indicators can be divided into 2 types. First, there are direct measures of ecosystem functions or processes such as community metabolism or net primary production. Second, functional traits (or simply 'traits') are used to measure ecological communities in terms of the functional attributes of their constituent species (e.g. the abundance of organisms with traits that facilitate tolerance of salinity). While direct measures of ecosystem function can be holistic performance indicators of key processes and can map to end-user understanding of ecosystem services, they also ignore the identity of component taxa. Accordingly, such measures are likely to be used with other 'structural' indicators if biodiversity is a management objective.

Functional traits

This refers to procedures that summarise the functions performed by the constituent species of a community, either as a complement to (Stuart-Smith et al. 2013) or substitute for taxonomically based measures of community structure (Schäfer et al. 2011, Kefford et al. 2020). For freshwater macroinvertebrates, an early attempt consisted of the functional feeding groups (ANZECC and ARMCANZ 2000). Some of the component metrics of the indices of biotic integrity for fish and macroinvertebrates (see Section 3.1.2) were early attempts at summarising functional attributes for these 2 biological communities. Classification schemes for a variety of functional traits have been developed for macroinvertebrates (Poff et al. 2006, Schäfer et al. 2011, Kefford et al. 2012), fish (Merigoux et al. 2001, Bergerot et al. 2015) and plants (Quetier et al. 2007), with applications for aquatic biomonitoring and assessment (Dolédec and Statzner 2008, Van den Brink et al. 2013, Arce et al. 2014).

Some biotic indices assume that the ratios of organisms with different feeding strategies will change with contamination (e.g. collectors will be more abundant than shredders under contaminated conditions), or that trophic generalist organisms will be more tolerant to contaminants than trophic specialist organisms. There is some doubt that these general rules hold true and that it is possible to assign taxa to different feeding strategies (Resh and Jackson 1993).

Organisms can be assigned to groups of species that feed in the same way, generally known as functional feeding groups. These can be used to assess the dynamics of food supplies in a waterbody and the balance of feeding strategies (food acquisition and morphology) in the community assemblage. Deviations from patterns at reference or control sites generally indicate stressed conditions. Functional feeding groups have been defined for bacteria and other microbes (e.g. photosynthetic autotrophs, bacterivores, saprotrophs), macroinvertebrates (e.g. scrapers, piercers, predators, shredders) and fish (e.g. piscivores, invertivores, algivores). The usefulness of functional feeding groups for assessing water quality has not been well demonstrated for benthic

macroinvertebrates, and the concept is not considered reliable in this case (Karr and Chu 1997). However, functional feeding groups may help identify the effects on the community of stressors other than water quality. For example, flow alterations may alter the relative abundances of feeding groups (Tullos et al. 2009) and aid the interpretation of field results.

Community metabolism and nutrient pathways

The community-metabolism approach is based on the concept that movement of organic carbon through an ecosystem can be used to measure community metabolism (see Section 8.1.2.1 of ANZECC and ARMCANZ (2000) for a detailed description). Closely allied to this approach are methods focused on tracing the movements, transformation and metabolism of key nutrients. These provide indicators of ecosystem health and may also improve system understanding of aspects such as lags and interactions with physical components of the aquatic system (Swanson et al. 2017).

For the community metabolism approach, 2 biological processes affect the movement of carbon: gross primary production (P) (via photosynthesis), and respiration (R). Community metabolism is considered to be sensitive to small changes in water quality (particularly organic contaminants and sedimentation) and to conditions that affect light input (e.g. Hopkinson and Vallino 1995, Bernot et al. 2010, Clapcott et al. 2016). As a result of this sensitivity, the community-metabolism approach may be able to detect a disturbance early, before it manifests in changes to organism assemblages (e.g. macroinvertebrate community composition).

The ratio of P:R measures ecosystem health in rivers and streams. Undisturbed forest-stream sites are typically heterotrophic (i.e. $P:R < 1$) and so are net consumers of carbon. Sites that are in cleared catchments or are nutrient enriched are typically autotrophic (i.e. $P:R > 1$). Consequently, the P:R ratio is an index that applies to freshwater, although there can be considerable regional differences in estimates of P and R and, therefore, P:R (Bernot et al. 2010, Clapcott et al. 2016). The community-metabolism approach is included in the suite of ecosystem-health monitoring programs developed by the Healthy Land and Water Partnership for the Moreton Bay catchments in south-east Queensland (Healthy Land and Water 2021). A variety of methods have been developed for rivers and streams of different sizes (Lamberti and Hauer 2017), and recent literature should be consulted to ensure that the methods chosen are up to date.

Measuring the rates and fates of key aquatic nutrients is important for interpreting and understanding the chemical and biological indicators of marine (e.g. Eyre and Ferguson 2009) and freshwater (e.g. Burrows et al. 2013) systems. As with the community-metabolism approach, methods must be appropriate for the spatial extent and temporal resolution of the program. The advantages of these methods are that they can directly measure ecosystem processes and are easily understood by stakeholders (e.g. denitrification rates in effluent treatment [Oakes et al. 2011]). Conversely, most measures of ecosystem function can only be implemented in field surveys. Demonstrating causal links, especially with toxicants, can be problematic for large-scale problems (Rohr et al. 2016). Rohr et al. (2016) also emphasise the uncertainty of links between ecosystem process measures and structural measures of biodiversity. For example, changes to the structure of the community or ecosystem may not manifest as changes to ecosystem processes, because the processes may be mediated by very different species. However, the structural changes may also be important to stakeholders because some species may be more valued than others (e.g. it may not be acceptable for an introduced or invasive species to substitute a native species to maintain an ecosystem function).

3.1.4 Additional guidance

Additional guidance on selecting biodiversity indicators is in Appendix A: Additional guidance for matching biodiversity indicators to problems.

3.2 Toxicity

The determination of the toxicity of water (i.e. wastewater or ambient water) or sediment samples is a form of rapid biological assessment that can be used as a line of evidence for assessing water or sediment quality. Toxicity assessments can be tailored to various environmental scenarios, including single or multiple toxicants, different durations (acute or chronic) and frequencies (pulsed, continuous) of exposure, and site-specific water-quality conditions that may affect toxicant bioavailability and toxicity. They can also be conducted in the laboratory or field. These aspects are discussed throughout this section.

Acute (i.e. short-term) toxicity can usually be determined within 4 days. Chronic tests, with the exception of microbial and algal bioassays, and sensitive early life-stage tests (e.g. echinoderm 1-hour egg fertilisation) require longer (see Warne et al. 2018, Table 1). Toxicity assessments typically require toxicity tests for at least 5 species (but preferably 8 or more) covering at least 4 different taxonomic groups. Fewer species may be acceptable depending on the assessment objective. Species vary in sensitivity to different stressors. For example, herbicides are generally more toxic to plants than to animals, and invertebrates are typically more sensitive than plants to insecticides. There are many influences on sensitivity, such as the nature and development of regulatory and excretory mechanisms (often making early life-history stages more sensitive than later life-history stages), and the exposure pathway (for example, absorption across gill surfaces commonly results in increased uptake compared to uptake via the gut). The drivers of sensitivity can be difficult to determine *a priori*, so generally a suite of species from different taxonomic or trophic groups is recommended (see 'Minimum requirements for using direct toxicity assessment to assess water quality' for more details). For the toxicants with DGVs, examination of the SSD can provide guidance on the sensitive species for a given toxicant.

Data from laboratory, single-chemical and single-species toxicity tests form the basis for deriving chemical-specific water-quality DGVs for Australia and New Zealand. As such, ANZG (2018) recommends direct toxicity assessment (DTA) as a useful technique for assessing the effects of mixtures of compounds in ambient waters (e.g. industrial effluent, wastewater treatment plant effluent, landfill leachates). In some countries (e.g. the USA), DTA is termed 'whole effluent toxicity' (WET) testing. The ANZECC and ARMCANZ (2000) guidance for toxicity testing remains largely valid today. However, Sections 3.2.1 and 3.2.2 describe this guidance with various updates as necessary.

3.2.1 Single-chemical toxicity testing

Benefits of single-chemical toxicity testing

A major benefit of single-chemical toxicity testing is that specific cause–effect information can be obtained on the toxicity of a particular chemical. This information is used to derive water-quality guideline values for the protection of aquatic ecosystems. Definitive limits can be set and, assuming there is an analytical detection method for the compound, it can be readily monitored in aquatic environments. In addition, the majority of single-chemical toxicity tests are carried out in the laboratory, where effects can be studied under controlled conditions with a limited number of variables (Sprague 1990). Assuming that such experiments are carried out correctly, there is a large

degree of certainty that the observed effects are caused by the chemical alone. For the majority of compounds, single-chemical toxicity tests are an appropriate way of determining toxicity and deriving water-quality guideline values.

Alternative approaches exist and may be more appropriate in certain circumstances, especially where site-specific guideline values are needed. Van Dam et al. (2019) provide guidance on how to select and evaluate approaches for deriving site-specific guideline values.

Limitations of single-chemical toxicity testing

An organism will rarely be exposed to just one toxicant in the environment (Sprague 1990), so single-chemical toxicity testing is not representative of the natural environment. In most circumstances, a particular chemical will be present in combination with many other chemicals, and interactions that alter their toxicity may occur (Holdway 1992, de Zwart and Posthuma 2005). Subsequently, mixtures of chemicals can result in additive toxicity, greater-than-additive toxicity (also known as synergism) or less-than-additive toxicity (antagonism) (Rand 1995, de Zwart and Posthuma 2005). Single-chemical toxicity tests do not account for such factors, and the extrapolation of the results to environmental impacts is uncertain. While methods exist for predicting the toxicity of mixtures by using data from single-chemical toxicity tests (Marking 1977, Warne 1998), and recent reviews have verified that the assumption of additive toxicity is the best default (Syberg et al. 2009, Verbruggen and Van den Brink 2010, Cedergreen 2014), this approach is only effective when all contaminants within an effluent are known. This is generally not the case for complex effluents and wastewaters.

While strict control of all variables except the one of interest is usually considered a benefit in laboratory experiments, it has also been recognised as their major limitation (Rand 1995). Manipulation of environmental factors can be incorporated into a laboratory toxicity test (e.g. water hardness for metals) but they cannot simulate all aspects of the natural environment. Other limitations include the use of a constant toxicant concentration, the use of a limited range of standard test organisms, and the need for optimal living conditions for test organisms. Additionally, condition and resilience of laboratory-cultured organisms can be very different to wild organisms. It is difficult to be sure that effects observed in laboratory experiments resemble those in the natural environment.

Various methods have been used to minimise limitations of single-chemical toxicity tests, including:

- the use of safety or assessment factors – a practice widely used in the derivation of water-quality guideline values worldwide, although questioned more recently (Batley et al. 2018)
- focusing on data from the most sensitive species tested
- using alternative statistical estimates, such as the EC10 (concentration generating a 10% effect response in the test population) as opposed to the EC50 (concentration generating a 50% effect response in the test population).

In addition, there has been much research focused on making the assessment of the effects of aquatic contaminants more realistic, such as the development of more relevant toxicity-test protocols, including laboratory multi-species and microcosm tests, outdoor mesocosm tests, *in-situ* tests, and tests for determining the toxicity of complex mixtures such as effluents and urban and industrial runoff water (see Section 3.2.4 for further details on field-based approaches such as *in-situ*

testing and mesocosms). Comparisons between outputs of single-species toxicity test-based SSD approaches and results from mesocosm and field experiments have generally indicated that:

- they are similar in terms of estimating protective concentrations (e.g. Versteeg et al. 1999, Hose and Van den Brink 2004)
- the use of single-species tests can under-predict toxicity depending on the mode of action of the toxicant (e.g. De Laender et al. 2009)
- protection is best afforded by concentrating on the taxa that are most sensitive to the toxicant (e.g. Van den Brink et al. 2006).

Therefore, there is no consensus of the efficacy of using an SSD based on single-species tests alone. However, the inclusion of taxa known to be sensitive to a toxicant mode of action (e.g. plants for herbicides, invertebrates for insecticides) generally provides adequate protection for real-world, multi-species exposures.

While this appears to be true for individual toxicants, it is less certain for mixtures of toxicants. Section 3.2.2 discusses the concepts behind and the advances in toxicity testing of complex mixtures of toxicants.

3.2.2 Direct toxicity assessment

DTA (or WET testing) is a well-established approach in ecotoxicology that has generally been adapted from conventional single-chemical toxicity testing approaches, using similar methods, species selection and extrapolation to receiving waters (Mount 1986). The types of mixtures that can be assessed include urban runoff water, sewage discharge, mining wastewater, petroleum production wastewater, agricultural runoff water, industrial effluent or any combination of compounds that occurs in or is likely to enter the environment and for which the toxicity is unknown. DTA can also include the assessment of the toxicity of ambient (natural) water that receives contaminant inputs. Therefore, DTA differs from single-chemical toxicity testing in that the combined effects of a number of compounds of unknown identity and concentration are assessed, as opposed to the effects of just one chemical.

Grothe and Johnson (1996) state that the primary aim of WET testing (and so DTA) is to ensure that wastewater releases into the aquatic environment do not harm aquatic life. This can also be broadened to account for physico-chemical changes in water quality, such as eutrophication, hypoxia and salinisation. DTA aims to do so by measuring the overall toxicity of a mixture of compounds. It is generally not concerned with individual components.

As with other methods, DTA has benefits and limitations. A more comprehensive review of the topic is provided by de Vlaming and Norberg-King (1999).

Benefits of direct toxicity assessment

DTA is an important tool for assessing the toxicity of complex wastewater and receiving (or ambient) water where there are thousands of unidentified components. The effects of complex mixtures cannot usually be predicted by determining the toxicity of individual components, which typically change with time and are often not fully known (Holdway 1992). DTA provides an integrated measure of the toxicity of all the chemicals within a mixture (de Vlaming et al. 2000) and so accounts

for interactions between compounds. Therefore, DTA (compared to single-chemical toxicity testing) more closely resembles the natural environment.

DTA also:

- provides a direct measure of toxicity and bioavailability
- is a reliable qualitative predictor of biological community impacts (Waller et al. 1996, de Vlaming and Norberg-King 1999, de Vlaming et al. 2000)
- provides early-warning capability so management actions can be implemented to minimise ecosystem impacts (van Dam et al. 1998, de Vlaming et al. 2000)
- can be performed relatively quickly and at less cost compared to other biological monitoring procedures (de Vlaming and Norberg-King 1999, de Vlaming et al. 2000).

Even so, the use of this approach internationally is dependent upon the regulatory regime in place (Power and Boumphrey 2004).

Limitations of direct toxicity assessment

Although it is considered more representative of the natural environment (compared to single-chemical toxicity testing), DTA also has its limitations (see also Chapman 2000). DTA can assess the toxicity of a mixture as a whole but it does not identify the toxic components of a mixture (Jop et al. 1991). While the toxic components might be obvious for a simple wastewater containing only a few well-defined contaminants, for the majority of cases there will be too many chemical components to easily identify those that are toxic. Identifying the toxic component(s) of a wastewater is an essential step in addressing a toxicity problem and improving treatment technology. DTA alone cannot provide this. DTA is only one step in an overall assessment of water quality (Chapman 1995, 2000). Specific methods for identifying the toxic components of effluent (toxicity identification evaluation [TIE]) exist and are discussed briefly in the '[Toxicity identification evaluation](#)' section.

Due to the variable nature of wastewater and effluent, and because their compositions are usually unknown, it may be difficult to obtain a representative sample of the mixture (Mount 1986). Therefore, the one-off testing of a chemical mixture will give little meaningful information if the representativeness of the sample is unknown. Repeated testing or continuous monitoring is desirable but may not be cost-efficient.

Standard protocols for preparing effluents for DTA are critical. Protocols have been developed in North America (e.g. US EPA 2002a, 2002b, 2002c, Environment Canada 1990a, 1990b, 1992a, 1992b, 1992c, 1992d) and are in practice at several laboratories in Australia and New Zealand since the release of ANZECC and ARMCANZ (2000). For commercial laboratories, these are often commercial-in-confidence documents, but some publicly available examples are available (e.g. Hall and Golding 1998, Trenfield et al. 2020). Aspects of effluent preparation include collection, storage, filtration, dilution, adjustment of physico-chemical parameters and aging. Aging is particularly important, as it relates to the persistence of chemicals and, consequently, toxicity (Mount 1986). In addition, chronic tests may exceed the optimum effluent holding storage time, in which case the experiment is conducted using different samples at different times. While this might increase the environmental realism of the assessment, it may also increase variation and uncertainty (Pifher and Egan 1989). Filtration is also a vital step, as micro-organisms (e.g. bacteria) and macro-organisms (e.g. predatory

copepods) can interfere with toxicity if effluent is not filtered correctly (Grothe and Johnson 1996). However, filtration can significantly reduce the environmental realism of an effluent sample. Even so, it is an important step in distinguishing the physical from the chemical stressor effects of an effluent. This can be important for guiding management decisions regarding treatment of the effluent.

The selection of appropriate DTA methods is a critical step. In the USA, standard methods are used to determine effluent toxicity. This has been criticised by industry (Pifher and Egan 1989). There has been a call for environmentally representative testing, although this too has been subjected to criticism. The advantages of standard and site-specific DTA are discussed in the following section.

Standard versus site-specific direct toxicity assessment

The basic differences between standard and site-specific DTA are in the methods. Standard DTA was developed to standardise ecotoxicological processes, enable comparison of the results of experiments conducted on different effluents from similar industries, and encourage the generation of scientifically sound data. Standard DTA is usually conducted with standard organisms, a standard synthetic water, and standard test conditions (e.g. pH, temperature, dissolved oxygen), duration and endpoints. An important criterion in the selection of suitable standard species is sensitivity to a wide range of toxicants. Combined with the use of conservative exposure conditions and application of safety factors to the results, this makes standardised tests more likely to be overprotective to the aquatic ecosystem of interest (Chapman 2000, Chapman et al. 1998). However, standard DTA is generally not representative of local environmental conditions and so has limited applicability for making conclusions about potential environmental effects. That is, a significant effect from a standard DTA does not necessarily mean there is a problem in the receiving water (Pifher and Egan 1989) while a negative result does not necessarily mean the wastewater is not impacting the receiving water.

For site-specific DTA, the tests are designed according to relevant environmental conditions. Organisms local to the area are chosen as test species (preferably with already-verified test protocols), and the local ambient water (upstream of the effluent source or from another appropriate reference site) is used as the control and dilution water. In addition, conditions such as test duration and test endpoints can be manipulated to best represent the nature of exposure to a particular effluent. Compared to standard DTA, site-specific DTA is more representative of the natural environment. However, site-specific DTA is not suitable for comparisons with other effluents if, for example, different species are used or the ambient water chemistry differs. In addition, using upstream water as dilution water may introduce other variables, such as background toxicity from further upstream (Pifher and Egan 1989). Finally, using local ambient water as control and dilution water increases the complexity of the testing and presents problems related to background effects, although treating and filtering the water may alter the toxicity of contaminants (Ruffier 1996). Examples of site-specific toxicity testing protocols, including the use of local species and local water, are provided by Trenfield et al. (2020). An example of the incorporation of locally relevant pulse-exposure conditions into site-specific toxicity tests to assess the effects of a locally relevant toxicant is provided by Hogan et al. (2013) and Prouse et al. (2015).

The development of site-specific toxicity test methods for local species requires significant investment, which needs to be weighed against the demand for application of such tests (van Dam et al. 2019). In some cases, efficiencies can be gained by selecting local species for which standard

toxicity-test methods for related species can be adapted (rather than developing new methods from scratch). When developing site-specific test methods, various aspects need to be considered, including species selection, endpoint selection, control or diluent water selection, experimental design, test conditions selection and optimisation, statistical analysis and QA/QC. It is beyond the scope of this report to provide detailed advice on these aspects. However, some information can be found in ANZECC and ARMCANZ (2000; Sections 8.3.6.7 and 8.3.6.8).

Another approach to site-specific testing is *in-situ* or in-stream testing, where organisms are exposed to the receiving water or wastewater in the environment (e.g. fish kept in cages). Examples of such approaches can be found in Supervising Scientist (2011) and Maltby et al. (2000). Effects are monitored and compared to organisms kept either upstream of the contaminant source or in a designated reference area. *In-situ* toxicity testing more easily accounts for the temporal variability of the wastewater or receiving water quality, including pulse exposures, than laboratory toxicity testing. *In-situ* testing is useful for assessing environmental impacts and determining confidence in predicting impacts from laboratory studies. For example, Eagleson et al. (1990) document the results of 43 comparisons between laboratory DTA and *in-situ* surveys. They found there was 88% agreement between the laboratory test and *in-situ* test. Several other studies have concluded that DTA, if used properly, is a reliable qualitative predictor of aquatic population impacts (Waller et al. 1996, de Vlaming and Norberg-King 1999, van Dam et al. 2014). Other studies have demonstrated poor correlations between effects measured in the laboratory and in the field (Clements and Kiffney 1994, Sarakinos and Rasmussen 1998), although there does not appear to be a unidirectional trend of laboratory bioassays overestimating or underestimating field effects (Chapman et al. 1998). So, field validation of laboratory experimentation is rarely possible (Chapman 2000). Finally, with *in-situ* testing, care needs to be taken to minimise risks associated with spread of diseases and parasites if introducing cultured species into natural systems. A summary of other field-based approaches for assessing site-specific impacts of physical and chemical stressors, including toxicants, is provided in Section 3.2.4.

In some situations, it may not be possible to collect, culture and test enough species that are truly local to the environment of interest. Where this is the case, appropriate surrogate species that are regionally relevant (rather than locally relevant) or that are from a similar ecosystem and water quality are preferred over standard species. In such situations, it is recommended that 2–3 locally relevant species that are taxonomically similar to some of the regionally relevant test species are assessed alongside the regionally relevant species. By comparing the results for the taxonomically similar locally relevant species and regionally relevant species, the appropriateness of using the regionally relevant species can be determined. An example of this is described by van Dam et al. (2014) for a diamond-mine wastewater.

The adoption of standard or site-specific DTA must be determined based on the objective of an investigation. For the purposes of Australian or New Zealand water managers, who generally oversee specific hydro-geographical regions and are concerned with local water quality, site-specific DTA is likely to be the most appropriate approach. In some cases, a hybrid approach will be appropriate or, at least, the best that is achievable. Guidance on selecting different approaches for deriving water-quality guideline values, much of which is transferable to DTA, is provided by van Dam et al. (2019).

Toxicity identification evaluation

TIEs are a set of toxicity-assessment procedures developed and modified to identify toxic components of effluents or contaminated natural waters (Jop et al. 1991, Maltby et al. 2000). They involve manipulating and fractionating the effluent or natural water and conducting subsequent toxicity tests to separate toxic components from non-toxic components (Burkhard and Ankley 1989). TIE methods have been extensively developed in North America (Jop et al. 1991, Norberg-King et al. 1991, Norberg-King et al. 1992, Durhan et al. 1993, Mount and Norberg-King 1993). They can be undertaken following DTA if necessary. Some TIE methods have also been developed in Australia (Manning et al. 1993, Pablo et al. 1996, Bailey et al. 2000a, 2000b, Strom et al. 2009) and New Zealand (Hall and Golding 1998). TIE has become an increasingly important tool but guidance for its use is not within the scope of the ANZG (2018) guidelines. The literature should be consulted for detailed information and guidance.

Application of direct toxicity assessment

Ideally, DTA should be carried out on every discrete mixture of chemicals that is known to enter the aquatic environment. However, this is impossible in practice. In addition, state and federal government legislation determines the priorities and uses for testing and whether DTA can be used as a regulatory tool. In the USA, there was considerable disagreement between government and industry as to whether WET testing, with its associated limitations, should be used to determine compliance with regulations (Pifher and Egan 1989, Moore et al. 2000). However, WET testing has now become an important component of many industrial and municipal National Pollutant Discharge Elimination System permits throughout the USA (Grothe et al. 1996), Europe and Canada (Power and Boumphrey 2004).

Potential applications of DTA include industries or processes where discrete complex effluent or wastewater is released into aquatic ecosystems. This could include wastewater from industries such as mining and minerals processing, oil and gas, pulp and paper, sewage treatment, seawater desalination and power generation, as well as from urban runoff water. If a water-quality monitoring program of a receiving water already exists, this should be carried out in conjunction with DTA of the specific discharge, as well as DTA of the receiving water. Due to the large number of chemicals likely to be present in wastewater, it is possible that some compounds could be missed if only suspected priority contaminants are measured. The use of DTA (in conjunction with measuring priority contaminants) overcomes this limitation as it integrates the toxicity of all the compounds in a complex mixture. Toxicity testing of wastewater prior to release into the aquatic environment aims to prevent contamination of a receiving water with wastewater that is toxic to aquatic life and to monitor the performance of wastewater treatment facilities. In addition, toxicity testing allows the determination of site-specific guidelines, such as 'safe' wastewater dilution and release rates. Van Dam and Chapman (2001) provide a table of potential applications of DTA as well as several, albeit relatively old, case studies of the use of DTA.

Where there are multiple small discharges to a single receiving water and individual sources are difficult to identify, it may not be economically or technically practicable to undertake DTA on all discharges. In such situations, DTA of the receiving water (rather than for each discharge) is the most appropriate option. Subsequently, if toxicity in a receiving water is observed, DTA of individual discharges may help to identify the source of toxicity.

In addition, many contaminants enter waterways over a broad spatial scale, with no particular point source, making assessment of their toxicity difficult. Again, DTA of the receiving water can be used. Alternatively, experiments can be designed to catch runoff water for laboratory DTA, or mixtures such as leachates can be prepared in the laboratory, following standard methods, for laboratory testing. Essentially, if an area is suspected to be contaminated, DTA can either be carried out *in situ* or in the laboratory using collected water samples and using representative clean water from elsewhere (e.g. upstream) as dilution water.

Additional applications of DTA include determining if naturally elevated background levels of inorganic compounds represent a risk to aquatic life or if other site-specific characteristics, such as salinity, pH and dissolved organic carbon, change the toxicity of particular compounds or mixtures of compounds. DTA can also be applied to assessing the bioavailability and toxicity of one or more toxicants under site-specific conditions (i.e. local species, local dilution water) to derive a site-specific guideline value (see van Dam et al. 2019).

Minimum requirements for using direct toxicity assessment to assess water quality

The data requirements for deriving a site-specific 'acceptable' dilution for an effluent, or a site-specific guideline value for a locally-relevant toxicant, based on DTA, are similar to those for DGVs for single chemicals (Batley et al. 2018, Warne et al. 2018). Some of the key requirements include:

- Data must meet appropriate quality requirements.
- Chronic toxicity data are typically preferred over acute toxicity data.
- Toxicity estimates such as no-effect concentrations or low-effect concentrations are preferred over no-observed-effect concentrations, but must be supported by appropriate experimental designs. See 'Glossary and acronyms' for definitions.
- A suite of toxicity tests using species from a range of taxonomic groups should be used because, for a variable, complex effluent, it cannot be certain which groups of organisms will be most sensitive for any particular period of sampling.

The reliability classification for toxicant guideline values requires chronic toxicity data for at least 8 species (from 4 taxonomic groups and with a good fit of the SSD model) for a high reliability guideline value and at least 15 species (from 4 taxonomic groups and with a good fit of the SSD model) for a very high reliability guideline value (Warne et al. 2018). For DGVs, where the toxicity data are collated from the national and international literature, a sample size of ≥ 15 can often be attained. However, the toxicity testing of effluent or ambient water is often an ongoing process due to their changing characteristics over time, and the regular assessment of toxicity to at least 8 (let alone 15) species can be difficult to achieve but is nevertheless encouraged. It may not be feasible to derive a high or very high reliability guideline value (or a 'safe' effluent dilution) for an effluent. Stakeholder expectations for this must be appropriately aligned.

Although Warne et al. (2018) do not specify types of taxonomic groups that need to be represented in a dataset for deriving guideline values, typically these might include aquatic algae, macrophytes, crustaceans, gastropods, bivalves, insects or fish. Although the minimum requirement of 4 taxonomic groups will often be achievable, this may be more of a challenge in some situations (e.g. temporary water ecosystems). In such cases, regionally relevant species may be used as discussed in Section 3.2.2 ('Standard versus site-specific direct toxicity assessment'). However, the appropriateness of this

would need to be discussed with regulators beforehand. If there are insufficient local or even regional species for deriving site-specific guideline values, then alternative methods should be explored, as discussed by Smith et al. (2020).

It is important to recognise that the reliability classification was developed primarily for DGVs. It does not take into account the additional relevance of the information obtained from testing locally relevant or regionally relevant species in locally relevant waters. For site-specific DTA to be used to develop a guideline value for an effluent or a locally-relevant toxicant, the use of at least 5 locally relevant or regionally relevant species from at least 4 taxonomic groups is required. Initially, long-term DTA programs should use the largest suite of relevant species possible, preferably at least 8. The suite can subsequently be refined as an understanding develops of the relative sensitivity of the species to the water being tested. The design and refinement of a DTA program would be largely guided by the objectives of the program. Van Dam et al. (2019) provide additional guidance to that in ANZG (2018) on selecting and evaluating approaches for deriving site-specific guideline values, much of which is transferable to DTA.

When conducting DTA on a suite of approximately 8 species and, for various reasons, some species might be temporarily or permanently removed from the testing suite, it is important to understand the Burrlioz 2.0 statistical distribution-fitting software for deriving guideline values. Where data exist for fewer than 8 species, the software uses a 2-parameter log-logistic distribution to generate the SSD. For data for 8 or more species, it uses a 3-parameter Burr distribution (see Batley et al. 2018, Warne et al. 2018). This shift in the distribution between sample sizes of 7 and 8 can result in markedly different protective concentrations, especially at the tails of the distributions, from which the guideline values are estimated. This can be problematic from a discharge licencing perspective, and it is a difficult situation to explain to dischargers and regulators. Future improvements in the statistical approach for deriving guideline values should overcome this problem to some extent (e.g. see description of model averaging in Fox et al. 2021). Ultimately, DTA programs should form one of several lines of evidence for assessing water quality, and best professional judgement may be needed where results present challenges that existing methods and guidance do not cover.

Frequency of assessment

The frequency of testing an effluent, or ambient water, will depend on the objective of the testing and on the nature of the effluent or receiving water, although some broad guidance can be offered. If the discharge is known to be of relatively constant composition, and the receiving water characteristics are well documented and understood, one-off testing may be appropriate, although periodic validation would be recommended. However, discharge composition is more often variable and, if DTA is determined to be an appropriate method, testing may be required on a more frequent basis (e.g. monthly or quarterly). If a discharge varies according to the process undertaken but is constant within that process, or if the receiving water varies seasonally but is relatively constant within seasons, testing can be carried out whenever such a change is known to occur. It may be possible to refine and reduce the size of more frequent DTA programs (e.g. by reducing the number of species and focusing on the most sensitive or consistently responding species). This must be supported by appropriate evidence and interspersed with periodic full DTAs to confirm acceptable dilution rates.

Some jurisdictions require multiple DTAs of a discharge during an operation's commissioning and early operational phases (e.g. 3 DTAs within the first 1–2 years), with the aim of confirming long-term acceptable dilution rates (i.e. to protect a specified percentage of species). How the data are used to set the acceptable dilution rates may vary depending on the actual objectives and the DTA results. Where the DTA results are consistent, an averaging approach may be appropriate. Where the DTA results are variable, it may be necessary to adopt a conservative approach and use the most sensitive of the DTA results (i.e. that resulting in the highest acceptable dilution rate) to set the receiving water dilution rates. Also, periodic DTA-based validation of the dilution rates is recommended. These are suggestions only. Detailed guidance on how to use DTA results to make regulatory and management decisions is outside the scope of this report, although could potentially be developed in the future.

Detailed guidance on how to best assess the effects of pulse exposures is also outside the scope of this report, but a good overview is provided by Chèvre and Vallotton (2013). There are numerous other studies that may also be useful (e.g. Ashauer et al. 2006, Hogan et al. 2013, Sinclair et al. 2014, Angel et al. 2015, Angel et al. 2017, Hamer et al. 2019). Also, Smith et al. (2020) consider the assessment of pulse exposures in the context of temporary waters.

3.2.3 Toxicity testing in sediments

Over the past decade, there have been significant advances in sediment ecotoxicology. In particular, there are now both chronic and acute whole-sediment toxicity tests for amphipods, benthic algae, mysids, copepods, bivalves and polychaete worms in marine sediments, and for amphipods, nematodes, oligochaete worms, bivalves, chironomids, mayflies, snails and water fleas in freshwater sediments (Simpson and Batley 2016). Details for some of these have been provided by Simpson and Batley (2016) and have been used for sediment DTA and the derivation of sediment-quality guideline values. Refer to Simpson et al. (2013) and Simpson and Batley (2016) for detailed guidance on the use of toxicity testing for sediments.

3.2.4 Semi-field-based assessments of toxicity

Much of the preceding guidance on toxicity assessment has focused on laboratory-based approaches. It is also worth providing some guidance for field-based assessments of toxicity, which have their own sets of strengths and limitations. Only a summary is provided here. Refer to ANZG (2018) for more detailed information. Although the information in ANZG (2018) is focused on the use of field-based approaches for deriving guideline values, it is equally applicable for assessments that have objectives of assessing direct and indirect effects of physical or chemical (including toxicants) stressors.

Field-based assessments can be separated into 2 broad types: semi-field assessments and field assessments. Field assessments include experimental manipulations in whole ecosystems or observations from ecosystems across existing gradients of physical or chemical stressor levels or concentrations. These are part of the biodiversity line of evidence for which detailed guidance is provided in Section 3.1. Semi-field assessments can include *in-situ* toxicity testing, as discussed in Section 3.2.2 ('Standard versus site-specific direct toxicity assessment'), and microcosm and mesocosm studies. The remainder of this section focuses on mesocosm studies.

Mesocosms are artificial systems containing complex and self-sustaining populations or communities set in natural (or semi-natural) environmental conditions. For assessing the impacts of physical and

chemical stressors, mesocosms can represent a useful middle ground between laboratory studies and full field studies by combining advantages of both (Buchwalter et al. 2017). Unlike laboratory studies, mesocosm studies can account for complex local interactions among the aquatic community, stressors and local environmental conditions (depending on mesocosm size and complexity). In doing so, they can capture both direct and indirect effects across multiple taxonomic communities (van Dam et al. 2019). They can also assess effects to a greater number and broader range of taxa, including sensitive taxa that may be difficult to culture and assess under laboratory conditions (e.g. many aquatic insect species) (ANZG 2018). Mesocosms also offer the capacity for experimental replication and greater control of confounding factors not often possible in field studies in natural settings (Buchwalter et al. 2017, ANZG 2018, van Dam et al. 2019). Mesocosms also have their limitations, including potentially high variability, enclosure biases, expense and being logistically complex. This often results in only a few exposure concentrations being tested (Perceval et al. 2011, Buchwalter et al. 2017, van Dam et al. 2019). However, data from mesocosm studies are considered appropriate to use in a weight-of-evidence assessment, to represent or support other data from the toxicity line of evidence in guideline value derivation, and, more broadly, in water or sediment quality assessments (e.g. European Commission 2018, ANZG 2018).

For further details on the use of mesocosm studies and other forms of semi-field toxicity assessments, refer to Perceval et al. (2011), Buchwalter et al. (2017), ANZG (2018) and van Dam et al. (2019).

3.3 Bioaccumulation

The bioaccumulation line of evidence includes indicators that are measured concentrations of contaminants accumulated by aquatic biota from water or sediment. These are compared with concentrations in the same species from a reference site(s). The measurements indicate whether the contaminants are present in forms that are bioavailable, complementing measurements of chemical concentrations in water and sediment. Bioaccumulation is traditionally considered in relation to food standards and the consumption of toxicant-containing biota. While this might be a management concern, bioaccumulation can have an important function in relation to ecological assessments: it indicates exposure to bioavailable contaminants and also provides a clear link between ecological and human-health assessments of water-quality risk. Measurements can involve the collection of field organisms, the use of caged bioaccumulators (e.g. mussels, oysters), transplanted organisms or biomimetic devices.

Commonly used measures of bioaccumulation include the bioconcentration factor (BCF) and the bioaccumulation factor (BAF). For aquatic biota, BCF is the ratio of the concentration of contaminant in the biota sample (e.g. whole tissue, organ-specific) to the concentration of contaminant in the water column. BAF is the ratio of the concentration of contaminant in the biota sample to the concentration of contaminant in both the water and dietary sources (McGeer et al. 2003). So, high BCFs and BAFs equate to high internal concentrations relative to the exposure media (e.g. water, diet), which indicates that bioaccumulation is occurring.

Bioaccumulation indicators provide information on the behaviour of contaminants in aquatic environments, exposure of aquatic biota to bioavailable contaminants, sources of contaminants in the environment, and causes of toxicity to aquatic biota (van der Oost et al. 2003, Luoma and Rainbow 2005). However, bioaccumulation is not an effective indicator of biological effects, and field

measurements of bioaccumulation can be difficult to interpret due to multiple confounding factors (Luoma and Rainbow 2005, Arnot and Gobas 2006). Therefore, it is important that bioaccumulation indicators are used alongside stressor and other ecosystem receptor indicators in a weight-of-evidence approach. Even so, for some substances (e.g. selenium), bioaccumulation data can better assess water quality than by measuring these substances in water or sediment alone (e.g. Hamilton 2002, Brix et al. 2005, Toll et al. 2005, US EPA 2016).

Care must be taken when using bioaccumulation, as the extent to which a substance accumulates depends on its inherent propensity to partition into biological tissues (particularly for xenobiotic substances) and on the ability of the organism to regulate uptake and excretion of the substance. Some groups of organisms detoxify toxins by sequestering them into biologically inert forms. Other groups detoxify by breaking the toxin into component parts and excreting it. Organisms that detoxify by excretion tend to bioaccumulate the toxin to a lesser degree than organisms that detoxify by sequestration. For some substances, different organs within an organism may fall into either of these groups. As a general rule, organisms that detoxify by sequestration are better monitors of substance bioavailability than the organisms that detoxify by excretion. Relevant reviews should be consulted before selecting what organisms or parts of organisms are selected for bioaccumulation monitoring (e.g. van der Oost et al. 2003, DeForest et al. 2007).

The ANZG (2018) list of DGVs for toxicants provides cautionary notes on the use of the DGVs for at least 90 substances that are known to bioaccumulate in some species or are expected to have a high potential for bioaccumulation due to having log- K_{ow} (octanol-water partition coefficient) values greater than 4. The [ANZG \(2018\) toxicants search tool](#) can be used to obtain details on whether a substance meets this criterion. If one or more of those substances are potential stressors of concern, or if a substance of potential concern does not have a DGV but is known to have a log- K_{ow} value greater than 4 (which would need to be determined from a literature review), the use of bioaccumulation as part of a weight-of-evidence assessment should be considered. ANZG (2018) also provides guidance on [how to apply DGVs for toxicants that bioaccumulate](#).

The following example is provided to demonstrate how bioaccumulation data can be a valuable component of a weight-of-evidence assessment when used with insight from the conceptual model and with reference to the literature on the ecotoxicology and organism regulation of that substance.

Copper is a very well-studied essential metal. Because it is required for a number of enzymatic processes but is toxic in excess, organisms have developed strategies to regulate their internal concentrations of copper (e.g. White and Rainbow 1985, Phillips and Rainbow 1989). Copper, as part of haemocyanin, plays a particularly important role in oxygen transport for crustaceans, in addition to its role in other enzyme systems common to other organisms. Even so, very different strategies for regulation of copper burdens have developed within groups of crustaceans. For example, in decapods, copper tends to be well regulated in general body tissues by sequestration or binding to metallothionines in the hepatopancreas and excretion via the antennal gland. However, in barnacles, copper is sequestered through binding into organic granules within the general body tissues, to the extent that excretion typically only occurs via the release of gametes (Rainbow and White 1989, Rainbow 2002). Therefore, in practice, biomonitoring of copper concentrations in populations of barnacles often provides information on lifetime exposure to bioavailable copper (for annually spawning species, effectively

integrating exposure over the course of a year). In contrast, in populations of decapod crustaceans, it provides evidence of shorter-term exposure but only via measurements in the hepatopancreas and antennal gland and not, for example, in abdominal flesh, which tends to have highly regulated copper concentrations (that are very sensitive to haemolymph loss during sample preparation).

Consideration must be given to the selection of organisms (and parts of organisms) to be measured for bioaccumulation of copper and to the duration of the exposure of the organisms to elevated copper concentrations. It is well established that the acute toxicity of copper is not caused by bioaccumulation of copper but is due to interference in sodium regulation and the associated loss of plasma sodium, at least for freshwater species (Grosell et al. 2002, Niyogi and Wood 2004, Grosell et al. 2007, Alsop and Wood 2011). Therefore, tissue copper concentrations associated with acute exposures can be poorly correlated with exposure to copper and its toxicity. In contrast, tissue copper concentrations associated with chronic exposures can be more reliably related to copper toxicity. However, in fishes, copper will typically remain well regulated in body parts such as flesh and plasma and can be more meaningfully monitored in the organs involved in uptake and regulation, such as the gills and liver.

Bioaccumulation can be a very useful part of a weight-of-evidence assessment, but care must be taken to consider the likely nature of the exposure (from the conceptual model), as well as the groups of organisms and the parts of those organisms to use – even for a well-understood toxicant like copper. For this line of evidence, the literature must be used to design programs and select indicators.

Finally, a cautionary note is provided on the use of bioaccumulation for assessing metals in aquatic environments. It has been well documented that BCFs and BAFs exhibit an inverse relationship for metals. They are highest (indicating hazard) at low exposure concentrations and are lowest (indicating no hazard) at high exposure concentrations, where impacts are likely (McGeer et al. 2003, DeForest et al. 2007, Fairbrother et al. 2007). So, using these metrics for metals can lead to conclusions that are inconsistent with the toxicological data (McGeer et al. 2003). Consequently, the metals risk-assessment framework of the United States Environmental Protection Agency states that the science does not support the use of a single, generic threshold BCF or BAF value as an indicator of metal hazard (Fairbrother et al. 2007). Where BCF and BAF values are being calculated, it is critical that they are interpreted in the context of the exposure concentration and what that means for potential impacts (DeForest et al. 2007).

3.4 Biomarkers

A biomarker is a measurable indicator of a sub-organismal biological response or condition (e.g. enzymatic or histopathological) as a result of exposure to a stressor. The biomarkers line of evidence includes sub-organismal indicators of either exposure to or effects of contaminants in water and sediment. Biomarkers have been equally applied worldwide in both freshwater and marine ecosystems. Molecular biomarkers are characteristic signatures of pollution expressed in enzymes, cell constituents or metabolism products within organs of animals and plants. Organisms respond to stress by invoking molecular responses, and these can be expressed as physiological or other changes. The molecular responses to contaminant stress are likely to be the earliest form of organism response. These could be used as early-warning indicators of changes induced by

contaminants and in a weight-of-evidence assessment as indicators of organism exposure to that contaminant(s). Of course, changes at the molecular level in an organism may not necessarily reduce its ecological fitness (its ability to function normally) because of various compensatory factors. Even so, molecular responses to various contaminants have been sought in a broad range of species, and many site-specific and host-specific responses have been detected for biomarkers based on the activity of specific enzymes in the liver, kidney, blood and other tissues.

Attributes of ideal biomarkers include the following (van Dam et al. 1998, Walker 1999).

- sensitive: responsive to exposure to a contaminant or contaminants, providing early-warning capability for the potential onset of effects at higher levels of biological organisation
- correlated to toxicity at the whole organism level (or higher): knowledge that the manifestation of a biomarker response leads to effects on whole-organism processes such as growth, development, reproduction or survival
- specific: provides a diagnostic capability for identifying the cause of an organism's response to contamination
- stable: useful, particularly in field studies
- simple: has a simple, rapid and cost-effective method
- non-destructive: reduces the number of animals used and also allowing for serial sampling.

No single biomarker is likely to possess all these attributes (van Dam et al. 1998, Walker 1999). A biomarker might provide an excellent indication of exposure to a particular contaminant but might not be correlated to effects at higher levels of biological organisation (e.g. individual, population, community, ecosystem). A biomarker may be very specific to a particular substance or group of chemicals (e.g. Matozzo et al. 2008) and may not be applicable to other contaminants. Similarly, a long-term monitoring program might provide excellent baseline data from which small perturbations will be obvious, but such a program may be neither time-efficient nor cost-efficient. Therefore, appropriate and achievable attributes for a particular purpose must be selected so that biomarker indicators can be chosen based on these.

The most intensively studied group of organisms in relation to biomarkers is fish, both freshwater and marine. Numerous studies have examined the usefulness of sub-cellular biomarkers to respond effectively to and generally indicate the effects of a number of pollutants. The most promising use of biomarkers is as screening tools for detecting pollution via unusual patterns in a suite of biomarkers (Gunther et al. 1997, van der Oost et al. 2003, Viarengo et al. 2007, Hook et al. 2014, Kroon et al. 2017). Several reviews have highlighted the potential usefulness of such approaches in environmental risk assessment (e.g. Behnisch et al. 2001, van der Oost et al. 2003).

As part of a weight-of-evidence assessment, biomarkers can play an important and cost-effective role in diagnostics and identifying contaminant sources. This has been recognised for the assessment of sediment (Martín-Díaz et al. 2008). Hook et al. (2014) provide a review of the general application of biomarkers, including application of exposure biomarkers and effect biomarkers, and provide a list of commonly used biomarkers according to different toxicant classes.

4 Summary

The ecosystem receptor line of evidence is a critical part of the weight-of-evidence process for assessing water and sediment quality. Ecosystem receptors represent the components (species, populations, communities, habitats) of the ecosystem that may be exposed to and impacted by pressures and their associated stressors. Consequently, in addition to assessing water chemistry against guideline values for water and sediment quality, assessing and understanding the responses of ecosystem receptors to stressors is critical to quantifying impacts to water and sediment quality. Additionally, ecosystem receptors are typically more acceptable to the general public than water chemistry. Their inclusion in monitoring programs provides greater confidence in concluding whether ecosystems are being impacted or protected.

Ecosystem receptor indicators must be appropriate, and they must be selected based on numerous factors that are typically considered in Step 1 and Step 2 of the WQMF and the associated monitoring requirements. Additional information on the selection of indicators for different lines of evidence across the pressure, stressor and ecosystem receptor causal pathway elements is provided in ANZG (2018).

Appendix A: Additional guidance for matching biodiversity indicators to problems

Table A1 Matching biodiversity indicators to a problem

Problem	Potentially suitable biodiversity indicators and measured responses	Advantages	Disadvantages
Nutrients, herbicides: early detection, changes to biodiversity.	Periphytic (benthic or epiphytic) algae: <ul style="list-style-type: none"> • biomass (chlorophyll <i>a</i>) • community structure (e.g. diatoms) • presence or absence, abundance categories (macroalgae). 	<ul style="list-style-type: none"> • There is a direct and rapid response to nutrients. • Techniques are well developed and generally applicable to most inland waters. • Diatoms preserved in sediments may be useful for paleolimnological monitoring. 	<ul style="list-style-type: none"> • Periphytic algae may not be useful in deep, highly turbid waters or in coloured (high gilvin content) wetlands. • Spatial and temporal variations in community structure may be very high (low power to detect effects).
	Phytoplankton: <ul style="list-style-type: none"> • biomass (chlorophyll <i>a</i> as surrogate) • frequency of algal blooms • community structure. 	<ul style="list-style-type: none"> • There is a direct and rapid response to nutrients. • Methods are simple. • Techniques are well developed and documented. • Some freshwater taxa are sensitive indicators of trophic state. • Algal blooms may be localised, enabling identification of source of effects. 	<ul style="list-style-type: none"> • Phytoplankton is not an applicable indicator for swift-flowing upland streams. • There is poor background knowledge of phytoplankton ecology. • Phytoplankton may not be useful in deep, highly turbid waters or in coloured (high gilvin content) wetlands. • Algal blooms are difficult to quantify and may not always be obvious at the surface. • Algal blooms are not always directly linked to nutrients, as blooms are often associated with dissolved nutrient forms stripped from the water column (i.e. at or near detection limit) and an interplay amongst nutrients, water flow (in streams), light and temperature (and hence season). • Spatial and temporal variations in community structure may be very high. • Mobility of algal blooms under certain conditions (e.g. due to currents especially in marine environments) will hamper identification of source of effects.
	Macrophytes: <ul style="list-style-type: none"> • emergent or submersed vegetation • seagrass depth limits. 	<ul style="list-style-type: none"> • Macrophytes can be useful indicators of water and sediment quality, depending upon the taxon. • Ground-survey techniques are easily applied, and GIS (geographic information systems) approaches are established. • Many species are habitat-forming (changes may cascade through system). • Drone-based methods are being developed. 	<ul style="list-style-type: none"> • There is a lack of knowledge about population dynamics and about how factors other than water quality affect distribution and growth. Hydrology can be a major determinant of distribution and growth. • Spatial and temporal variations in community structure may be very high. • Macrophytes are less useful in systems with naturally low numbers of taxa (e.g. much of south-west Western Australia and temporary systems).

Problem	Potentially suitable biodiversity indicators and measured responses	Advantages	Disadvantages
Nutrients: early detection, changes to biodiversity.	Stream metabolism: <ul style="list-style-type: none"> gross primary productivity (GPP), respiration (R), gross primary production:respiration ratio (P:R) and net daily metabolism. 	<ul style="list-style-type: none"> Stream metabolism may provide inexpensive, advanced warning of nutrient enrichment in some nutrient-poor forest streams. 	<ul style="list-style-type: none"> Unless correlated with structure changes of aquatic ecosystem components, stream metabolism may lack information about ecological relevance and importance. It can be difficult to obtain consistent results. Closed systems may not reliably simulate <i>in-situ</i> substrata and metabolism. Open-channel methods can be unreliable or inconsistent in turbulent (fast-flowing) water.
Nutrients: changes to biodiversity.	Macroinvertebrates: <ul style="list-style-type: none"> community structure (quantitative assessment, rapid biological assessment [RBA]). 	<ul style="list-style-type: none"> Macroinvertebrates are found in most habitats. A large number of taxa offers a wide range of responses (diagnostic value). The limited mobility of the aquatic larval stages allows effective spatial analyses of disturbance effects. Larval stages often extend up to and beyond one month, integrating effects of prolonged or intermittent exposure. Identification is relatively easy at the family level. Methods are well established and ecological knowledge base is improving. RBA methods are faster, allowing for greater geographical coverage, while quantitative methods are slower but more sensitive which can benefit site-specific assessments. . 	<ul style="list-style-type: none"> If direct responses to nutrient enrichment are sought, then aquatic flora would be a preferred indicator and may be less expensive to sample. There can be difficulties with diagnosis, as nutrients elicit both a positive and negative response in macroinvertebrate species at different levels (increasing and decreasing with phytoplankton levels and associated changes in water quality). Identification is difficult at the species level, which can be important given the variability in species responses within families (particularly in systems with low number of families). RBA methods incur some loss of information, while quantitative methods are more expensive. Adult stages may be very mobile and air breathing, masking passage-barrier impacts, and may be very insensitive to toxicants if not identified separately from larval phases.
General organic and inorganic contaminants (including metals, pesticides): early detection of changes (water column and sediments).	Generic biodiversity-based indices (excluding similarity indices)	<ul style="list-style-type: none"> Responses for a number of indicators have been linked to particular stressors and systems and so have diagnostic value. Where the response has been shown to be highly specific for particular contaminants, biodiversity-based indices have high sensitivity, allowing timely detection of the effects of particular substances at specific sites (i.e. prior to or indicating the onset of environmental impacts). 	<ul style="list-style-type: none"> The outputs for indices that lack contaminant specificity or reference to the identity of taxa can be difficult to interpret. Biodiversity-based indices may lack ecological relevance: very few responses have been linked to effects at higher levels of biological organisation (e.g. ecosystems). The sensitivity of the selected test species in single-species tests or studies may be unrepresentative of the wider assemblage of organisms in the field. Timeliness may be compromised by the need for adequate baseline data and post-disturbance data from the field, and (usually) dose-response or exposure-response relationships from laboratory studies, to interpret results and strengthen inferences. Site-specific assessments may be difficult with mobile species depending on site connectivity, species habits and life stage present (e.g. fish, frogs). The biology of the organism being measured must be understood.
General organic and inorganic contaminants (including metals, pesticides): early detection of changes (water column).	Physiological or biochemical (sub-organismal) changes in organisms	In addition to the advantages listed under 'General organic and inorganic contaminants (including metals, pesticides): early detection of changes (water column and sediments)': <ul style="list-style-type: none"> Physiological or biochemical changes in organisms can indicate the bioavailability of a contaminant. 	In addition to the disadvantages listed under 'General organic and inorganic contaminants (including metals, pesticides): early detection of changes (water column and sediments)': <ul style="list-style-type: none"> This indicator may not be specific to particular contaminants.
	Whole-body responses of organisms (field and laboratory toxicological assessments of ambient water or effluent): includes lethality, growth, reproduction	In addition to the advantages listed under 'General organic and inorganic contaminants (including metals, pesticides): early detection of changes (water column and sediments)': <ul style="list-style-type: none"> Endpoints are regarded as being of greater ecological relevance than many other 'early-detection' methods. 	In addition to the disadvantages listed under 'General organic and inorganic contaminants (including metals, pesticides): early detection of changes (water column and sediments)': <ul style="list-style-type: none"> Field application of this indicator can be expensive (non-<i>in-situ</i> methods).
	Whole-body responses of organisms (field surveys): abnormalities	In addition to the advantages listed under 'General ... organic and inorganic contaminants (including metals, pesticides): early detection of changes (water column and sediments)':	In addition to the disadvantages listed under 'General organic and inorganic contaminants (including metals, pesticides): early detection of changes (water column and sediments)':

Problem	Potentially suitable biodiversity indicators and measured responses	Advantages	Disadvantages
		<p>Frogs</p> <ul style="list-style-type: none"> Skeletal abnormalities are purported to indicate exposure to low concentrations of contaminants, including metals and radionuclides. See also: advantages listed for frogs in 'General organic and inorganic contaminants (including metals, pesticides): changes to biodiversity'. <p>Birds (eggshell thinning)</p> <ul style="list-style-type: none"> Responses linked to specific contaminants. <p>Imposex in gastropods</p> <ul style="list-style-type: none"> There is a direct relationship between exposure and organotins. Gastropods have limited mobility, allowing for effective spatial analyses of disturbance effects. Robust models and concepts are available for this indicator. 	<p>Frogs</p> <ul style="list-style-type: none"> See also: disadvantages listed for frogs in 'General organic and inorganic contaminants (including metals, pesticides): changes to biodiversity'. <p>Birds (eggshell thinning)</p> <ul style="list-style-type: none"> Effects are sub-lethal but may show a slow or lagged response. Birds are highly mobile, so site-specific assessments may not be possible. High specificity to contaminant classes limits applicability. <p>Imposex in gastropods</p> <ul style="list-style-type: none"> The observed effects require laboratory verification of response thresholds for the species concerned.
	Sentinel organisms (bioaccumulation, body burdens): includes macrophytes, long-lived invertebrates (molluscs, crustaceans), fish	<ul style="list-style-type: none"> The species selected can absorb or adsorb high concentrations of contaminants without direct toxicity. Sentinel organisms provide an indication of the bioavailability of contaminants. Methods for sentinel organisms are well established. 	<ul style="list-style-type: none"> Some waters may lack an indicator organism of sufficiently large size or age. For correct interpretation, knowledge on whether or not the contaminant bioaccumulates and the retention time in the organism must be determined.
General organic and inorganic contaminants (including metals, pesticides): early detection of changes (sediments).	'Whole-sediment' laboratory toxicity assessment	<p>In addition to the advantages listed under 'General organic and inorganic contaminants (including metals, pesticides): early detection of changes (water column and sediments)':</p> <ul style="list-style-type: none"> The endpoints of lethality, growth, reproduction have ecological importance. 	<p>In addition to the disadvantages listed under 'General organic and inorganic contaminants (including metals, pesticides): early detection of changes (water column and sediments)':</p> <ul style="list-style-type: none"> There are a limited number of sediment tests available in Australia (see Simpson & Batley 2016). Laboratory results might not accurately reflect effects that can occur at the ecosystem level (i.e. test conditions may not simulate actual environmental conditions).
	Morphological deformities of sediment-dwelling organisms	<p>In addition to the advantages listed under 'General organic and inorganic contaminants (including metals, pesticides): early detection of changes (water column and sediments)':</p> <ul style="list-style-type: none"> Methods for specific groups are well established (e.g. chironomids). 	<p>In addition to the disadvantages listed under 'General organic and inorganic contaminants (including metals, pesticides): early detection of changes (water column and sediments)':</p> <ul style="list-style-type: none"> The degree of applicability to specific stressors needs to be ascertained.
General organic and inorganic contaminants (including metals, pesticides): changes to biodiversity.	Algae	<ul style="list-style-type: none"> Other than useful indicators of nutrient enrichment, attached or planktonic algae may be sensitive to salinity, dissolved oxygen, pH, suspended solids, a number of metals, and general contaminant classes including pesticides. Databases on physiological optima or tolerance to water-quality constituents of algal species are available. Algae form the base of many food chains, respond quickly to water quality stressors, are relatively cheap and easy to measure, while attached forms allow effective spatial analyses of disturbance effects. 	<ul style="list-style-type: none"> Planktonic forms may be too 'dilute' and transported readily in upland portions of streams. Spatial and temporal variations in community structure may be very high. Field studies require measurement of nutrients and other indicators of organic enrichment if the effects of other contaminant classes are the subject of investigation (i.e. to eliminate or quantify the contribution of these other stressors as causes of effects) .
	Macrophytes	<ul style="list-style-type: none"> Many species are habitat-forming (changes may cascade through system). Ground-survey techniques are easily applied, and GIS approaches are established (emergents). Useful as indicators of herbicides, suspended solids, acidification and as bioaccumulators. 	<ul style="list-style-type: none"> There is a lack of knowledge about population dynamics and how factors other than water quality affect distribution and growth. Few macrophytes are directly sensitive to metals, so they are not likely to be good indicators of metal contamination.
	Zooplankton and other microfauna	<ul style="list-style-type: none"> Microfauna are sensitive to wide range of contaminants with well-established response thresholds from laboratory testing available for some species. 	<ul style="list-style-type: none"> Spatial and temporal variations in community structure may be very high. Microfauna may be too 'dilute' and transported readily in upland portions of streams.

Problem	Potentially suitable biodiversity indicators and measured responses	Advantages	Disadvantages
General organic and inorganic contaminants (including metals, pesticides): changes to biodiversity.	Macroinvertebrates: <ul style="list-style-type: none"> community structure. 	<ul style="list-style-type: none"> See advantages listed for macroinvertebrates in 'Nutrients: changes to biodiversity'. 	<ul style="list-style-type: none"> Sample processing and identification of samples is labour intensive. Adult stages may be very mobile and air breathing, masking passage-barrier impacts, and may be very insensitive to toxicants if not identified separately from larval phases.
	Fish	<ul style="list-style-type: none"> Fish have a high public profile. Fish are sensitive to a wide range of contaminants. Fish are long-lived and so exposures can be long-term. Taxonomy is usually simple; sample processing costs generally 'small'. Sampling methods are well established. Where physical or contaminant passage barriers to the movement of fish are important, identifying and monitoring of these potential migration impediments can be an effective assessment technique. Some species and life stages are quite sedentary, so if these species and life stages can be identified, effective spatial analyses of disturbance effects may be possible. 	<ul style="list-style-type: none"> Assemblage-based approaches are compromised in most of southern and inland Australia where species diversity is low, fluctuations in species abundance and occurrence are extreme, and relative dominance of exotic species is high. The diadromous nature of most of native fish in New Zealand precludes the use of natural freshwater fish communities for bioassessment of water quality except where impediments to recruitment are of concern. Fish are highly mobile and so the interpretation of data must take into account factors affecting the entire catchment or region, including connectivity between sites.
	Frogs	<ul style="list-style-type: none"> There is a high public profile and level of concern about frogs. Frogs are sensitive to a wide range of contaminants (but generally less than microflora and microfauna). Skeletal abnormalities are purported to indicate exposure to low concentrations of contaminants. 	<ul style="list-style-type: none"> There are no standard techniques developed in Australia and New Zealand. Most species have a semi-aquatic life cycle. The fully aquatic (larval) phase is often highly seasonal and transient. Adults are highly mobile. There is high inter-annual variability of breeding adults, spawn and recruitment. Little is known about the response of frogs to changing environmental conditions.
	Waterbirds	<ul style="list-style-type: none"> Good baseline information exists (e.g. Birdlife Australia). Where physical or contaminant passage barriers to the movement of birds are important, identifying and monitoring of these potential migration impediments can be an effective assessment technique. Some species are quite sedentary, for which effective spatial analyses of disturbance effects may be possible. 	<ul style="list-style-type: none"> Waterbirds are only indirectly affected by most contaminants through either effects on prey or through the food chain. The highly mobile and migratory nature of many species means that the interpretation of data must take into account factors affecting the entire catchment or region, including connectivity between sites.
Rapid biological assessment (RBA) for early detection.	Macroinvertebrates	<ul style="list-style-type: none"> Macroinvertebrates can be measured at relatively low cost at a large number of sites or over large areas. In their broad coverage, macroinvertebrates may identify and detect problem locations and stressors that would otherwise pass unnoticed. Macroinvertebrates have ecological, regional and social relevance. Ecogenomic approaches are being actively developed. 	<ul style="list-style-type: none"> There is low power to detect changes. For particular sites, macroinvertebrates are not sufficiently sensitive to detect subtle impacts at an early stage. Macroinvertebrates assessed using these methods have generally poor sensitivity to point-source impacts of contaminants.
Rapid biological assessment (RBA) for biodiversity, broad-scale ecosystem health.	Macroinvertebrates	<ul style="list-style-type: none"> The same advantages as listed for macroinvertebrates in 'Nutrients: changes to biodiversity' apply here. The method is widely regarded as adequately reflecting ecological condition or integrity at the catchment or region scale. Approaches to sampling and data analysis are highly standardised. The response is measured rapidly, inexpensively and with rapid turnaround of results. Results are readily understood by non-specialists. The response has some diagnostic value. Ecogenomic approaches are being developed. 	<ul style="list-style-type: none"> The method is not designed to detect minor or subtle impacts (negating sole use in regions of higher conservation value). There are a number of limitations for site-specific assessments reported under 'Predictive modelling'.

Problem	Potentially suitable biodiversity indicators and measured responses	Advantages	Disadvantages
	<p>Fish:</p> <ul style="list-style-type: none"> presence or absence, indices. 	<p>In addition to the advantages listed for fish in 'General organic and inorganic contaminants (including metals, pesticides): changes to biodiversity':</p> <ul style="list-style-type: none"> These indicators can be cheaply applied over large spatial scales. 	<p>In addition to the disadvantages listed for fish in 'General organic and inorganic contaminants (including metals, pesticides): changes to biodiversity':</p> <ul style="list-style-type: none"> There is a lack of understanding of population dynamics and ecology for some species in Australian systems. These indicators are not appropriate in New Zealand.
	<p>Frogs:</p> <ul style="list-style-type: none"> presence or absence. 	<p>In addition to the advantages listed for frogs in 'General organic and inorganic contaminants (including metals, pesticides): changes to biodiversity':</p> <ul style="list-style-type: none"> Frog calls are a simple approach. Some background data are available (e.g. FrogWatch). The availability of larval identification keys means that using the larval stage is a viable option. 	<p>In addition to the disadvantages listed for frogs in 'General organic and inorganic contaminants (including metals, pesticides): changes to biodiversity':</p> <ul style="list-style-type: none"> Frog calls are highly dependent upon seasons and environmental conditions.
	<p>Birds:</p> <ul style="list-style-type: none"> presence or absence. 	<ul style="list-style-type: none"> Good baseline information exists (e.g. Birdlife Australia). 	<ul style="list-style-type: none"> Birds are only indirectly affected by most contaminants through either effects on prey or through the food chain. The highly mobile and migratory nature of many species means that the interpretation of data must take into account factors affecting the entire catchment or region, including connectivity between sites.
Suspended solids, sedimentation, dredging.	<p>Macrophytes:</p> <ul style="list-style-type: none"> emergent or submersed vegetation seagrass depth limits. 	<ul style="list-style-type: none"> See advantages listed in 'Nutrients: early detection, changes to biodiversity'. There is a direct response of submerged macrophytes to shading caused by suspended sediments 	<ul style="list-style-type: none"> See disadvantages listed in 'Nutrients: early detection, changes to biodiversity' Vegetation can be patchy and variable in extent, both spatially and seasonally. Use of these indicators may require significant sample numbers, controls and replication to increase power for change detection.
	<p>Seagrass:</p> <ul style="list-style-type: none"> seagrass health and mortality. 	<ul style="list-style-type: none"> There is a direct response of seagrasses to shading caused by suspended sediments. Seagrasses are a widely accepted indicator of light stress. Water-quality triggers are established as indicators of pressure. Established biological responses are linked to water quality. 	<ul style="list-style-type: none"> Triggers are not available for all species. Some species' adaptation to highly turbid environments may require the development of site-specific triggers. Some methods, such as leaf counts, can be resource intensive and require diving (increasing safety risks). Some species have naturally variable cover on a seasonal and inter-annual basis.
	<p>Phototrophic sponges:</p> <ul style="list-style-type: none"> sponge health and mortality. 	<ul style="list-style-type: none"> There is a direct response of phototrophic sponges to shading caused by suspended sediments and sedimentation. Water-quality triggers are established as indicators of pressure. Established biological responses are linked to water quality. 	<ul style="list-style-type: none"> The variability of response between different phototrophic species is not well known. There is evidence that some phototrophic sponges adapt rapidly to increased turbidity.
	<p>Coral communities:</p> <ul style="list-style-type: none"> sponge health and mortality. 	<ul style="list-style-type: none"> There is a direct response to shading and other effects (e.g. fouling, clogging) caused by suspended sediments and sedimentation. Water-quality triggers are established as indicators of pressure. Established biological responses are linked to water quality. 	<ul style="list-style-type: none"> Triggers may not be appropriate for very shallow or intertidal areas. Bleaching events (climate-driven) can negate the usefulness of ongoing water-quality monitoring programs. Spatial variability can mean large sample sizes are required or that specific corals are marked and monitored. Both approaches can be resource intensive.
	<p>Macroinvertebrates (especially benthic and infauna):</p> <ul style="list-style-type: none"> community structure. 	<ul style="list-style-type: none"> See advantages listed in 'Nutrients: early detection, changes to biodiversity'. 	<ul style="list-style-type: none"> See disadvantages listed in 'Nutrients: early detection, changes to biodiversity'.
	<p>Stream metabolism:</p> <ul style="list-style-type: none"> gross primary productivity (GPP), respiration (R), gross primary productivity:respiration ratio (P:R) and net daily metabolism (NDM). 	<ul style="list-style-type: none"> In some nutrient-poor forest streams, depressed GPP may indicate sedimentation or degradation in water quality. 	<ul style="list-style-type: none"> Unless correlated with changes in structure of aquatic ecosystem components, these indicators may lack information about ecological relevance and importance.

Glossary and acronyms

Term	Definition
ambient water	All water generally of natural occurrence (e.g. lakes, rivers, wetlands, estuaries, oceans).
AUSRIVAS	Australian River Assessment System.
BAF	Bioaccumulation factor.
BCF	Bioconcentration factor.
bioaccumulation	The process by which chemical substances are accumulated by aquatic organisms by all routes of exposures (dietary and the ambient environment).
biodiversity (biological diversity)	The variety of life forms, including the plants, animals and microorganisms, the genes they contain, and the ecosystems and ecological processes of which they are a part.
biomarker	A measurable indicator of a sub-organismal biological response or condition (e.g. enzymatic, histopathological) as a result of exposure to a stressor.
default guideline value (DGV)	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 <i>Australian and New Zealand guidelines for fresh and marine water quality</i> . Formerly known as 'trigger values'.
direct toxicity assessment (DTA)	The use of toxicity tests to determine the acute or chronic toxicity of wastewater discharges or total pollutant loads in receiving waters. Assesses the toxicity of mixtures of chemicals rather than individual chemicals
ecosystem condition	The current or desired state of health of an ecosystem, relative to the degree of human disturbance.
ecosystem receptor	Any living organism or natural habitat that could be exposed to a stressor.
eDNA	Environmental DNA.
GPP	Gross primary productivity.
guideline value	A measurable quantity (e.g. concentration) or condition of an indicator for a specific community 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 community value. Guideline values for more than one indicator should be used simultaneously in a multiple lines-of-evidence approach
indicator	A parameter that can be used to provide a measure of a pressure, stressor or ecosystem-condition response.
infauna	Benthic organisms that live within the bottom substratum of a waterbody.
low-effect concentration	Concentration generating an x% effect response, where $x \leq 10\%$.
MBACI	Multiple before–after, control–impact.
meiofauna	Small benthic invertebrates that live in both marine and freshwater environments.
multiple lines of evidence	Two or more lines of evidence that can be combined to monitor, assess or manage water or sediment quality.
NGS	Next-generation sequencing.
no-effect concentration	Maximum concentration of a toxicant that causes no adverse effect in a target organism.
no-observed-effect concentration	Highest concentration of a toxicant used in a toxicity test that does not have a statistically significant effect, compared to the controls.
P:R	Gross primary productivity:respiration ratio.

Term	Definition
periphytic	Living on the surfaces of rooted aquatic plants.
pressure	Any human activity or biophysical change that has the potential to have an impact on the natural environment.
QA/QC	Quality assurance and quality control.
R	Respiration.
RBA	Rapid biological assessment.
receiving water	An ambient water into which contaminants are being discharged.
reference condition	An environmental quality or condition that is defined from as many similar systems as possible and used as a benchmark for determining the environmental quality or condition to be achieved or maintained in a particular system of equivalent type.
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.
SSD	Species sensitivity distribution.
stressor	Any physical, chemical or biological substance or process arising from a pressure that has the potential to induce an adverse environmental response to a community value.
toxicity	The inherent potential or capacity of a material to cause adverse effects in a living organism.
toxicity identification evaluation (TIE)	Toxicity characterisation procedures involving use of selective chemical manipulations or separations and analyses, coupled with toxicity testing to identify specific classes of chemicals and ultimately individual chemicals that are responsible for the toxicity observed in a particular sample.
weight of evidence	A qualitative, semi-quantitative or quantitative combination of multiple lines of evidence to make an overall assessment of water or sediment quality or associated management. This assessment incorporates judgments about the quality, quantity, relevance and congruence of the data contained in the different lines of evidence.
WET	Whole effluent toxicity.
WQMF	Water quality management framework.

References

- Alsop D and Wood CM (2011) 'Metal uptake and acute toxicity in zebrafish: common mechanisms across multiple metals', *Aquatic Toxicology*, 105:385–393, doi:10.1016/j.aquatox.2011.07.010.
- Anderson MJ (2001) 'Permutation tests for univariate or multivariate analysis of variance and regression', *Canadian Journal of Fisheries and Aquatic Sciences*, 58:626–639, doi:10.1139/f01-004.
- Anderson MJ and Robinson J (2003) 'Generalized discriminant analysis based on distances', *Australian and New Zealand Journal of Statistics*, 45:301–318, doi:10.1111/1467-842X.00285.
- Anderson MJ and Ter Braak CJF (2003) 'Permutation tests for multi-factorial analysis of variance', *Journal of Statistical Computation and Simulation*, 73:85–113, doi:10.1080/00949650215733.
- Anderson MJ, Ellingsen KE and McArdle BH (2006) 'Multivariate dispersion as a measure of beta diversity', *Ecology Letters*, 9:683–693, doi:10.1111/j.1461-0248.2006.00926.x.
- Angel BM, Simpson SL, Chariton AA, Stauber JL and Jolley DF (2015) 'Time-averaged copper concentrations from continuous exposures predicts pulsed exposure toxicity to the marine diatom, *Phaeodactylum tricornutum*: importance of uptake and elimination', *Aquatic Toxicology*, 164:1–9, doi:10.1016/j.aquatox.2015.04.008.
- Angel BM, Simpson SL, Granger E, Goodwyn K and Jolley DF (2017) 'Time-averaged concentrations are effective for predicting chronic toxicity of varying copper pulse exposures for 2 freshwater green algae species', *Environmental Pollution*, 230:787–797, doi:10.1016/j.envpol.2017.07.013.
- Ankley GT, Bennett RS, Erickson RJ, Hoff DJ, Hornung MW, Johnson RD, Mount DR, Nichols JW, Russom CL, Schmieder PK, Serrano JA, Tietge JE and Villeneuve DL (2010) 'Adverse outcome pathways: a conceptual framework to support ecotoxicology research and risk assessment', *Environmental Toxicology and Chemistry*, 29:730–741, doi:10.1002/etc.34.
- ANZECC and ARMCANZ (Australian and New Zealand Environment and Conservation Council and Agriculture and Resource Management Council of Australia and New Zealand) (2000) *Australian water quality guidelines for fresh and marine waters*, ANZECC and ARMCANZ.
- ANZG (Australian and New Zealand Guidelines) (2018) *Australian and New Zealand guidelines for fresh and marine water quality*, Australian and New Zealand governments and Australian state and territory governments.
- Arce E, Archambault V, Mondy CP and Usseglio-Polatera P (2014) 'Recovery dynamics in invertebrate communities following water-quality improvement: taxonomy- vs trait-based assessment', *Freshwater Science*, 33:1060–1073, doi:10.1086/678673.
- Arnot JA and Gobas FAPC (2006) 'A review of bioconcentration factor (BCF) and bioaccumulation factor (BAF) assessments for organic chemicals in aquatic organisms', *Environmental Reviews*, 14:257–297, doi:10.1139/a06-005.

Ashauer R, Boxall A and Brown C (2006) '[Predicting effects on aquatic organisms from fluctuating or pulsed exposure to pesticides](#)', *Environmental Toxicology and Chemistry*, 25:1899–1912, doi:10.1897/05-393r.1.

Aylagas E, Borja A, Tangherlini M, Dell'Anno A, Corinaldesi C, Michell CT, Irigoien X, Danovaro R and Rodríguez-Ezpeleta N (2017) '[A bacterial community-based index to assess the ecological status of estuarine and coastal environments](#)', *Marine Pollution Bulletin*, 114:679–688, doi:10.1016/j.marpolbul.2016.10.050.

Bailey HC, Krassoi R, Elphick JR, Mulhall AM, Hunt P, Tedmanson L and Lovell A (2000a) '[Whole effluent toxicity of sewage treatment plants in the Hawkesbury-Nepean watershed, New South Wales, Australia, to *Ceriodaphnia dubia* and *Selenastrum capricornutum*](#)', *Environmental Toxicology and Chemistry*, 19:72–81 doi:10.1002/etc.5620190108.

Bailey HC, Krassoi R, Elphick JR, Mulhall AM, Hunt P, Tedmanson L and Lovell A (2000b) '[Application of *Ceriodaphnia dubia* for whole effluent toxicity tests in the Hawkesbury-Nepean Watershed, New South Wales, Australia: method development and validation](#)', *Environmental Toxicology and Chemistry*, 19:88–93, doi:10.1002/etc.5620190110.

Baird DJ and Hajibabaei M (2012) '[Biomonitoring 2.0: a new paradigm in ecosystem assessment made possible by next-generation DNA sequencing](#)', *Molecular Ecology*, 21:2039–2044, doi:10.1111/j.1365-294x.2012.05519.x.

Batley GE, van Dam R, Warne MStJ, Chapman JC, Fox DR, Hickey CW and Stauber JL (2018) [Technical rationale for changes to the method for deriving Australian and New Zealand water quality guideline values for toxicants](#), report prepared for the revision of the *Australian and New Zealand guidelines for fresh and marine water quality*, Australian and New Zealand governments and Australian state and territory governments.

Behnisch PA, Hosoe K and Sakai S (2001) '[Bioanalytical screening methods for dioxins and dioxin-like compounds — a review of bioassay/biomarker technology](#)', *Environment International*, 27:413–439, doi:10.1016/s0160-4120(01)00028-9.

Bergerot B, Hugueny B and Belliard J (2015) '[Relating life-history traits, environmental constraints and local extinctions in river fish](#)', *Freshwater Biology*, 60:1279–1291, doi:10.1111/fwb.12561.

Bernot M, Sobota D, Hall R, Mulholland PJ, Dodds W, Webster J, Tank J, Ashkenas L, Cooper L, Dahm C, Gregory S, Grimm N, Hamilton S, Johnson S, Mcdowell W, Meyer J, Peterson B, Poole G, Valett H, Arango C, Beaulieu J, Burgin A, Crenshaw C, Helton A, Johnson L, Merriam J, Niederlehner B, O'Brien J, Potter J, Sheibley R, Thomas S and Wilson K (2010) '[Inter-regional comparison of land-use effects on stream metabolism](#)', *Freshwater Biology*, 55:1874–1890, doi:10.1111/j.1365-2427.2010.02422.x.

Borcard D, Gillet F and Legendre P (2011) *Numerical ecology with R*, Springer, New York, doi:10.1007/978-1-4419-7976-6.

Brix KV, Toll JE, Tear LM, Deforest DK and Adams WJ (2005) '[Setting site-specific water-quality standards by using tissue residue thresholds and bioaccumulation data. Part 2. Calculating site-](#)

specific selenium water-quality standards for protecting fish and birds', *Environmental Toxicology and Chemistry*, 24:231–237, doi:10.1897/03-471.1.

Buchwalter DB, Clements WH and Luoma SN (2017) 'Modernising water quality criteria in the United States: a need to expand the definition of acceptable data', *Environmental Toxicology and Chemistry*, 36:285–291, doi:10.1002/etc.3654.

Burkhard LP and Ankley JL (1989) 'Identifying toxicants: NETAC's toxicity-based approach', *Environmental Science and Technology*, 23:1438–1443, doi:10.1021/es00070a001.

Burrows RM, Fellman JB, Magierowski RH and Barmuta LA (2013) 'Greater phosphorus uptake in forested headwater streams modified by clearfell forestry', *Hydrobiologia*, 703:1–14, doi:10.1007/s10750-012-1332-5.

Carew ME, Pettigrove VJ, Metzeling L and Hoffmann AA (2013) 'Environmental monitoring using next generation sequencing: rapid identification of macroinvertebrate bioindicator species', *Frontiers in Zoology*, 10:45, doi:10.1186/1742-9994-10-45.

Cedergreen N (2014) 'Quantifying synergy: a systematic review of mixture toxicity studies within environmental toxicology', *PLoS ONE*, 9(5):e96580, doi:10.1371/journal.pone.0096580.

Chapman JC (1995) 'The role of ecotoxicity testing in assessing water quality', *Australian Journal of Ecology*, 20:20–27, doi:10.1111/j.1442-9993.1995.tb00518.x.

Chapman PM (2000) 'Whole effluent toxicity (WET) testing — usefulness, level of protection and risk assessment', *Environmental Toxicology and Chemistry*, 19:3–13, doi:10.1002/etc.5620190102.

Chapman PM, Fairbrother A and Brown D (1998) 'A critical evaluation of safety (uncertainty) factors for ecological risk assessment', *Environmental Toxicology and Chemistry*, 17:99–108, doi:10.1002/etc.5620170112.

Chariton AA, Ho KT, Proestou D, Bik HM, Simpson SL, Portis LM, Cantwell MG, Baguley JG, Burgess RM, Pelletier MM, Perron M, Gunsch C and Matthews RA (2014) 'A molecular-based approach for examining responses of eukaryotes in microcosms to contaminant-spiked estuarine sediments', *Environmental Toxicology*, 33:359–369, doi:10.1002/etc.2450.

Chariton AA, Sun M, Gibson J, Webb JA, Leung KMY, Hickey CW and Hose GC (2016) 'Emergent technologies and analytical approaches for understanding the effects of multiple stressors in aquatic environments', *Marine and Freshwater Research*, 67:414, doi:10.1071/MF15190.

Chessman BC (1995) 'Rapid assessment of rivers using macroinvertebrates: a procedure based on habitat-specific sampling, family level identification and a biotic index', *Australian Journal of Ecology*, 20:122–129, doi:10.1111/j.1442-9993.1995.tb00526.x.

Chessman BC (2003) 'New sensitivity grades for Australian macroinvertebrates', *Marine and Freshwater Research*, 54:95–103, doi:10.1071/MF02114.

Chessman BC (2021) 'What's wrong with the Australian River Assessment System (AUSRIVAS)?', *Marine and Freshwater Research*, 72:1110–1117, doi:10.1071/MF20361.

Chèvre N and Vallotton N (2013) '[Pulse exposure in ecotoxicology](#)', in Féraud JF and Blaise C (eds) *Encyclopedia of aquatic ecotoxicology*, Springer, Dordrecht, doi:10.1007/978-94-007-5704-2_84.

Clapcott JE, Young RG, Neale MW, Doehring K and Barmuta LA (2016) '[Land use affects temporal variation in stream metabolism](#)', *Freshwater Science*, 35:1164–1175, doi:10.1086/688872.

Clarke K (1993) '[Non-parametric multivariate analyses of changes in community structure](#)', *Austral Ecology*, 18:117–143, doi:10.1111/j.1442-9993.1993.tb00438.x.

Clements WH and Kiffney PM (1994) '[Integrated laboratory and field approach for assessing impacts of heavy metals at the Arkansas River, Colorado](#)', *Environmental Toxicology and Chemistry*, 13:397–404, doi:10.1002/ETC.5620130306.

Cox B, Oeding S and Taffs K (2019) '[A comparison of macroinvertebrate-based indices for biological assessment of river health: a case example from the sub-tropical Richmond River catchment in northeast New South Wales, Australia](#)', *Ecological Indicators*, 106:105479, doi:10.1016/j.ecolind.2019.105479.

Crook D, Wedd D, Adair B, King A, Mooney T, Harford A and Humphrey C (2021) *Fish migration in Magela Creek and potential impacts of mining-related solutes*, report to the Northern Australia Environmental Resources Hub of the Australian Government's National Environmental Science Program, Charles Darwin University.

Dafforn KA, Johnston EA, Ferguson A, Humphrey C, Monk W, Nichols S, Simpson S, Tulbure M and Baird DJ (2016) '[Big data opportunities for assessing multiple stressors across scales in aquatic ecosystems](#)', *Marine and Freshwater Research*, 67:393–413, doi:10.1071/MF15108.

De Laender F, De Schampelaere KAC, Vanrolleghem PA and Janssen CR (2009) '[Comparing ecotoxicological effect concentrations of chemicals established in multi-species vs. single-species toxicity test systems](#)', *Ecotoxicology and Environmental Safety*, 72:310–315, doi:10.1016/j.ecoenv.2008.07.014.

de Vlaming V and Norberg-King TJ (1999) *A review of single species toxicity tests: are the tests reliable predictors of aquatic ecosystem community responses?*, EPA 440/5-80-029, Office of Research and Development, United States Environmental Protection Agency, United States Government.

de Vlaming V, Connor V, DiGiorgio C, Bailey HC, Deanovic LA and Hinton DE (2000) '[Application of whole effluent toxicity test procedures to ambient water quality assessment](#)', *Environmental Toxicology and Chemistry*, 19:42–62, doi:10.1002/etc.5620190106.

de Vos A, Faux CE, Marthick J, Dickinson J and Jarman SN (2018) '[New determination of prey and parasite species for northern Indian Ocean blue whales](#)', *Frontiers in Marine Science*, 5:104, doi:10.3389/fmars.2018.00104.

de Zwart D and Posthuma L (2005) '[Complex mixture toxicity for single and multiple species: proposed methodologies](#)', *Environmental Toxicology and Chemistry*, 24:2665–2676, doi:10.1897/04-639r.1.

DeForest DK, Brix KV and Adams WJ (2007) 'Assessing metal bioaccumulation in aquatic environments: the inverse relationship between bioaccumulation factors, trophic transfer factors and exposure concentration', *Aquatic Toxicology*, 84:236–246, doi:10.1016/j.aquatox.2007.02.022.

Dolédec S and Statzner B (2008) 'Invertebrate traits for the biomonitoring of large European rivers: an assessment of specific types of human impact', *Freshwater Biology*, 53:617–634, doi:10.1111/j.1365-2427.2007.01924.x.

Downes BJ, Barmuta LA, Fairweather PG, Faith DP, Keough MJ, Lake PS, Mapstone BD and Quinn GP (2002) *Monitoring ecological impacts: concepts and practice in flowing waters*, Cambridge University Press, Cambridge, doi: 10.1017/CBO9780511542015.

Durhan EJ, Norberg-King TJ and Burkhard LP (eds) (1993) *Methods for aquatic toxicity identification evaluations: phase II toxicity identification procedures for samples exhibiting acute and chronic toxicity*, EPA/600/R-92/080, Office of Research and Development, United States Environmental Protection Agency, United States Government.

Egleson KW, Lenat DL, Ausley LW and Winborne FB (1990) 'Comparison of measured instream biological responses with response predicted using the *Ceriodaphnia dubia* chronic toxicity test', *Environmental Toxicology and Chemistry*, 9:1019–1028, doi:10.1002/ETC.5620090808.

Environment Canada (1990a) 'Biological test method: reference method for determining acute lethality of effluents to rainbow trout', *Environmental Protection Series*, Report EPS 1/RM/13, Environment Canada, Conservation and Protection, Canadian Government, amended May 1996.

Environment Canada (1990b) 'Biological test method: reference method for determining acute lethality of effluents to *Daphnia magna*', *Environmental Protection Series*, Report EPS 1/RM/14, Environment Canada, Conservation and Protection, Canadian Government, amended May 1996.

Environment Canada (1992a) 'Biological test method: test of larval growth and survival using fathead minnows', *Environmental Protection Series*, Report EPS 1/RM/22, Environment Canada, Conservation and Protection, Canadian Government, amended November 1997 and March 1998.

Environment Canada (1992b) 'Biological test method: test of reproduction and survival using the cladoceran *Ceriodaphnia dubia*', *Environmental Protection Series*, Report EPS 1/RM/21, Environment Canada, Conservation and Protection, Canadian Government, amended November 1997.

Environment Canada (1992c) 'Biological test method: growth inhibition test using the freshwater alga *Selenastrum capricornutum*', *Environmental Protection Series*, Report EPS 1/RM/25, Environment Canada, Conservation and Protection, Canadian Government, amended November 1997 and March 1998.

Environment Canada (1992d) 'Biological test method: fertilization assay using echinoids (sea urchins and sand dollars)', *Environmental Protection Series*, Report EPS 1/RM/27, Environment Canada, Conservation and Protection, Canadian Government, amended November 1997.

European Commission, Directorate-General for Health and Food Safety (2018) *Technical guidance for deriving environmental quality standards*, European Commission, doi:10.2875/018826.

- Eyre BD and Ferguson AJP (2009) '[Denitrification efficiency for defining critical loads of carbon in shallow coastal ecosystems](#)', *Hydrobiologia*, 629:137–146, doi:10.1007/978-90-481-3385-7_12.
- Fairbrother A, Wenstel R, Sappington K and Wood W (2007) '[Framework for metals risk assessment](#)', *Ecotoxicology and Environmental Safety*, 68:145–227, doi:10.1016/j.ecoenv.2007.03.015.
- Faith DP, Dostine PL and Humphrey CL (1995) '[Detection of mining impacts on aquatic macroinvertebrate communities: results of a disturbance experiment and the design of a multivariate BACIP monitoring programme at Coronation Hill, Northern Territory](#)', *Austral Ecology*, 20:167–180, doi:10.1111/j.1442-9993.1995.tb00530.x.
- Ferrier S, Drielsma M, Manion G and Watson G (2002) '[Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New South Wales. II. Community-level modelling](#)', *Biodiversity and Conservation*, 11:2309–2338, doi:10.1023/A:1021374009951.
- Forbes VE, Salice CJ, Birnir B, Bruins RJF, Calow P, Ducrot V, Galic N, Garber K, Harvey BC, Jager H, Kanarek A, Pastorok R, Railsback SF, Rebarber R and Thorbek P (2017) '[A framework for predicting impacts on ecosystem services from \(sub\)organismal responses to chemicals](#)', *Environmental Toxicology and Chemistry*, 36:845–859, doi:10.1002/etc.3720.
- Fore LS, Karr JR and Wisseman RW (1996) '[Assessing invertebrate responses to human activities: evaluating alternative approaches](#)', *Journal of the North American Benthological Society*, 15:212–231, doi:10.2307/1467949.
- Fox DR, van Dam RA, Fisher R, Batley GE, Tillmans AR, Thorley J, Schwarz CJ, Spry DJ and McTavish K (2021) '[Recent developments in species sensitivity distribution modelling](#)', *Environmental Toxicology and Chemistry*, 40:293–308, doi:10.1002/etc.4925.
- Gholizadeh MH, Melesse AM and Reddi L (2016) '[A comprehensive review on water quality parameters estimation using remote sensing techniques](#)', *Sensors*, 16:1298; doi:10.3390/s16081298.
- Golob JL, Margolis E, Hoffman NG and Fredricks DN (2017) '[Evaluating the accuracy of amplicon-based microbiome computational pipelines on simulated human gut microbial communities](#)', *BMC Bioinformatics*, 18:283, doi:10.1186/s12859-017-1690-0.
- Gray PC, Ridge JT, Poulin SK, Seymour AC, Schwantes AM, Swenson JJ and Johnston DW (2018) '[Integrating drone imagery into high resolution satellite remote sensing assessments of estuarine environments](#)', *Remote Sensing*, 10:1257, doi:10.3390/rs10081257.
- Groom S, Sathyendranath S, Ban Y, Bernard S, Brewin R, Brotas V, Brockmann C, Chauhan P, Choi J, Chuprin A, Ciavatta S, Cipollini P, Donlon C, Franz B, He X, Hirata T, Jackson T, Kampel M, Krasemann H, Lavender S, Pardo-Martinez S, Mélin F, Platt T, Santoleri R, Skakala J, Schaeffer B, Smith M, Steinmetz F, Valente A and Wang M (2019) '[Satellite ocean colour: current status and future perspective](#)', *Frontiers in Marine Science*, 6:485, doi:10.3389/fmars.2019.00485.
- Grosell M, Nielsen C and Bianchini A (2002) '[Sodium turnover rate determines sensitivity to acute copper and silver exposure in freshwater animals](#)', *Comparative Biochemistry and Physiology Part C: Toxicology and Pharmacology*, 133:287–303., doi:10.1016/s1532-0456(02)00085-6.

- Grosell M, Blanchard J, Brix KV and Gerdes R (2007) 'Physiology is pivotal for interactions between salinity and acute copper toxicity to fish and invertebrates', *Aquatic Toxicology*, 84:162–172 doi:10.1016/j.aquatox.2007.03.026.
- Grothe DR and Johnson DE (1996) 'Bacterial interference in whole-effluent toxicity tests', *Environmental Toxicology and Chemistry*, 15:761–764 doi:10.1002/etc.5620150523.
- Grothe DR, Dickson KL and Reed-Judkins DK (eds) (1996) 'Whole effluent toxicity testing: an evaluation of methods and prediction of receiving system impacts', in *Proceedings of the Pellston workshop on whole effluent toxicity, 16–21 September 1995, University of Michigan Biological Station, Douglas Lake, Pellston, Michigan*, Society of Environmental Toxicology and Chemistry Press, Pensacola.
- Growns JE, Chessman BC, McEvoy PK and Wright IA (1995) 'Rapid assessment of rivers using macroinvertebrates: case studies in the Nepean River and Blue Mountains, NSW', *Australian Journal of Ecology*, 20:130–141, doi:10.1111/j.1442-9993.1995.tb00527.x.
- Gunther AJ, Spies RB, Stegeman J, Woodin B, Carney D, Oakden J and Hain L (1997) 'EROD activity in fish as an independent measure of contaminant-induced mortality of invertebrates in sediment bioassays', *Marine Environmental Research*, 44:41–49, doi:10.1016/S0141-1136(96)00101-8.
- Gwinn DC, Beesley LS, Close P, Gawne B and Davies PM (2016) 'Imperfect detection and the determination of environmental flows for fish: challenges, implications and solutions', *Freshwater Biology*, 61:172–180, doi:10.1111/fwb.12684.
- Hall JA and Golding L (1998) *Standard methods for whole effluent toxicity testing: development and application*, National Institute of Water and Atmospheric Research Ltd, Hamilton.
- Hamer M, Maynard SK and Schneider S (2019) 'A pulsed-dose study evaluating chronic toxicity of chlorothalonil to fish: a case study for environmental risk assessment', *Environmental Toxicology and Chemistry*, 38:1549-1559, doi:10.1002/etc.4421.
- Hamilton SJ (2002) 'Rationale for a tissue-based selenium criterion for aquatic life', *Aquatic Toxicology*, 57:85–100, doi:10.1016/s0166-445x(01)00262-4.
- Harford AJ, Simpson SL, Chariton AA, van Dam RA and Humphrey CL (2013) 'The toxicity of uranium (U) to sediment biota of Magela Creek backflow billabong environments', in Supervising Scientist Division (ed) *ERISS research summary 2012–2013*, Supervising Scientist Report 205, Supervising Scientist Division, Australian Government.
- Harris JH and Silveira R (1999) 'Large-scale assessments of river health using an index of biotic integrity with low diversity fish communities', *Freshwater Biology*, 41:235–252, doi:10.1046/j.1365-2427.1999.00428.x.
- Harvey JBJ, Johnson SB, Fisher JL, Peterson WT and Vrijenhoek RC (2017) 'Comparison of morphological and next generation DNA sequencing methods for assessing zooplankton assemblages', *Journal of Experimental Marine Biology and Ecology*, 487:113–126, doi:10.1016/j.jembe.2016.12.002.

Healthy Land and Water (2021) *Report card methods manual, updated July 2021*. Healthy Land and Water, Brisbane.

Hogan AC, Trenfield MA, Harford AJ and van Dam RA (2013) 'Toxicity of magnesium pulses to tropical freshwater species and the development of a duration-based water quality guideline', *Environmental Toxicology and Chemistry*, 32:1969–1980, doi:10.1002/etc.2251.

Holdway DA (1992) 'Control of metal pollution in tropical rivers in Australia', in Connell DW and Hawker DW (eds) *Pollution in tropical aquatic systems*, CRC Press, Boca Raton, doi:10.1201/9781351075879.

Hook SE, Gallagher EP and Batley GE (2014) 'The role of biomarkers in the assessment of aquatic ecosystem health', *Integrated Environmental Assessment and Management*, 10:327–341, doi:10.1002/ieam.1530.

Hopkinson CS Jr and Vallino JJ (1995) 'The relationships among man's activities in watersheds and estuaries: a model of runoff effects on patterns of estuarine community metabolism', *Estuaries*, 18:598–621, doi:10.2307/1352380.

Hose GC and Van den Brink PJ (2004) 'Confirming the species-sensitivity distribution concept for Endosulfan using laboratory, mesocosm and field data', *Archives of Environmental Contamination and Toxicology*, 47:511–520.

Hui FKC, Taskinen S, Pledger S, Foster SD and Warton DI (2015) 'Model-based approaches to unconstrained ordination', *Methods in Ecology and Evolution*, 6:399–411, doi:10.1111/2041-210X.12236.

Jalal A, Salman A, Mian A, Shortis M and Shafait F (2020) 'Fish detection and species classification in underwater environments using deep learning with temporal information', *Ecological Informatics*, 57:101088, doi:10.1016/j.ecoinf.2020.101088.

Jarman SN, Wiseman N, Baker CS and Gales NJ (2006) *Incidence of prey DNA types in Bryde's whale scats*, report to the Scientific Committee of the International Whaling Commission, SC/58 E 29, International Whaling Commission, Cambridge.

Jop KM, Kendall TZ, Askew AM and Foster RB (1991) 'Use of fractionation procedures and extensive chemical analysis for toxicity identification of a chemical plant effluent', *Environmental Toxicology and Chemistry*, 10:981–990, doi:10.1002/etc.5620100802.

Karr RK and Chu EW (1997) *Biological monitoring and assessment: using multimetric indexes effectively*, EPA 235-R97-001, University of Washington, Seattle.

Kefford BJ, Schaefer RB and Metzeling L (2012) 'Risk assessment of salinity and turbidity in Victoria (Australia) to stream insects' community structure does not always protect functional traits', *Science of the Total Environment*, 415:61–68, doi:10.1016/j.scitotenv.2011.05.056.

Kefford BJ, Botwe PK, Brooks AJ, Kunz S, Marchant R, Maxwell S, Metzeling L, Schäfer RB and Thompson RM (2020) 'An integrated database of stream macroinvertebrate traits for Australia: concept and application', *Ecological Indicators*, 114:106280., doi:10.1016/j.ecolind.2020.106280

Kerans BL and Karr JR (1994) 'A benthic index of biotic integrity (B-IBI) for rivers of the Tennessee Valley', *Ecological Applications*, 4:768–85, doi:10.2307/1942007.

Kroon F, Streten C and Harries S (2017) 'A protocol for identifying suitable biomarkers to assess fish health: a systematic review', *PLoS ONE*, 12(4):e0174762, doi:10.1371/journal.pone.0174762.

Lamberti GA and Hauer FR (eds) (2017) *Methods in stream ecology*, 3rd edn, Academic Press, San Diego.

Laroche O, Wood SA, Tremblay LA, Ellis JI, Lear G, Pochon X (2018) 'A cross-taxa study using environmental DNA/RNA metabarcoding to measure biological impacts of offshore oil and gas drilling and production operations', *Marine Pollution Bulletin*, 127:97–107, doi:10.1016/j.marpolbul.2017.11.042.

Legendre P and Legendre L (2012) *Numerical ecology*, 3rd edn, Elsevier, Amsterdam.

Legge R (2012) *Analysis of microbial diversity by amplicon pyrosequencing*, PhD dissertation, Food Science and Technology Department, University of Nebraska–Lincoln.

Luoma SN and Rainbow PS (2005) 'Why is metal bioaccumulation so variable? Biodynamics as a unifying concept', *Environmental Science and Technology*, 39:1921–1931, doi:10.1021/es048947e.

MacDonald AJ and Sarre SD (2017) 'A framework for developing and validating taxon-specific primers for specimen identification from environmental DNA' *Molecular Ecology Resources*, 17:708–720, doi:10.1111/1755-0998.12618.

Mackenzie DI and Royle JA (2005) 'Designing occupancy studies: general advice and allocating survey effort', *Journal of Applied Ecology*, 42:1105–1114, doi:10.1111/j.1365-2664.2005.01098.x.

MacKenzie DI (2006) *Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence*, Elsevier, Burlington.

Madden JC, Rogiers V and Vinken M (2014) 'Application of *in silico* and *in vitro* methods in the development of adverse outcome pathway constructs in wildlife', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369:20130584–20130584, doi:10.1098/rstb.2013.0584.

Magurran AE (2013) *Measuring biological diversity*, 2nd edn, Blackwell Publishing, Oxford.

Maltby L, Clayton SA, Yu H, McLoughlin N, Wood RM and Yin D (2000) 'Using single-species toxicity tests, community-level responses, and toxicity identification evaluations to investigate effluent impacts', *Environmental Toxicology and Chemistry*, 19:151–157, doi:10.1002/etc.5620190118.

Manning TM, Evans JL and Chapman JC (1993) 'The development of toxicity identification and evaluation procedures in Australia', *Chemistry in Australia*, August 1993:398–400.

Marking LL (1977) 'Method for assessing additive toxicity of chemical mixtures', in Mayer FL and Hamelink JL (eds) *Aquatic toxicology and hazard evaluation*, ASTM STP 634, American Society for Testing and Materials, Philadelphia.

Martín-Díaz ML, DelValls TA, Riba I and Blasco J (2008) '[Integrative sediment quality assessment using a biomarker approach: review of 3 years of field research](#)' *Cell Biology and Toxicology*, 24:513–526, doi:10.1007/s10565-008-9062-1.

Matozzo V, Gagné F, Marin MG, Ricciardi F and Blaise C (2008) '[Vitellogenin as a biomarker of exposure to estrogenic compounds in aquatic invertebrates: a review](#)', *Environment International*, 34:531–545, doi:10.1016/j.envint.2007.09.008.

McGeer JC, Brix KV, Skeaff JM, DeForest DK, Brigham SI, Adams WJ and Green A (2003) '[Inverse relationship between bioconcentration factor and exposure concentration for metals: implications for hazard assessment of metals in the aquatic environment](#)', *Environmental Toxicology and Chemistry*, 22:1017–1037, doi:10.1002/etc.5620220509.

McInerney PJ, Rees GN, Gawne B, Suter P, Watson G and Stoffels RJ (2016) '[Invasive willows drive instream community structure](#)', *Freshwater Biology*, 61:1379–1391, doi:10.1111/fwb.12778.

McLellan SL, Huse SM, Mueller-Spitz SR, Andreishcheva EN and Sogin ML (2010) '[Diversity and population structure of sewage-derived microorganisms in wastewater treatment plant influent](#)', *Environmental Microbiology*, 12:378–392, doi:10.1111/j.1462-2920.2009.02075.x.

Merigoux S, Dolédec S and Stutzner B (2001) '[Species traits in relation to habitat variability and state: neotropical juvenile fish in floodplain creeks](#)', *Freshwater Biology*, 46:1251–1267, doi:10.1046/j.1365-2427.2001.00744.x.

Miller DL, Leonard PM, Hughes RM, Karr JR, Moyle PB, Schrader LH, Thompson BA, Daniels RA, Fausch KD, Fitzhugh GA, Gammon JR, Halliwell DB, Angermeier PL and Orth DJ (1988) '[Regional applications of an index of biotic integrity for water resource management](#)', *Fisheries*, 13:12–20, doi:10.1577/1548-8446(1988)013%3C0012:RAOAI0%3E2.0.CO;2.

Moore TF, Canton SP and Grimes M (2000) '[Investigating the incidence of Type I errors for chronic whole effluent toxicity testing using *Ceriodaphnia dubia*](#)', *Environmental Toxicology and Chemistry*, 19:118–122, doi:10.1897/1551-5028(2000)019%3C0118:ITIOTI%3E2.3.CO;2.

Mount DI and Norberg-King TJ (eds) (1993) *Methods for aquatic toxicity identification evaluations: phase III toxicity confirmation procedures for samples exhibiting acute and chronic toxicity*, EPA/600/R-92/081, Office of Research and Development, United States Environmental Protection Agency, United States Government.

Mount DI (1986) 'Principles and concepts of effluent testing', in Bergman L, Kimerle RA and Maki AW (eds) *Environmental hazard assessment of effluents*, Pergamon Press, Oxford.

Niku J, Hui FKC, Taskinen S and Warton DI (2019) '[gllvm: fast analysis of multivariate abundance data with generalized linear latent variable models in R](#)', *Methods in Ecology and Evolution*, 10:2173–2182., doi:10.1111/2041-210X.13303

Niyogi S and Wood CM (2004) '[Biotic ligand model, a flexible tool for developing site-specific water quality guidelines for metals](#)', *Environmental Science and Technology*, 38:6177–6192, doi:10.1021/es0496524.

Noon BR, Bailey LL, Sisk TD and Mckelvey KS (2012) '[Efficient species-level monitoring at the landscape scale](#)', *Conservation Biology*, 26:432–441, doi:0.1111/j.1523-1739.2012.01855.x.

Norberg-King TJ, Mount DI, Durhan EJ, Ankely GT and Burkhard LP (eds) (1991) *Methods for aquatic toxicity identifications: phase I toxicity characterisation procedures*, 2nd edn, EPA/600/6-91/003, Office of Research and Development, United States Environmental Protection Agency, United States Government.

Norberg-King TJ, Mount DI, Amato JR, Jensen DA and Thompson JA (eds) (1992) *Toxicity identification evaluation: phase I characterisation of chronically toxic effluents*, EPA/600/6-91/005F, Office of Research and Development, United States Environmental Protection Agency, United States Government.

Oakes JM, Eyre BD and Ross DJ (2011) '[Short-term enhancement and long-term suppression of denitrification in estuarine sediments receiving primary- and secondary-treated paper and pulp mill discharge](#)', *Environmental Science and Technology*, 45:3400–3406, doi:10.1021/es103636d.

Olsen A, Konovalov D, Philippa B, Ridd P and Wood JC (2020) '[Deep Weeds: a multiclass weed species image dataset for deep learning](#)', *Scientific Reports*, 9:2058, doi:10.1038/s41598-018-38343-3.

Ovaskainen O, Tikhonov G, Norberg A, Guillaume Blanchet F, Duan L, Dunson D, Roslin T and Abrego N (2017) '[How to make more out of community data? A conceptual framework and its implementation as models and software](#)', *Ecology Letters*, 20:561–576., doi:10.1111/ele.12757

Pablo F, Chapman JC and Patra R (1996) 'Development of toxicity characterisation (TIE) procedures using Australian *Ceriodaphnia dubia*', in *Proceedings of the International Symposium on Environmental Chemistry and Toxicology (InterSECT)*, Sydney.

Perceval O, Caquet T, Lagadic L, Bassères A, Azam D, Lacroix G and Poulsen V (2011) '[Mesocosms: their value as tools for managing the quality of aquatic environments](#)', in *Recap document of the Ecotoxicology Symposium 14–16 October 2009 in Le Croisic*, Onema, Vincennes.

Phillips DJH and Rainbow PS (1989) '[Strategies of trace metal sequestration in aquatic organisms](#)', *Marine Environmental Research*, 28:207–210, doi:10.1016/0141-1136(89)90226-2.

Pifher MT and Egan JT (1989) 'Biomonitoring and toxics control: the POTW perspective', *Natural Resources and the Environment*, 4:13(5).

Poff NL, Olden JD, Vieira NKM, Finn DS, Simmons MP and Kondratieff BC (2006) '[Functional trait niches of North American lotic insects: traits-based ecological applications in light of phylogenetic relationships](#)', *Journal of the North American Benthological Society*, 25:730–755 doi:10.1899/0887-3593(2006)025[0730:FTNONA]2.0.CO;2.

Power EA and Boumphrey RS (2004) '[International trends in bioassay use for effluent management](#)', *Ecotoxicology*, 13:377–398, doi:10.1023/B:ECTX.0000035290.89590.03.

Prouse AE, Hogan AC, Harford AJ, van Dam RA and Nuggeoda D (2015) '[Hydra viridissima \(green hydra\) rapidly recover from multiple magnesium pulse exposures](#)', *Environmental Toxicology and Chemistry*, 34:1734–1743, doi:10.1002/etc.2966.

Quetier F, Lavorel S, Thuiller W and Davies I (2007) 'Plant-trait-based modeling assessment of ecosystem-service sensitivity to land-use change', *Ecological Applications*, 17:2377–2386, doi:10.1890/06-0750.1.

Rainbow PS and White SL (1989) 'Comparative strategies of heavy metal accumulation by crustaceans: zinc, copper and cadmium in a decapod, an amphipod and a barnacle', *Hydrobiologia*, 174:245–262, doi:10.1007/BF00008164.

Rainbow PS (2002) 'Trace metal concentrations in aquatic invertebrates: why and so what?' *Environmental Pollution*, 120:497–507, doi:10.1016/S0269-7491(02)00238-5.

Rand GM (ed) (1995) *Fundamentals of aquatic toxicology: effects, environmental fate and risk assessment*, 2nd edn, Taylor and Francis, Washington, DC, doi:10.1201/9781003075363.

Rees HC, Maddison BC, Middleditch DJ, Patmore JRM and Gough KC (2014) 'Review: the detection of aquatic animal species using environmental DNA – a review of eDNA as a survey tool in ecology', *Journal of Applied Ecology*, 51:1450–1459, doi:10.1111/1365-2664.12306.

Resh VH and Jackson JK (1993) 'Rapid assessment approaches to biomonitoring using benthic macroinvertebrates', in Rosenberg DM and Resh VH (eds) *Freshwater biomonitoring and benthic macroinvertebrates*, Chapman and Hall, New York.

Resh VH, Norris RH and Barbour MT (1995) 'Design and implementation of rapid bioassessment approaches for water resource monitoring using benthic macroinvertebrates' *Australian Journal of Ecology*, 20:108–21, doi:10.1111/j.1442-9993.1995.tb00525.x.

Richardson AJ, Walne AW, John AWG, Jonas TD, Lindley JA, Sims DW, Stevens D and Will M (2006) 'Using continuous plankton recorder data', *Progress in Oceanography*, 68:27–74, doi:10.1016/j.pocean.2005.09.011.

Ritchie JC, Zimba PV and Everitt JH (2003) 'Remote sensing techniques to assess water quality', *Photogrammetric Engineering and Remote Sensing*, 69:695–704, doi:10.14358/PERS.69.6.695.

Robson HLA, Noble TH, Saunders RJ, Robson SKA, Burrows DW and Jerry DR (2016) 'Fine-tuning for the tropics: application of eDNA technology for invasive fish detection in tropical freshwater ecosystems', *Molecular Ecology Resources*, 16:922–932, doi:10.1111/1755-0998.12505.

Rohr JR, Salice CJ and Nisbe RM (2016) 'The pros and cons of ecological risk assessment based on data from different levels of biological organization', *Critical Reviews in Toxicology*, 46:756–784, doi:10.1080/10408444.2016.1190685.

Ruffier PJ (1996) 'Whole effluent toxicity testing: the municipal perspective', in Grothe DR, Dickson KL and Reed-Judkins DK (eds) *Whole effluent toxicity testing: an evaluation of methods and prediction of receiving system impacts*, Society of Environmental Toxicology and Chemistry Press, Pensacola.

Sarakinos HC and Rasmussen JB (1998) 'Use of bioassay-based whole effluent toxicity (WET) tests to predict benthic community responses to a complex industrial effluent', *Journal of Aquatic Ecosystem Stress and Recovery*, 6:141–157, doi:10.1023/A:1009947824130.

Schäfer RB, Kefford BJ, Metzeling L, Liess M, Burgert S, Marchant R, Pettigrove V, Goonan P and Nugegoda D (2011) 'A trait database of stream invertebrates for the ecological risk assessment of single and combined effects of salinity and pesticides in south-east Australia', *Science of the Total Environment*, 409:2055–63, doi:10.1016/j.scitotenv.2011.01.053.

Shaw JLA, Clarke LJ, Wedderburn SD, Barnes TC, Weyrich LS and Cooper A (2016) 'Comparison of environmental DNA metabarcoding and conventional fish survey methods in a river system', *Biological Conservation*, 197:131–138, doi:10.1016/j.biocon.2016.03.010.

Simmonds J and MacLennan J (2005) *Fisheries acoustics: theory and practices*, 2nd edn, Blackwell Publishing, Oxford.

Simpson SL and Batley GE (2016) *Sediment quality assessment: a practical handbook*, CSIRO Publishing, Clayton.

Simpson SL, Batley GE and Chariton AA (2013) *Revision of the ANZECC/ARMCANZ sediment quality guidelines*, CSIRO Land and Water Report 8/07, CSIRO Land and Water, doi: 10.4225/08/5894c6184320c.

Sinclair A, Tayler K, van Dam R and Hogan A (2014) 'Site-specific water quality guidelines: 2. Development of a water quality regulation framework for pulse exposures of mine water discharges at a uranium mine in northern Australia', *Environmental Science and Pollution Research* 21:131–140, doi:10.1007/s11356-013-1922-4.

Smith REW, Boulton AJ, Baldwin DS, Humphrey CL, Butler B and Halse S (2020) *Assessing and managing water quality in temporary waters, Australian and New Zealand guidelines for fresh and marine water quality*, Australian and New Zealand governments and Australian state and territory governments.

Sprague JB (1990) 'Aquatic toxicology', in Schreck CB and Moyle PB (eds) *Methods for fish biology*, American Fisheries Society, Maryland, doi:10.47886/9780913235584.

Stark D (1993) 'Performance of the macroinvertebrate community index: effects of sampling method, sample replication, water depth, current velocity and substratum on index values', *New Zealand Journal of Marine and Freshwater Research*, 27:463–478, doi:10.1080/00288330.1993.9516588.

Stark JD and Maxted JR (2007) *A user guide for the macroinvertebrate community index*, Cawthron Report No 1166, Cawthron Institute, Nelson.

Strom D, Ralph PJ and Stauber JL (2009) 'Development of a toxicity identification evaluation protocol using chlorophyll-a fluorescence in a marine microalga' *Archives of Environmental Contamination and Toxicology*, 56:30–38, doi:10.1007/s00244-008-9174-x.

Stuart-Smith RD, Bates AE, Lefcheck JS, Duffy JE, Baker SC, Thomson RJ, Stuart-Smith JF, Hill NA, Kininmonth SJ, Airoidi L, Becerro MA, Campbell SJ, Dawson T, Navarrete SA, Soler GA, Strain EMA, Willis TJ and Edgar GJ (2013) 'Integrating abundance and functional traits reveals new global hotspots of fish diversity', *Nature*, 501:539–42, doi:10.1038/nature12529.

Supervising Scientist (2011) *Environmental monitoring protocols to assess potential impacts from Ranger minesite on aquatic ecosystems: in situ toxicity monitoring – freshwater snail, *Amerianna cumingi*, reproduction test*, Internal Report 588, Supervising Scientist Division, Australian Government.

Supervising Scientist (2013) *Environmental monitoring protocols to assess potential impacts from Ranger minesite on aquatic ecosystems: macroinvertebrate community structure in streams*, Internal Report 591, Supervising Scientist Division, Australian Government.

Swanson S, Kozlowski D, Hall R, Heggem D and Lin J (2017) ‘Riparian proper functioning condition assessment to improve watershed management for water quality’, *Journal of Soil and Water Conservation*, 72:168–182, doi:10.2489/jswc.72.2.168.

Syberg K, Jensen TS, Cedergreen N and Rank J (2009) ‘On the use of mixture toxicity assessment in REACH and the water framework directive: a review’, *Human and Ecological Risk Assessment: an International Journal*, 15:1257–1272, doi:10.1080/10807030903304922.

ter Braak CJF (2019) ‘New robust weighted averaging- and model-based methods for assessing trait–environment relationships’, *Methods in Evolution and Ecology*, 10:1962–1971, doi:10.1111/2041-210X.13278.

Toll JE, Tear LM, DeForest DK, Brix KV and Adams WJ (2005) ‘Setting site-specific water-quality standards by using tissue residue criteria and bioaccumulation data. Part 1. Methodology’, *Environmental Toxicology and Chemistry*, 24:224–230, doi:10.1897/03-472.1.

Trenfield MA, Pease CJ, Walker SL, Mooney T, Tybell L, Humphrey C, van Dam RA and Harford AJ (2020) ‘Standardised chronic toxicity test protocols and culturing methods for a suite of tropical freshwater species’, *Australasian Bulletin of Ecotoxicology and Environmental Chemistry*, 6:1–80.

Tullos DD, Penrose DL, Jennings GD and Cope WG (2009) ‘Analysis of functional traits in reconfigured channels: implications for the bioassessment and disturbance of river restoration’, *Journal of the North American Benthological Society*, 28:80–92, doi:10.1899/07-122.1.

US EPA (United States Environmental Protection Agency) (2002a) *Methods for measuring the acute toxicity of effluents and receiving waters to freshwater and marine organisms*, 5th edn, EPA-821-R-02-012, Office of Water, US EPA, United States Government, Washington, DC.

US EPA (United States Environmental Protection Agency) (2002b) *Short-term methods for estimating the chronic toxicity of effluents and receiving waters to freshwater organisms*, 4th edn, EPA-821-R-02-013, Office of Water, US EPA, United States Government, Washington, DC.

US EPA (United States Environmental Protection Agency) (2002c) *Short-term methods for estimating the chronic toxicity of effluents and receiving waters to marine and estuarine organisms*, 4th edn, EPA-821-R-02-014, Office of Water, US EPA, United States Government, Washington, DC.

US EPA (United States Environmental Protection Agency) (2016) *Aquatic life ambient water quality criteria for selenium– freshwater*, EPA 822-R-16-006, Office of Water and Office of Science and Technology, US EPA, United States Government, Washington, DC.

van Dam RA and Chapman JC (2001) 'Direct toxicity assessment (DTA) for water quality guidelines in Australia and New Zealand', *Australasian Journal of Ecotoxicology*, 7:175–198.

van Dam RA, Camilleri C and Finlayson CM (1998) 'The potential of rapid assessment techniques as early warning indicators of wetland degradation: a review', *Environmental Toxicology and Water Quality*, 13(4):297–312, doi:10.1002/(SICI)1098-2256(1998)13:4%3C297::AID-TOX3%3E3.0.CO;2-2.

van Dam RA, Humphrey CL, Harford AJ, Sinclair A, Jones DR, Davies S and Storey AW (2014) 'Site-specific water quality guidelines: 1. Derivation approaches based on physicochemical, ecotoxicological and ecological data', *Environmental Science and Pollution Research*, 21:118–130, doi:10.1007/s11356-013-1780-0.

van Dam RA, Hogan, AC, Humphrey CL and Harford AJ (2019) 'How specific is site-specific? A review and guidance for selecting and evaluating approaches for deriving local water quality benchmarks', *Integrated Environmental Assessment and Management*, 15:683–702, doi:10.1002/ieam.4181.

Van den Brink PJ, Blake N, Brock TCM and Maltby L (2006) 'Predictive value of species sensitivity distributions for effects of herbicides in freshwater ecosystems', *Human and Ecological Risk Assessment: An International Journal*, 12:645–674, doi:10.1080/10807030500430559.

Van den Brink PJ, Baird DJ, Baveco H and Focks A (2013) 'The use of traits-based approaches and eco(toxico)logical models to advance the ecological risk assessment framework for chemicals', *Integrated Environmental Assessment and Management*, 9:E47–E57, doi:10.1002/ieam.1443.

van der Oost R, Beyer J and Vermeulen NPE (2003) 'Fish bioaccumulation and biomarkers in environmental risk assessment: a review', *Environmental Toxicology and Pharmacology*, 13:57–149, doi:10.1016/s1382-6689(02)00126-6.

Verbruggen EMJ and Van den Brink PJ (2010) *Review of recent literature concerning mixture toxicity of pesticides to aquatic organisms*, Report 601400001/2010, RIVM, Bilthoven.

Versteeg DJ, Belanger SE and Carr GJ (1999) 'Understanding single-species and model ecosystem sensitivity: data-based comparison', *Environmental Toxicology and Chemistry*, 18:1329–1346, doi:10.1002/etc.5620180636.

Viarengo A, Lowe D, Bolognesi C, Fabbri E and Koehler A (2007) 'The use of biomarkers in biomonitoring: a 2-tier approach assessing the level of pollutant-induced stress syndrome in sentinel organisms', *Comparative Biochemistry and Physiology Part C: Toxicology and Pharmacology*, 146:281–30, doi:10.1016/j.cbpc.2007.04.011.

Vinken M (2013) 'The adverse outcome pathway concept: a pragmatic tool in toxicology', *Toxicology*, 312:158–165, doi:10.1016/j.tox.2013.08.011.

Walker CH (1999) 'The use of biomarkers in vertebrates', in Peakall DB, Walker CH and Migula P (eds) *Biomarkers: a pragmatic basis for remediation of severe pollution in eastern Europe*, NATO Science Series 2: Environmental Security, vol 54, Springer Science+Business Media, Dordrecht.

Waller WT, Ammann LP, Birge WJ, Dickson KL, Dorn PB, LeBlanc NE, Mount DI, Parkhurst BR, Preston HR, Schimmel SC, Spaci A and Thursby GB (1996) 'Predicting instream effects from WET tests', in

Grothe DR, Dickson KL and Reed-Judkins DK (eds) *Whole effluent toxicity testing: an evaluation of methods and prediction of receiving system impacts*, Society of Environmental Toxicology and Chemistry Press, Pensacola.

Warne MStJ (1998) *Critical review of methods to derive water quality guidelines for toxicants and a proposal for a new framework*, Supervising Scientist Report 135, Supervising Scientist Division, Australian Government.

Warne MStJ, Batley GE, van Dam RA, Chapman JV, Fox DR, Hickey CW and Stauber JL (2018) *Revised method for deriving Australian and New Zealand water quality guideline values for toxicants*, report prepared for the revision of the *Australian and New Zealand guidelines for fresh and marine water quality*, Australian and New Zealand governments and Australian state and territory governments.

Warton DI, Wright ST and Wang Y (2012) 'Distance-based multivariate analyses confound location and dispersion effects', *Methods in Ecology and Evolution*, 3:89–101, doi:10.1111/j.2041-210X.2011.00127.x.

Warton DI, Blanchet FG, O'Hara RB, Ovaskainen O, Taskinen S, Walker SC and Hui FKC (2015a) 'So many variables: joint modeling in community ecology', *Trends in Ecology and Evolution*, 30:766–779, doi:10.1016/j.tree.2015.09.007.

Warton DI, Foster SD, De'ath G, Stoklosa J and Dunstan PK (2015b) 'Model-based thinking for community ecology' *Plant Ecology*, 216:669–682, doi:10.1007/s11258-014-0366-3.

Washington HG (1984) 'Diversity, biotic and similarity indexes – a review with special relevance to aquatic ecosystems', *Water Research*, 18:653–694, doi:10.1016/0043-1354(84)90164-7.

White SL and Rainbow PS (1985) 'On the metabolic requirements for copper and zinc in molluscs and crustaceans', *Marine Environmental Research*, 16:215–229, doi:10.1016/0141-1136(85)90139-4.