Environ Sci Pollut Res (2014) 21:51–60
DOI 10.1007/s11356-013-1779-6

Revisions to the derivation of the Australian and New Zealand guidelines for toxicants in fresh and marine waters

M. St. J. Warne · G. E. Batley · O. Braga · J. C. Chapman · D. R. Fox · C. W. Hickey · J. L. Stauber · R. Van Dam

Received: 31 January 2013 / Accepted: 26 April 2013 / Published online: 25 June 2013

Abstract The Australian and New Zealand Guidelines for Fresh and Marine Water Quality are a key document in the Australian National Water Quality Management Strategy. These guidelines released in 2000 are currently being reviewed and updated. The revision is being co-ordinated by the Australian Department of Sustainability, Environment, Water, Population and Communities, while technical matters are dealt with by a series of Working Groups. The revision will be evolutionary in nature reflecting the latest scientific developments and a range of stakeholder desires. Key changes will be: increasing the types and sources of data that can be used; working collaboratively with industry to permit the use of commercial-in-confidence data; increasing the minimum data requirements; including a measure of the uncertainty of the trigger value; improving the software used to calculate trigger values; increasing the rigour of site-specific trigger values; improving the method for assessing the reliability of the trigger values; and providing guidance of measures of toxicity and toxicological endpoints that may, in the near future, be appropriate for trigger value derivation. These changes will markedly improve the number and quality of the trigger values that can be derived and will increase end-users' ability to understand and implement the guidelines in a scientifically rigorous manner.

Keywords Environmental quality standards · Toxicity · Species sensitivity distribution · Australia · New Zealand

Responsible editor: Philippe Garrigues

M. St. J. Warne (✉)
Water Quality and Investigations, Environmental Monitoring and Assessment Science, Science Delivery, Department of Science, Information Technology, Innovation and the Arts, Brisbane, QLD, Australia
e-mail: michael.warne@science.dsitia.qld.gov.au

G. E. Batley · J. L. Stauber
Centre for Environmental Contaminants Research, CSIRO Land and Water, Lucas Heights, NSW, Australia

O. Braga
Department of Sustainability, Environment, Water, Population and Communities, Canberra, Australia

J. C. Chapman
Office of Environment & Heritage, Lidcombe, NSW, Australia

D. R. Fox
Environmetrics, Melbourne, VIC, Australia

C. W. Hickey
National Institute of Water and Atmospheric Research (NIWA), Hamilton, New Zealand

R. Van Dam
Environmental Research Institute of the Supervising Scientist, Darwin, NT, Australia

Introduction

The quality of water in Australia is managed via the National Water Quality Management Strategy (NWQMS; ANZECC 1992). The NWQMS currently consists of 24 documents that cover policies, processes and guidelines. The guidelines cover many aspects of water quality; however, two of the key documents are the Australian and New Zealand Guidelines for Fresh and Marine Water Quality (hereafter referred to as the Guidelines; ANZECC/ARMCANZ 2000a) and the Australian Guidelines for Water Quality Monitoring and Reporting (ANZECC/ARMCANZ 2000b). These two guidelines are currently under review to ensure that they remain up to date with the latest scientific advances, maintain their relevance and retain their value as a national best practice tool for water quality management. The Department of Sustainability, Environment, Water, Population and

Communities is co-ordinating the review, while technical matters are dealt with by a series of Working Groups. Initial scoping of the guideline revision requirements was undertaken by a series of Working Groups—each consisting of experts with appropriate knowledge. The review of the portion of the Guidelines relating to toxicants was undertaken by the Fresh and Marine Water Quality Working Group 4 (FMWQ Working Group 4)—a group of experts in the field of ecotoxicology, statistics and the derivation of trigger values. Given the focus of the International Conference on Deriving Environmental Quality Standards for the Protection of Aquatic Ecosystems Conference (Hong Kong, 3–7 December 2011) and the current Special Issue of Environmental Science and Pollution Research, this manuscript focuses on the key drivers of change to the Guidelines for toxicants and the recommendations of the FMWQ Working Group 4 to address these. The recommendations of the FMWQ Working Group 4 or the other Working Groups have not been approved at the time of writing. Therefore, the manuscript represents the views of the members of FMWQ Working Group 4 and not those of the Australian or New Zealand Governments.

The key drivers for change are:

- Expanding the use of different types of toxicity data
- Increasing regional specificity and the use of site-specific investigations
- Increasing the usefulness of site-specific investigation data
- Incorporating and presenting uncertainty
- Improving the BurrliOZ software
- Increasing the sources of toxicity data that can be used
- Improving the assessment of the reliability of trigger values
- Increasing flexibility of the guideline derivation process and providing guidance on how to address issues that may arise in the future
- Increasing international collaboration and harmonisation

Recommended approaches to address each of these drivers are discussed below.

Expanding the use of different types of toxicity data

Chronic rather than acute toxicity data are preferred to derive guideline trigger values for toxicants, as they are more appropriate to achieve the overall aim of the Guidelines to provide life-long protection for aquatic organisms and hence, it is assumed, for aquatic ecosystems. The majority of chronic toxicity data are hypothesis-based values such as the no observed effect concentrations (NOECs) and the lowest observed effect concentrations (LOECs), but both of these types of data have come under persistent criticism since the 1990s (e.g. Hoekstra and Van Ewijk 1993; Noppert et al. 1994; Chapman et al. 1996; OECD 2006), with a recent revived push to limit their generation and use (Newman 2008; Warne and Van Dam 2008; Landis and Chapman 2011; Jager 2012; Van Dam et al. 2012a). No observed effect concentrations have been incorporated into various legislative compliance measures for discharge consent and monitoring. This will make it more difficult and slower to stop the use of NOECs, but having clear guidance in the Guidelines will greatly facilitate this process.

The most accepted replacements of NOECs have been low effect measures of toxicity such as the concentration that causes a 10 % effect (EC10) and that causes 10 % lethality (LC10) (Van der Hoeven 1997; CCME 2007; Warne and Van Dam 2008); however, several authors have advocated the use of no effect concentrations (NECs; Van der Hoeven 1997; Fox 2009, 2010). The NEC is represented by a modelled threshold concentration, below which no adverse effects occur, and above which adverse effects do occur. The NEC concept has considerable merit, and is theoretically well suited to the purpose of protecting aquatic ecosystems. However, there are doubts over whether a threshold concentration exists for all chemicals and there has also been concern that there will seldom be sufficient data available to validate the choice of model used to determine the NECs (Van der Hoeven 1997). To overcome, some of these issues Fox (2009, 2010) proposed the NEC be calculated using a Bayesian approach. Use of the NEC has merit, but it was felt that additional testing of the robustness of NECs to a wide variety of toxicity data was required. This should include analysis of the sensitivity of the NEC to the prior distributions required for the Bayesian method, and development of a relatively simple to use and understand software, before it could be considered a suitable toxicity estimate for deriving trigger values. Nonetheless, the NEC is viewed as a potential preferred measure of toxicity in the future.

In revising the Guidelines, it was considered important to: (1) permit the use of more measures of toxicity to derive trigger values; (2) provide a stated of preference for using the permitted various measures; and (3) provide recommendations for future work and revision of the guidelines. Increasing the types of toxicity measures that can be used has the benefit of increasing sample sizes used for species sensitivity distributions (SSDs) which, consequently, increases the confidence in the derived trigger values. The recommended approach is similar to that adopted by Canada (CCME 2007), which has a list of preferred measures of toxicity. This will have a clearly stated preference for toxicity estimates derived from regression-based approaches (e.g. concentrations that are lethal to a certain percentage ‘x’ of the individuals (LCx) or the concentrations that cause ‘x’ per cent of individuals to experience a given effect or the concentration which on average causes a ‘x’ per cent effect

(ECx)) rather than hypothesis-based approaches (e.g. NOEC and LOEC). It was acknowledged, however, that as the vast majority of existing chronic toxicity data are NOEC data, they will have to continue to be used to derive trigger values until there are sufficient ECx data. The current Guidelines (ANZECC/ARMCANZ 2000a, b), reflecting the recommendations of Warne (1998), recognised the limitations of NOEC data and recommended their use for guideline trigger value derivation be phased out as EC10 data became available. However, no specific means of pursuing this recommendation were provided. To encourage the generation of EC10 data, LC10 and NOEC data will initially be deemed to be equivalent, and combinations of such data can be used to derive trigger values. However, once there are sufficient EC10 and/or LC10 data to meet the minimum data requirements (refer to "Increasing the regional specificity and the use of site-specific investigations" section), NOEC data should not be used to calculate trigger values.

The move from NOEC to EC10 and LC10 data will necessitate changes to the experimental design of ecotoxicity tests. In order to use hypothesis-based statistical methods to calculate NOEC values, all treatments must be conducted at least in triplicate. In contrast, for regression-based methods used to calculate ECx values, replication is less important than having more treatments, particularly those that are likely to exert biological effects of less than 50 %. Thus, it was recommended that the revised Guidelines include guidance on the design of experiments for concentration–response modelling similar to that provided by the OECD (2006) and Canada (CCME 2007).

Increased flexibility should also be provided by permitting the use of actual chronic and estimates of chronic toxicity data to derive trigger values. The Guidelines currently do not permit this; rather trigger values are either generated using only chronic or only acute toxicity data, with a decreasing level of reliability assigned to the resultant trigger values, accordingly. Estimates of chronic toxicity could be derived by dividing acute toxicity data by default assessment factors or acute to chronic ratios. Obviously, the use of estimates of chronic toxicity data will affect the reliability of the resulting trigger value, and this will need to be addressed by the scheme that assesses and assigns the reliability of trigger values (refer to "Improving the assessment of the reliability of trigger values" section).

Increasing regional specificity and the use of site-specific investigations

Regional specificity

The Guidelines do not place any preference on the use of toxicity data from Australian or New Zealand species. Rather, given the statistical nature of the SSD method that is the preferred method of deriving trigger values, greater emphasis is placed on maximising the number of species and taxa for which toxicity data are available.

The US water quality guidelines (USEPA 2007a) require all of the toxicity data used to derive limits are for species that live and breed in continental North America. Canada previously had similar requirements, but has recently modified its position so that "species that are non-resident to Canada can be used if it can be demonstrated that they are acceptable surrogate species for Canadian resident species and the studies were conducted under exposure conditions representative of Canadian waters" (CCME 2007). While not explicitly stated, it is assumed that there was concern in both these jurisdictions that non-resident overseas species might have different sensitivities to North American species and, thus, the resulting trigger values may be either over- or under-protective.

The potential issue of Australasian species having different sensitivities to those from elsewhere has long been of interest in Australia, particularly given our geologically long separation from most Northern Hemisphere continents. However, the studies that have addressed this have all suffered from only comparing data for a limited number of chemicals or a limited number of species (e.g. Johnston et al. 1990; Sunderam et al. 1992; Davies et al. 1994; Hickey and Martin 1995; Markich and Camilleri 1997; Mulhall 1997; Rose et al. 1998; Hickey 2000; Hose and Van den Brink 2004; Phyu 2004; Westbury et al. 2004) and their conclusions were contradictory. Thus, a general conclusion could not be reached.

A number of more recent larger studies have attempted to address the issue of different sensitivities by comparing the sensitivities of species from different zoogeographical areas in the USA (Dyer et al. 1997), Europe (Maltby et al. 2003), Australasia (Hobbs et al. 2004) as well as in tropical (Kwok et al. 2007; Rombke et al. 2008; Daam and Van den Brink 2010; Sanchez-Bayo and Hyne 2011) and polar (Chapman et al. 2006) regions. The results of these studies have been inconsistent. Hobbs et al. (2004) and Kwok et al. (2007) both found species from different regions had higher, equal and lower sensitivities with no apparent reason for the differences. While Chapman et al. (2006) did not find any consistent trend in sensitivity across three climatic zones (temperate, tropical and polar), they nonetheless concluded that "toxicity data from one geographic region will not be universally protective of other regions."

In revising the toxicant trigger value derivation method, it was considered that there were only a small percentage of cases where differences in sensitivities of species from different regions occurred and that these differences were not large (typically less than 1 order of magnitude). Thus, it was not recommended that regional requirements be included in

the rules governing data that can be used to derive national TVs. However, the case for increased regionalisation was recognised as being appropriate for deriving site-specific trigger values (see next section).

Site-specific investigations

When guideline trigger values have been exceeded at a site there are two basic options: (1) to commence management actions to rectify the situation (e.g. remediate the site, decrease the concentration of contaminants being released) or (2) to conduct further investigation (termed site-specific investigation) to determine the relevance of the default trigger values to the site and ultimately to derive site-specific trigger values. The decision on how to proceed is a matter for the proponent and to a large degree this is a risk or cost-benefit management decision, a process that most commercial organisations are familiar with. The current Guidelines strongly encourage site-specific investigation and have provided a number of decisions trees which provide a scientifically logical and cost effective means of conducting these. However, one of the key drivers of the Guidelines is flexibility in the approach taken—in other words, it is result rather than process driven—providing the process used is scientifically rigorous and defensible. Van Dam et al. (2013) and Sinclair et al. (2013) provide several examples of the use of site-specific ecotoxicity data for deriving site-specific trigger values.

Prior to the release of the 2000 Guidelines (ANZECC/ARMCANZ 2000a), there was considerable resistance to the concept of site-specific investigation. Industry and commercial entities believed that they would be forced to undertake site-specific investigations. However, subsequently, these entities have come to value and appreciate the benefit of having such a flexible system, and many are now strong advocates of site-specific investigations. Hundreds of site-specific investigations have been conducted throughout Australia and New Zealand, and very frequently they are being conducted as part of the ecological impact study and development approval stages of proposed developments. For example, in response to the prolonged drought in eastern Australia up to 2008, many states proposed to build desalination plants. In every case, site-specific ecotoxicological investigations were conducted in the approvals phase (e.g. Warne 2010).

These site-specific investigations have often taken the form of direct toxicity assessments (DTAs), which is the equivalent of whole effluent toxicity testing. There are a number of factors that control the number and types of test species used in DTA tests (Van Dam and Chapman 2001). However, one of the key drivers from the public and conservation groups' point of view has been the desire that the test species be locally relevant. While such desires are sound, they run up against practicality issues such as the availability of toxicity test methods (in most jurisdictions and particularly Australia and New Zealand, there are a limited number of tests that have been developed for indigenous species) (e.g. Hall and Golding 1998; Van Dam et al. 2008) and the time available to undertake the approval process (it takes quite a reasonable time to develop new toxicity tests and this may not meet the development timeframe). Nonetheless, researchers and regulators (Van Dam, personal communication) and commercial ecotoxicity testing organisations are responding by developing a range of new local test species including more marine and tropical test species (Krassoi, personal communication). For example, as part of the DTA for the proposed Olympic Dam desalination plant in South Australia, a series of new toxicity tests were developed for the locally important Australian Giant Cuttlefish (*Sepia aparma*) by Geotechnical Services (2006).

It is highly unlikely that this desire to use locally relevant species in DTA will diminish. Rather, it is likely that, as the public and conservation groups become increasingly knowledgeable about DTA and the Guidelines, their demands will require ever-increasingly site-relevant species and information. This in turn will drive further developments, innovations and new, faster and more efficient ways of developing ecotoxicity tests.

The increased public scrutiny of the results of site-specific investigations in development approvals, discharge compliance and dredge-spoil monitoring has already led to improved scientific rigour and this trend will continue. Increasingly, the science involved in site-specific investigations is subject to independent peer review (e.g. the Independent Review Panel for the Victorian Desalination Plant or the South Australian and Australian Government review of the EIS for the Olympic Dam Desalination plant) or public scrutiny via public hearings (e.g. for the Victorian Desalination Plant) or court cases. One outcome has been the realisation that some site-specific investigations have simply met the 'minimum requirements' to derive site-specific trigger values. Many governments and public stakeholders are now demanding more than the minimum and industrial and commercial entities that place a high importance on their environmental credentials, image and goodwill are now pre-emptively meeting this and doing additional work. For example, while the Guidelines only require DTA to use a minimum of five test species that belong to at least four taxonomic groups (at the phyla level), the DTA for the proposed Olympic Dam desalination plant determined the toxicity of the saline return water to 16 species and used chronic toxicity data from 7 species that belonged to 6 taxonomic groups to derive the site-specific trigger values (Warne 2010).

Concern has been expressed that the trigger values in the Guidelines, which are often based on far more than the minimum data requirements, are being replaced by site-specific trigger values based, in many cases, on the minimum number of data. Although ANZECC/ARMCANZ (2000a, b) cautions against excluding comprehensive datasets of non-local species in favour of potentially small datasets of local species, this practice is known to often occur. This concern, combined with the increased importance that site-specific investigations and site-specific trigger values are playing in regulating and managing chemicals in the environment, has been one driver to increase the minimum number of species and taxa that can be used to derive trigger values. The other driver has been recent moves to increase the number of species by the European Commission (EC 2011). In the revision of the Guidelines, it was proposed that the minimum toxicity data requirements to derive a national trigger value be increased to toxicity data for at least eight species that belong to at least four taxonomic groups. The SSD method assumes that the toxicity data are for a random selection of species and phyla in the ecosystem being considered. Therefore, in both the current and proposed versions of the Guidelines, there are no requirements for specific phyla or specific organisms to be part of the minimum data requirements. This is a considerable increase on the current minimum data requirements (by 60 %), but still not as large as the USA (USEPA 1999) or EU requirements (EC 2011). However, the minimum data requirements of all these jurisdictions fall considerably short of the minimum data requirements recommended by Newman et al. (2000) of between 15 and 55 with a median of 30 and by Wheeler et al. (2002) of 10–15 species. The reason for not increasing the minimum data requirements to at least 10 (as per Wheeler et al. 2002) was a matter of balancing competing factors. For many chemicals, there are high quality toxicity data for less than 10 species. It was considered preferable to maximise the number of chemicals which could have trigger values derived by the more scientifically rigorous SSD method, albeit with a less stringent data requirement, than imposing more stringent data requirements and having fewer trigger values derived by the SSD method.

The proposed new minimum data requirements to derive national trigger values for Australia and New Zealand are based on pragmatic decisions to incrementally increase the scientific rigour of the trigger values while acknowledging the current situation with limited DTA tests available in Australia and New Zealand. However, due to the limited number of toxicity tests for Australian species, the recommended minimum data requirement to derive a site-specific trigger value is toxicity data for at least five species that belong to at least four taxonomic groups.

Increasing the usefulness of site-specific investigation data

Considerable amounts of toxicity data are being generated through the site-specific investigations. Yet, the toxicity data generated are only used to assess the site being investigated. However, where there are similarities in water chemistry, ecosystems and toxicants at sites, it might be possible to use data from previously conducted site-specific investigations at new sites. This often does not happen, because the site-specific data are the property of the company paying for the site-specific investigation, and they do not want to provide a commercial advantage to potentially competing companies or to inadvertently reveal commercially sensitive information. Even if these issues could be overcome, there would need to be a central repository from which to store and extract the necessary reports and toxicity data. As with all data being considered for deriving trigger values, data quality and appropriateness must be assessed. If the commercial considerations of the data owners could be addressed, there would be significant financial benefits to all companies and the generation of far more extensive information on which to make environmentally responsible decisions.

Incorporating and presenting uncertainty

In the Guidelines, each trigger value is presented as a single unique value without any indication of uncertainty. The Monitoring Guidelines (ANZECC/ARMCANZ 2000b) provide a means of comparing monitoring data against the trigger values to determine if the latter have been exceeded, but this only considers the uncertainty in the monitoring data, not the uncertainty in the trigger values.

In the current revision of the Guidelines, it was recommended that information about the uncertainty associated with trigger values be included. The trigger values should be presented with their 95 % confidence limits (CLs). In addition, the BurrliOZ software (Campbell et al. 2000), which is used to calculate the national and site-specific trigger values (see "Improving the BurrliOZ software program" section below) is currently being improved to graphically and numerically present the 95 % CLs associated with the trigger values (JSC 2010). The 95 % CLs are a measure of the uncertainty associated with predicting the trigger value using the statistical distribution that best fits the available toxicity data. The smaller the 95 % CLs the less uncertainty there is in predicting the trigger value and conversely the larger the 95 % CLs the larger the uncertainty in the trigger value. Another major source of uncertainty that will remain unquantified results from the fact that the sensitivity of each species is represented by a single value in the SSD, despite being based on multiple values, from

multiple tests, sources, endpoints, durations and life stages. The confidence limits also do not include the uncertainties in the derivation of the toxicity datum used to derive trigger values. A long-term goal would be the inclusion of statistical uncertainty throughout the entire process of calculating trigger values, as proposed by Shao and Warne (2002).

A potential perverse outcome associated with incorporating measures of uncertainty may be that it increases the complexity of the Guidelines and decreases their comprehension and useability. This could be overcome by providing specific guidance to users on the purpose of the confidence limits and how they are intended to be used.

Improving the BurrliOZ software program

BurrliOZ was developed by the Commonwealth Scientific and Industrial Research Organisation (Campbell et al. 2000) based on the earlier work by Shao (2000). It was the SSD method used to derive all of the toxicant trigger values in the Guidelines.

In addition, to graphically and mathematically present the 95 % CLs of the trigger values (as discussed in the preceding section), the BurrliOZ software is being improved in the following ways:

- the type of taxonomic organism and whether the data are chronic, acute or a chronic estimate will be presented in the graphical outputs
- the software will be written using the R computer code
- it will automatically fit the log-logistic distribution (with two parameters) to data sets that contain toxicity data for five to seven species (that can only be used to derive site-specific trigger values) and fit the best distribution, including the Burr type III distribution to data sets that contain toxicity data for eight or more species; and it will have improved graphical output editability and quality

These changes to the BurrliOZ software will considerably improve the useability of the software and are consistent with changes being made to the Guidelines.

Increasing the sources of data that can be used to derive trigger values

In the current Guidelines, only toxicity data sourced from peer-reviewed scientific journals were used to derive trigger values. However, it was realised that this requirement was being overly restrictive and reducing the amount of data that could be used to derive trigger values. Many data were being unnecessarily excluded, particularly as the Guidelines include data suitability and data quality assessment schemes to determine if data are of appropriate scientific rigour to be used.

It was recommended that any published data (including internal reports and consultancy reports) could be used to calculate trigger values provided that:

- a copy of the document was publically available (if necessary, documents will be hosted on a web-site associated with the revised Guidelines)
- the document could be independently reviewed by an assessor with expertise in trigger value derivation
- the data passed the data suitability and data quality assessment schemes

For many organic chemicals, particularly industrial chemicals, pesticides and pharmaceuticals, there are very limited amounts of public domain toxicity data, yet manufacturers of these chemicals often provide at least the OECD minimum data set (an invertebrate, a fish and a plant) to regulatory authorities in order to determine if the chemicals can be used in Australia. These data are provided on a commercial-in-confidence basis. In order to use this commercial-in-confidence data to derive trigger values, an approach is being investigated whereby agreements will be reached with individual companies for staff of Australian authorities to derive trigger values (using all the accepted methods) and for these to be peer-reviewed on a confidential basis. The resulting trigger values would be presented in the revised Guidelines, but whether the underpinning toxicity data used to derive them will be presented will depend on the individual agreements with the companies. This is in direct conflict with the current standard practise in the Guidelines where all data used to derive trigger values are presented. Despite this lack of transparency, it was considered preferable to not having trigger values for these chemicals.

Improving the assessment of the reliability of trigger values

In the 2000 Guidelines, there were four classifications of the reliability of the trigger values: high reliability, moderate reliability, low interim (LR interim) and low environmental concern level. These classifications were based on two key factors:

- the ecological relevance of the toxicity data (field and chronic laboratory exposures were deemed to be of higher reliability than acute laboratory exposures)
- the number of species and taxa for which there were toxicity data (the greater the number the higher the reliability)

The number and types of toxicity data required for trigger values to achieve each of the four reliability classifications

in the current Guidelines (ANZECC/ARMCANZ 2000a, b) are provided in Table 1. Further details on the types of toxicity data that can be used to derive trigger values can be obtained from Warne (2001).

The current Guideline approach led to trigger values based on widely differing numbers of chronic toxicity data being classified as high reliability, e.g. some trigger values based on chronic toxicity data for 30 species and others based on 5 species. In hindsight, this reliability classification system does not adequately indicate the numbers of data used to derive the trigger values. Consequently, it was recommended that a new reliability classification scheme be developed as part of the revision of the Guidelines. The exact form of the new classification scheme has not yet been resolved but it would be good for it to contain three components to indicate the number of species for which toxicity data were available, whether the data are chronic or acute or a mixture of the two, and how well the SSD fitted the toxicity data. The latter could be indicated by the 95 % CLs for each trigger value (refer to "Incorporating and presenting uncertainty" section).

Increasing flexibility and providing guidance on how to address issues that may arise in the future

As stated earlier, the Guidelines are result-driver, rather than process-driven, thus new methods or information can be used, provided they are scientifically rigorous and defensible. Presently, only 'ecologically relevant' endpoints that measure detrimental effects on populations, communities and ecosystems (e.g. death, immobilisation, growth (individual or population) and reproductive impairment) are used to derive trigger values. No toxicity data that measure effects below the individual level of organisation (e.g. subcellular, biochemical) can be used. However, in light of the rapid expansion of this field of ecotoxicology, such data could in the future be used, provided their ecological relevance can be demonstrated. This would have to be done on a case by case basis. At the other end of the spectrum, the Guidelines promote the use of field ecological data for deriving site-specific trigger values, but provide little specific guidance. Van Dam et al. (2013) provide several examples about how this can be done, and also identify several recent methods that may be useful for using field data to derive trigger values. This is an area that is receiving increasing attention (e.g. Crane et al. 2007; Kwok et al. 2009).

It is well known that many environmental factors can modify the toxicity and bioavailability of chemicals to aquatic organisms. Examples in the current Guidelines include algorithms on how water hardness modifies toxicity for some metals and the effect of pH and temperature on the toxicity of ammonia (ANZECC/ARMCANZ 2000a). The biotic ligand model (BLM) has been applied to the toxicity of copper (e.g. Santore et al. 2001; De Schamphelaere et al. 2002; De Schamphelaere and Janssen 2004; USEPA 2007b), nickel (Keithly et al. 2004; Hoang et al. 2004), silver (Paquin et al. 1999) and zinc (Heijerick et al. 2002a, b); however, there are very few data available to develop quantitative relationships between environmental factors and toxicity, irrespective of their type, that would permit the calculation of site-specific trigger values.

A key development in the future, as has been done with the revised Australian sediment quality guidelines (Simpson et al. 2010) and the Australian ecological investigation levels for contaminated sites (NEPC 2011a, b), should be to develop more of these relationships that can modify the trigger values under varying environmental conditions. Examples of this are the work by Van Dam et al. (2012a, b) on the influence of dissolved organic carbon on uranium toxicity, and the nickel BLM project. In the latter, representatives of Australian regulatory agencies, academics and consultants are working collaboratively with the Nickel Producers Environmental Research Association (NiPERA) and WCA Environment Limited (England) to determine the validity of the nickel BLM, which was initially developed for Europe, to suit Australian conditions and biota.

Increasing international collaboration and harmonisation

The National Water Quality Management Strategy and many of the underlying water quality guidelines (e.g. the Australian and New Zealand Guidelines for Fresh and Marine Water Quality) are examples of international collaboration and harmonisation. While there are always hurdles

Table 1 The minimum data requirements needed for each classification of trigger value reliability (based on ANZECC/ARMCANZ 2000a)

Trigger value reliability classification	Minimum data requirements
High reliability (HR)	Chronic field or laboratory data for at least five species that belong to at least four phyla
Moderate reliability (MR)	Acute laboratory data for at least five species that belong to at least four phyla
Low-reliability interim (LR interim)	Acute or chronic data for at least a fish, an invertebrate and a plant
Low-reliability environmental concern level (LR ECL)	Acute or chronic data for at least one species

to overcome in such ventures (such as agreement of scientific methodologies or developing a system that can be implemented within the regulatory and policy frameworks of all participating jurisdictions), there are substantial benefits to be gained through such a process, both scientifically and financially. Australian and New Zealand scientists have been active in explaining, training and providing expert advice to other jurisdictions developing their own water quality guidelines and in commencing a program to develop water quality guidelines for the Australian Antarctic Territories.

There are clear efforts underway to increase international collaboration and harmonisation of methods to derive water quality guidelines. The ideas in Merrington et al. (2013) represent logical first steps in progressing international collaboration and harmonisation. There is likely to be strong support for such endeavours within the Australian and New Zealand community involved with deriving water quality guidelines. Even without any specific program aimed at increased international harmonisation, this process will continue as regulatory scientists monitor the international literature and developments in other jurisdictions. Two examples of this are the adoption of the SSD approach as the preferred method of WQG derivation by all major organisations involved in WQG derivation, and the influence that the Technical Guidance for Deriving Environmental Quality Standards (EC European Commission 2011) and the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) program of the European Union (ECHA 2007, 2008) are having beyond Europe.

Summary

The current revision of the Australian and New Zealand Guidelines for Fresh and Marine Water Quality will include significant changes to the derivation of trigger values for toxicants. The changes will be evolutionary rather than revolutionary as there have been no major paradigm changes since the current Guidelines were released in 2000. The changes will reflect the latest scientific findings and enhance the risk-based approach which encourages site-specific investigation and the derivation of site-specific trigger values. By increasing the types and sources of data that can be used, more guideline trigger values will be able to be derived using the SSD approach. Working collaboratively and on a confidential basis with industry will permit the use of commercial-in-confidence data to derive trigger values for many chemicals where there are limited public domain data. Increasing the minimum data requirements will increase the accuracy and reliability of the resulting trigger values. Including a measure of the uncertainty of the trigger value will provide more information to end-users and enable more robust decision making. Improving the software used to derive trigger values will improve the users' ability to interpret the results. Providing guidance of measures of toxicity and toxicological endpoints that may, in the near future, be appropriate for trigger value derivation will help to ensure the Guidelines do not become obsolete and instead, keep up with the latest developments. Similarly, having scientific rigour as the main requirement of the suitability of any work, rather than whether it follows a prescriptive method will help future-proof the guidelines. Finally, the Australian and New Zealand Guidelines for Fresh and Marine Water Quality are an outstanding example of the benefits of international collaboration and harmonisation. There is much to be gained through such endeavours and the authors encourage further work in this area.

References

ANZECC (Australian and New Zealand Environment and Conservation Council) (1992) National Water Quality Management Strategy. ANZECC, Canberra, Australia. http://www.environment.gov.au/water/policy-programs/nwqms/. Accessed 10 Jan 2013

ANZECC/ARMCANZ (Australian and New Zealand Environment and Conservation Council and Agriculture and Resource Management Council of Australia and New Zealand) (2000a) National Water Quality Management Strategy, Document 4—Australian and New Zealand Guidelines for Fresh and Marine Water Quality. ANZECC/ARMCANZ, Canberra, Australia. http://www.environment.gov.au/water/publications/quality/nwqms-guidelines-4-vol1.html, http://www.environment.gov.au/water/publications/quality/nwqms-guidelines-4-vol2.html, and http://www.environment.gov.au/water/publications/quality/nwqms-guidelines-4-vol3.html. Accessed 10 Jan 2013

ANZECC/ARMCANZ (Australian and New Zealand Environment and Conservation Council and Agriculture and Resource Management Council of Australia and New Zealand) (2000b) National Water Quality Management Strategy, Document 7—Australian Guidelines for Water Quality Monitoring and Reporting. ANZECC/ARMCANZ, Canberra, Australia. http://www.environment.gov.au/water/publications/quality/nwqms-monitoring-reporting.html. Accessed 10 Jan 2013

Campbell E, Palmer MJ, Shao Q, Wilson D (2000) BurrliOZ: a computer program for calculating toxicant trigger values for the ANZECC and ARMCANZ water quality guidelines. Perth, Western Australia, Australia. http://www.cmis.csiro.au/envir/burrlioz/Download1.htm. Accessed 10 Jan 2013

CCME (Canadian Council of Ministers of the Environment) (2007) A protocol for the derivation of water quality guidelines for the protection of aquatic life 2007. In: Canadian Environmental Quality Guidelines, Canadian Council of Ministers of the Environment, 1999, Winnipeg, MB, Canada. 37pp

Chapman PM, Cardwell RS, Chapman PF (1996) A warning: NOECs are inappropriate for regulatory use. Environ Toxicol Chem 15:77–79

Chapman PM, McDonald B, Kickham PE, McKinnon S (2006) Global geographic differences in marine metals toxicity. Mar Pollut Bull 52:1081–1084

Crane M, Kwok KWH, Wells C, Whitehouse P, Lui GCS (2007) Use of field data to support European Water Framework Directive Quality Standards for dissolved metals. Environ Sci Technol 41:5014–5021

Daam MA, Van den Brink PJ (2010) Implications of differences between temperate and tropical freshwater ecosystems for the ecological risk assessment of pesticides. Ecotoxicology 19:24–37

Davies PE, Cook LSJ, Goenarso D (1994) Sublethal responses to pesticides of several species of Australian freshwater fish and crustaceans and rainbow trout. Environ Toxicol Chem 13:1341–1354

De Schamphelaere KAC, Janssen CR (2004) Development and field validation of a biotic ligand model predicting chronic copper toxicity to *Daphnia magna*. Environ Toxicol Chem 23:1365–1375

De Schamphelaere K, Heijerick D, Janssen C (2002) Refinement and field validation of a biotic ligand model predicting acute copper toxicity to *Daphnia magna*. Comp Biochem Physiol C 133:243–258

Dyer SD, Belanger SE, Carr GJ (1997) An initial evaluation of the use of Euro/North American fish species for tropical effects assessments. Chemosphere 35:2767–2781

EC (European Commission) (2011) Common implementation strategy for the Water Framework Directive (2006/60/EC). Guidance document no. 27. Technical guidance for deriving environmental quality standards. European Commission, Brussels, Belgium, 204pp

ECHA (European Chemical Agency) (2007) Guidance for identification and naming of substances under REACH. European Chemicals Agency, Helsinki, Finland. http://guidance.echa.europa.eu/. Accessed 12 Jan 2013

ECHA (European Chemical Agency) (2008) The guidance on information requirements and chemical safety assessment. Guidance for the implementation of REACH. European Chemical Agency, Helsinki, May 2008

Fox DR (2009) Opinion: NECs, NOECs and the ECx. Australas J Ecotoxicol 14:7–10

Fox DR (2010) An integrated Bayesian approach for determining the no effect and hazardous concentrations in ecotoxicology. Environ Toxicol Chem 73:123–131

Geotechnical Services (2006) Effects of RO brine on the development of giant cuttlefish (*Sepia apama*) embryos. Report ENV06-128. Report prepared for BHP-Billiton. 73pp

Hall JA, Golding LA (1998) Standard methods for whole effluent toxicity testing: development and application. Report no. MFE80205. NIWA report for the Ministry for the Environment, Wellington, New Zealand. pp. 53. http://www.mfe.govt.nz/publications/water/whole-effluent-toxicity-nov98/index.html

Heijerick DG, De Schamphelaere KAC, Janssen CR (2002a) Biotic ligand model development predicting Zn toxicity to the alga *Pseudokirchneriella subcapitata*: possibilities and limitations. Comp Biochem Physiol C 133:207–218

Heijerick DG, De Schamphelaere KAC, Janssen CR (2002b) Predicting acute zinc toxicity for *Daphnia magna* as a function of key water chemistry characteristics: development and validation of a biotic ligand model. Environ Toxicol Chem 21:1309–1315

Hickey CW (2000) Ecotoxicology: laboratory and field approaches. In: Collier KC, Winterbourn M (eds) New Zealand stream invertebrates: ecology and implications for management. New Zealand Limnological Society, Christchurch, New Zealand, pp 313–343

Hickey CW, Martin ML (1995) Relative sensitivity of five benthic invertebrate species to reference toxicants and resin acid contaminated sediments. Environ Toxicol Chem 14:1401–1409

Hoang TC, Tomasso JR, Klaine SJ (2004) Influence of water quality and age on nickel toxicity to fathead minnows (*Pimephales promelas*). Environ Toxicol Chem 23:86–92

Hobbs DA, Warne MStJ, Markich SJ (2004) Utility of northern hemisphere metal toxicity data in Australasia. SETAC Globe 5:38–39

Hoekstra JA, Van Ewijk PH (1993) Alternatives for the no-observed effect level. Environ Toxicol Chem 12:187–194

Hose GC, Van den Brink PJ (2004) Confirming the species sensitivity distribution concept for endosulfan using laboratory, mesocosm and field data. Arch Environ Contam Toxicol 47:511–520

Jager T (2012) Bad habits die hard: the NOEC's persistence reflects poorly on ecotoxicology. Environ Toxicol Chem 31:228–229

Johnston N, Skidmore J, Thompson GB (1990) Applicability of OECD data to Australian aquatic species. A report to the Advisory Committee on Chemicals in the Environment. Australian and New Zealand Environment and Conservation Council, Canberra, Australia

JSC (Joint Steering Committee for the Revision of the National Water Quality Management Strategy) (2010) Frequently asked questions. http://www.scew.gov.au/archive/water/pubs/wq_fmwq_revision_of_guidelines_faq_final_230910.pdf. Accessed 10 Jan 2013

Keithly J, Brooker JA, DeForest DK, Wu BK, Brix KV (2004) Acute and chronic toxicity of nickel to a cladoceran (*Ceriodaphnia dubia*) and an amphipod (*Hyalella azteca*). Environ Toxicol Chem 23:691–696

Kwok KWH, Leung KMY, Chu VKH, Lam PKS, Morritt D, Maltby L, Brock TCM, Van den Brink PJ, Warne MStJ, Crane M (2007) Comparison of tropical and temperate freshwater species sensitivities to chemicals: implications for deriving safe extrapolation factors. Integr Environ Assess Manag 3:49–67

Kwok KWH, Bjorgesaeter A, Leung KMY, Lui GCS, Gray JS, Shin PKS, Lam PKS (2009) Deriving site-specific sediment quality guidelines for Hong Kong marine environments using field-based species sensitivity distributions. Environ Toxicol Chem 27:226–234

Landis WG, Chapman PM (2011) Well past time to stop using NOELs and LOELs. Integrated Environmental Assessment and Management 7, vi–viii

Maltby L, Blake N, Brock TCM, Van den Brink PJ (2003) Addressing interspecific variation in sensitivity and the potential to reduce this source of uncertainty in ecotoxicological assessments. Science and Research Report PN0932. Department for Environment, Food and Rural Affairs, London, UK. www.defra.gov.uk/science. Accessed 15 Jan 2013

Markich SJ, Camilleri C (1997) Investigation of metal toxicity to tropical biota: Recommendations for revision of the Australian water quality guidelines. Supervising Scientist Report SSR 127, Supervising Scientist, Canberra, Australia

Merrington, Youn-Joo An, Grist EPM, Jeong S-W, Rattikansukha C, Roe S, Schneider U, Sthiannopkao S, Suter GW II, Van Dam R, Van Sprang P, Wang J-Y, Warne MStJ, Yillia PT, Zhang X-W, Leung KMY (2013) Water quality guidelines for chemicals: learning lessons to deliver meaningful environmental metrics. Env Sci Poll Res. doi:10.1007/s11356-013-1732-8

Mulhall A (1997) Models to predict the toxicity of selected phenols and benzamines to a cladoceran and a marine bacterium. Honours thesis, University of Technology Sydney, Australia, 112 pp

NEPC (National Environment Protection Council) (2011a) National Environment Protection (Assessment of Site Contamination) Measure. Schedule B5b—Guideline on methodology to derive Ecological Investigation Levels in contaminated soils. NEPC, Adelaide, Australia. 87 pp. http://www.ephc.gov.au/sites/default/files/Schedule_B5b__Guideline_on_methodology_to_derive_EILs__SEP10.pdf. Accessed 12 Jan 2013

NEPC (National Environment Protection Council) (2011b) National Environment Protection (Assessment of Site Contamination) Measure. Schedule B5c—Guideline on soil quality guidelines for arsenic, chromium III, copper, DDT, lead, naphthalene, nickel and zinc. NEPC, Adelaide, Australia. 177pp. http://www.ephc.gov.au/sites/default/files/Schedule_B5c__Guideline_on_soil_quality_guidelines__SEP10.pdf. Accessed 12 Jan 2013

Newman MC (2008) What exactly are you inferring? A closer look at hypothesis testing. Environ Toxicol Chem 27:1013–1019

Newman MC, Ownby DR, Mezin LCA, Powell DC, Christensen T, Lerberg SB, Anderson BA (2000) Applying species-sensitivity distribution in ecological risk assessment: assumptions of distribution type and sufficient numbers of species. Environ Toxicol Chem 19:508–515

Noppert F, Van der Hoeven N, Leopold A (1994) How to measure no effect? Towards a new measure of chronic toxicity in ecotoxicology. Netherlands Working Group on Statistics and Ecotoxicology, Delft, The Netherlands

OECD (2006) Current approaches in the statistical analysis of ecotoxicity data: a guide to application. OECD Environmental Health and Safety Publication, Series on Testing and Assessment, No. 54, Environment Directorate, ENV/JM/MONO(2006)18, Paris, France

Paquin PR, Di Toro DM, Santore RC, Trivedi D, Wu KB (1999) A biotic ligand model of the acute toxicity of metals: III. Application to fish and *Daphnia magna* exposure to silver. US Government Printing Office, Washington DC, 1999. EPA 822-E-99-001

Phyu YL (2004) Assessment of toxicity, bioavailability, partitioning and hazard of the herbicides, atrazine and molinate. PhD thesis. University of Technology Sydney, Australia, 281pp

Rombke J, Waichman AV, Garcia MVB (2008) Risk assessment of pesticides for soils of the central Amazon, Brazil: comparing outcomes with temperate and tropical data. Integr Environ Assess Manag 4:94–104

Rose RM, Warne MStJ, Lim RP (1998) Quantitative structure–activity relationships and volume fraction analysis for nonpolar narcotic chemicals to the Australian cladoceran *Ceriodaphnia* cf. *dubia*. Arch Environ Contam Toxicol 34:248–252

Sanchez-Bayo F, Hyne RV (2011) Comparison of environmental risks of pesticides between tropical and non-tropical regions. Integr Environ Assess Manag 7:577–586

Santore RC, Di Toro DM, Paquin PR, Allen HE, Meyer JS (2001) A biotic ligand model of the acute toxicity of metals. II. Application to acute copper toxicity in freshwater fish and daphnia. Environ Toxicol Chem 20:2397–2402

Shao Q (2000) Estimation for hazardous concentrations based on NOEC toxicity data: an alternative approach. Environmetrics 11:583–595

Shao Q, Warne MStJ (2002) Critical assessment of current water quality guidelines and potential further development. In 'Interact 2002—Program and Abstract Book. Sydney, 21–25 July, 2002', p. 158

Simpson SL, Batley GE, Chariton AA (2010) Revision of the ANZECC/ARMCANZ Sediment Quality Guidelines. CSIRO Land and Water Science Report 08/07. August 2008, revised July 2010. Report prepared for the Department of Environment, Water, Heritage and the Arts. 115pp

Sinclair A, Tayler K, van Dam R, Hogan A (2013) 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. Environ Sci Pollut Res. doi:10.1007/s11356-013-1922-4

Sunderam RIM, Cheng DMH, Thompson GB (1992) Toxicity of endosulfan to native and introduced fish in Australia. Environ Toxicol Chem 11:1469–1476

USEPA (2007a) Water Quality Standards Handbook. 2nd Ed. EPA-823-B-94-0059. 2007 Update. United States Environmental Protection Agency, Washington DC, USA

USEPA (2007b) Aquatic life ambient freshwater quality criteria—copper. 2007 Revision. No. EPA-822-R-07-001. United States Environmental Protection Agency, Criteria and Standards Division, Washington D.C

USEPA (U.S. Environmental Protection Agency) (1999) Water quality guidance for the Great Lakes system, Federal Register, 40 CFR Part 142, July 1. USEPA, Washington, DC, USA

Van Dam RA, Chapman JC (2001) Direct toxicity assessment (DTA) for water quality guidelines in Australia and New Zealand. Australas J Ecotoxicol 7:175–198

Van Dam RA, Harford AJ, Houston MA, Hogan AC, Negri A (2008) Tropical marine toxicity testing in Australia: a review and recommendations. Australas J Ecotoxicol 14:55–88

Van Dam R, Harford A, Warne MStJ (2012a) Time to get off the fence: the need for definitive international guidance for statistical analysis of ecotoxicity data. Integr Environ Assess Manag 8:242–245

Van Dam RA, Trenfield MA, Markich SJ, Harford AJ, Humphrey AC, Stauber JL (2012b) Re-analysis of uranium toxicity data for freshwater organisms and the influence of dissolved organic carbon. Environ Toxicol Chem 31:2606–2614

Van Dam RA, Humphrey CL, Harford AJ, Frostick A, Jones DR, Davies S, Storey AW (2013) Site-specific water quality guidelines: 1. Derivation approaches based on physicochemical, ecotoxicological and ecological data. Environ Sci Pollut Res. doi:10.1007/s11356-013-1780-0

Van der Hoeven N (1997) How to measure no effect. Part 1: towards a new measure of chronic toxicity? Environmetrics 8:241–248

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, Canberra, ACT, Australia. ISBN 0 642 24338 7. 82pp

Warne MStJ (2001) Derivation of the ANZECC and ARMCANZ water quality guidelines for toxicants. Australas J Ecotoxicol 7:123–136

Warne MStJ (2010) A refined assessment of the selection of species and other factors that affect dilution factors for the proposed desalination plant at Point Lowly, South Australia. CSIRO Land and Water Science Report 07/10. CSIRO, Adelaide, South Australia, 43pp. http://www.bhpbilliton.com/home/aboutus/regulatory/Documents/Olympic%20Dam%20Supplementary%20EIS/Appendices/Appendix%20H4_Ecotoxicology.pdf. Accessed 11 May 2013

Warne MStJ, Van Dam R (2008) NOEC and LOEC data should no longer be generated or used. Australas J Ecotoxicol 14:1–5

Westbury A-M, Warne MStJ, Lim RP (2004) Toxicity of substituted phenols to *Ceriodaphnia* cf. *dubia* and *Vibrio fischeri* and the development of predictive models. Australas J Ecotoxicol 10:33–42

Wheeler JR, Grist EPM, Leung KMY, Morritt D, Crane M (2002) Species sensitivity distributions: data and model choice. Mar Pollut Bull 45:192–202