DERIVATION OF THE AUSTRALIAN AND NEW ZEALAND WATER QUALITY GUIDELINES FOR TOXICANTS

Michael St. J. Warne
Ecotoxicology Section, New South Wales Environment Protection Authority (EPA) located at the EPA/UTS Centre for Ecotoxicology, University of Technology Sydney (UTS), Westbourne St, Gore Hill, NSW 2065, Australia.

ABSTRACT

A new framework was developed and used to derive the Australian and New Zealand (ANZECC & ARMCANZ 2000) water quality guidelines for metal, inorganic and organic toxicants. This framework incorporates the principles of risk and recent advances in ecotoxicity modelling. As such, it is a significant advance on the previous ANZECC (1992) and other international water quality guidelines. Two different methods were used to derive the guidelines: a modification of the Canadian (CCME 1991) assessment factor method, and a new statistical distribution method called the Burr Type III method (Campbell *et al.* 2000) which was developed from the Aldenberg and Slob (1993) method. This paper provides the rules governing the framework, details of the collation and screening of the toxicity data and the methods used to derive the guidelines.

Key words: toxicants, trigger values, statistical distribution, assessment factor.

INTRODUCTION

The National Water Quality Management Strategy of Australia (ANZECC & ARMCANZ 1994) provides for the derivation and periodic review of the water quality guidelines (WQGs). The previous Australian water quality guidelines (ANZECC 1992) were published in 1992 and therefore required review of both the guideline values and the methods used to derive them. As part of this review, the author was commissioned by the Environmental Research Institute of the Supervising Scientist of Australia to review the current methods of deriving WQGs for toxicants and determine the most appropriate toxicity data to use for this purpose. The review (Warne 1998) examined the merits of the assessment factor (AF) and the more recently developed statistical distribution (SD) approaches in terms of: the assumptions of each approach; the magnitude of the assessment factors; whether or not the approaches provided the stated degree of environmental protection; how well they adhered to the Precautionary Principle; and how conservative were the guidelines they derived. The review led to the development and approval by ANZECC, of a framework for deriving toxicant trigger values (TVs) (Warne 1998). The key features of this framework were the use of both a statistical distribution and an assessment factor method to calculate the TVs and a risk based approach (Warne 1998).

The framework used to derive the toxicant TVs in the earlier Draft ANZECC & ARMCANZ water quality guidelines (1999), was that recommended by Warne (1998). The framework used in ANZECC & ARMCANZ (2000) was modified in order to reflect the advances that occurred in the three-year period between commencing the review in 1996 and the completion of the public consultation phase in 1999. While details of the framework used in ANZECC & ARMCANZ (2000) were different from that recommended by Warne (1998), it retained the key features described above. This paper describes the framework that was used to derive the TVs in the new water quality guidelines (ANZECC & ARMCANZ 2000).

OVERVIEW OF THE NEW FRAMEWORK

In the previous ANZECC WQGs (ANZECC 1992) the environmentally safe levels were termed guidelines, whereas in the new guidelines (ANZECC & ARMCANZ 2000) they are termed Trigger Values (TVs). This term is used because, if the TVs are exceeded, it triggers one or more of the following: further investigation, development and implementation of management strategies, or remediation. This concept is discussed in broad terms by McAlpine and Humphrey (2001) and in greater detail by Chapman (2001).

The new framework is hierarchical, having three grades of TVs - high, moderate and low reliability (HR, MR and LR respectively). Within the LR TVs there are two types of TVs, the interim (LR (interim) TV) and environmental concern level (LR (ECL) TV). The different grades of TVs reflect the certainty that they would provide adequate environmental protection, which in turn was related to the quantity, type and representativeness of the available toxicity data from which they were derived.

The highest possible grade of TV for which there were adequate, suitable toxicity data, was derived. If there were insufficient suitable data to derive a HR TV, then the hierarchy was descended until the available data met the minimum data requirements for a particular grade of TV.

The framework used two different methods to calculate TVs. The preferred method was the Burr Type III statistical distribution (BT III SD) method developed by Shao (2000) which was based on the Aldenberg and Slob (A&S) (1993) SD method. The alternative method was a modification of the Canadian (CCME 1991) AF method, developed by Warne (1998), which is henceforth referred to as the ANZECC & ARMCANZ AF method.

Whenever the available toxicity data permitted, TVs were calculated for both marine and freshwater environments. However, this was not always possible. In such cases, the TVs for organic and inorganic chemicals (excluding metals) derived for one medium were adopted for the other medium but the grade was reduced to LR. For example, a freshwater MR TV would be adopted as a marine LR TV. This was done because the chemistry and toxicity of organic chemicals is not greatly affected by the change in salinity. For metals, TVs were not adopted from one medium to another because their chemistry and hence toxicity could be markedly different under saline and fresh conditions.

The new ANZECC & ARMCANZ (2000) framework is similar to the USEPA (1986), OECD (1992), Danish (Petersen and Pedersen 1995), South African (Roux *et al.* 1996) and The Netherlands' (Van de Plassche *et al.* 1993) frameworks in that it preferred SD methods to the AF methods. It differed from some overseas frameworks in three ways. Firstly, it used a new SD method. Secondly, the toxicity of mixtures was taken into account, which is not done in any of the overseas frameworks. How this issue was dealt with is described in Chapman (2001). Thirdly, the potential for secondary poisoning is accounted for but not in a direct manner, as is done by the USEPA and The Netherlands.

THE NEW FRAMEWORK

Collecting toxicity and physicochemical data

Acute, chronic, laboratory, field, mesocosm and microcosm toxicity data were obtained by conducting searches of the USEPA (1994) AQUIRE database, the Australasian Ecotoxicology Database (Warne *et al.* 1998; Warne and Westbury 1999; Markich *et al. in press*), the in-house literature collection of the Ecotoxicology Section of the New South Wales Environment Protection Authority and various water quality documents of Canada, Denmark, The Netherlands, The United Kingdom and the USA. Additional searches of abstracting services such as Cambridge Abstract Service, Biological Abstracts and Pollution Abstracts were conducted for meso- and micro-cosm toxicity data. As a general rule, toxicity data published prior to 1980 were not included, as they were considered to be unreliable due to advances in experimental and analytical capabilities since that time (Warne 1998). All TV values, except for the LR TVs of non-polar narcotic chemicals, were derived using only toxicity data that were obtained from the above sources. Low reliability TVs for non-polar organics also used data generated by quantitative structure-activity relationships (QSARs) that are described in a following section.

Wherever possible, the following information was obtained for every chemical for which a TV was derived: Chemical Abstract Services number (CAS no.), IUPAC name, common name, aqueous solubility, boiling and melting point, chemical formula, half-life in water and sediment, molecular weight, octanol-water partition coefficient, bioconcentration factor, specific gravity and vapour pressure. This information was obtained from sources such as Hansch *et al.* (1995), Mackay *et al.* (1992a and b, 1993, 1995), Shiu *et al.* (1994), Tomlin (1994), Verscheuren (1983) and Weast (1987).

Screening the toxicity data

The quality of the toxicity data obtained from AQUIRE had already been assessed (USEPA 1994). This assessment examined how the toxicity data were generated and a score was awarded on the basis of answers to a series of questions similar to those presented in Table 1. Toxicity data were classed as: complete (C) with a score between 85 and 100, moderate (M) with a score of 51-84 or incomplete (I) with a score of 50 or less. Only complete and moderate quality data were used to derive the TVs.

The quality of toxicity data used by the Dutch and Danish (eg. Petersen and Pedersen 1995; RIVM 1995) had also been assessed. Both countries only used data that passed their quality system to derive their WQGs. Such data were assumed to be equivalent to the C and M classes of AQUIRE and were therefore also used to derive the new TVs. The quality of all other toxicity data (eg. that obtained from journal articles) was assessed using a system based on the AQUIRE method (Table 1) and classed as complete, moderate or incomplete using the USEPA ranges stated previously. It is worth noting that the quality assessment scheme used to derive the new TVs has subsequently been substantially improved and used in the published version of the Australasian Ecotoxicology Database (Warne *et al.* 1998; Warne and Westbury 1999; Markich *et al. in press*).

The above system for assessing the quality of toxicity data was not suitable for either data generated by quantitative structure-activity relationships (QSARs) or for multiple species (MS) toxicity data generated in meso- and/or micro-cosms. The quality of the former was not assessed as the quality of the QSARs had already been assessed (refer to the section on QSARs for details). The quality of the latter was assessed using the following rules that were based on those recommended by the OECD (1992). In order for MS data to have been used to derive TVs, the mesocosms and/or microcosms:

1. should have included fish or shellfish or the endpoints measured should have been directly relate to these species;
2. must have represented the basic properties of ecosystems including photosynthesis, nutrient cycling and trophic structure;
3. should have had at least three dose treatments and a suitable control and all treatments should have been at least duplicated;
4. should have measured the chemical and physical properties that can affect exposure to the toxicant or the bioavailability;
5. should have covered individual, population and community level biological endpoints; and
6. tests should have been of sufficient duration to account for the life-history of the organisms and the fate of the toxicant.

The most common failings of the MS toxicity data that were examined were that they had insufficient treatments, replication and/or controls.

Once the quality of the toxicity data had been determined and all incomplete quality data removed, the remainder were screened further using a number of other variables that are presented in Table 2. Data that had these characteristics were not used to derive TVs. The toxicological endpoints of the data used to derive the TVs were limited to those that were considered to have ecological relevance (Table 2) (Warne 1998).

Only chronic no observed effect concentration (NOEC) data were used to derive HR TVs while only acute fifty percent effect data (LC50 and EC50) were used to derive MR TVs. Thus, it was necessary to define these terms and determine whether each datum was chronic or acute. Chronic exposures was defined for multi-celled organisms, as being greater than 96 hours, while for single-celled organisms it was defined as being equal to or greater than 72 hours. Thus, data classified as chronic in this study contained data that would normally be classified as chronic and sub-chronic. Acute exposure was defined as being greater than 24 hours but shorter than the duration for chronic data. Data based on exposures of less than 24 hours were not used to derive TVs.

An ACCESS® database (Sunderam *et al.* 2000) containing all the toxicity data used to derive the TVs and the physicochemical data for every chemical that a TV was calculated for, is supplied on a CD-ROM as part of the ANZECC & ARMCANZ (2000) water quality guidelines. The database will be integral to the implementation of the toxicant TVs, particularly if site-specific investigations are conducted or site-specific TVs are derived. How these data can be used for these purposes is described in Chapman *et al.* (2001).

Table 1. The questions used and the marks awarded to determine the quality score and quality class of the toxicity data. Modified from the USEPA (1994).

Question	Possible Marks[1]
Was the duration of the exposure stated?	20 or 0
Were there appropriate controls (eg. a solvent control if solvents are used)?	5 or 0
Were the characteristics of the test organism stated?	5 or 0
Were the chemical concentrations measured?	5 or 0
Was the type of exposure (eg. static, flow through) stated?	5 or 0
Was the test location stated?	4 or 0
Was the grade or purity of the test chemical stated?	4 or 0
Was the type of test media used stated?	4 or 0
Was the hardness (for freshwater) or the salinity (for saltwater) measured and stated?	2 or 0
Was the alkalinity (for freshwater) or salinity (for saltwater) measured and stated?	2 or 0
Was the dissolved oxygen content of the test water measured at some stage during or after the test?	2 or 0
Was the temperature measured during the test?	2 or 0
Was the pH of the test water measured at some time during the test?	2 or 0
Was the biological endpoint clearly defined?	20 or 0
Was there a concentration-response relationship either observable or stated?	5 or 0
Was the biological effect quantified ie. 50% effect, 25% effect?	5 or 0
Was the statistical level of significance for any statistical tests stated (for NOEC/LOEC data)? Was a valid model used to derive the LC50/EC50 values (for LC/EC data)?	4 or 0
Was the stated significance level 0.05 or less (for NOEC/LOEC data)? Was there an estimate of the variability of the LC50 or EC50 (for LC/EC data)?	4 or 0
Total Score	
Class (C, M, I)	

[1] There are only two marks that can be awarded in answering a question – the full mark or zero.

Table 2. The types of toxicity data that were excluded from the calculation of trigger values and environmental concern levels.

Type of Variable	Conditions excluded
For all Chemicals	
Concentration	ranges, >, ≥, <, and ≤ values
Experimental design	Where the test concentrations differed by a large amount (eg. 10 or greater)
Duration of exposure	not stated and/or < 24 hours duration
Toxicological endpoint	not stated and/or endpoints other than lethality, immobilisation, growth, population growth, and reproduction or the equivalent
Measure of toxicity	not stated and/or measures of toxicity other than 50% effects, NOEC, LOEC and MATC values
Aqueous solubility	values greater than twice the aqueous solubility
For Metals in Freshwater	
pH	Not stated and/or the pH was outside the range of 6.5 to 9
Water hardness	Not stated and/or varying considerably during the test

Table 3. Quantitative structure-activity relationships (QSARs) used to generate toxicity data for non-polar narcotics (from Van Leeuwen et al. 1992).

Species	QSAR (NOEC expressed as mol/L)
Bacteria	
Clostridium botulinum	Log NOEC = -0.82 log Kow - 0.29
Bacillus subtilis	Log NOEC = -0.64 log Kow - 2.03
Pseudomonas putida	Log NOEC = -0.64 log Kow - 1.60
Vibrio fischeri (previously called *Photobacterium phosphoreum*)	Log NOEC = -0.68 log Kow - 1.52
Algae	
Skeletonema costacum	Log NOEC = -0.72 log Kow - 1.42
Scenedesmus subspicatus	Log NOEC = -0.86 log Kow - 1.41
Pseudokirchneriella subcapitata (previously called *Selenastrum capricornutum*)	Log NOEC = -1.00 log Kow - 1.71
Fungi	
Saccharomyces cerevisiae	Log NOEC = -0.78 log Kow - 0.35
Protozoans	
Tetrahymena pyriformis	Log NOEC = -0.80 log Kow - 1.28
Coelenterates	
Hydra oligactis	Log NOEC = -0.86 log Kow - 2.05
Molluscs	
Lymnaea stagnalis	Log NOEC = -0.86 log Kow - 2.08
Arthropods	
Nitocra spinipes	Log NOEC = -0.78 log Kow - 2.14
Daphnia magna	Log NOEC = -1.04 log Kow - 1.70
Aedes aegypti	Log NOEC = -1.09 log Kow - 1.36
Culex pipiens	Log NOEC = -0.86 log Kow - 1.98
Fish	
Pimephales promelas and *Brachydanio rerio*	Log NOEC = -0.87 log Kow - 2.35
Amphibia	
Ambystoma mexicanum	Log NOEC = -0.88 log Kow - 1.89
Rana temporaria	Log NOEC = -1.09 log Kow - 1.47
Xenopus laevis	Log NOEC = -0.90 log Kow - 1.79

Estimation of the chronic toxicity of non-polar narcotic chemicals using Quantitative Structure-Activity Relationships (QSARs)

Non-polar narcotics are chemicals, such as alkanes, alkenes and alkyl and halogen substituted benzenes, that exert their toxicity in a non-specific, reversible manner and are the least toxic group of chemicals (eg. Warne *et al.* 1991; OECD 1995). The chemical characteristics of these chemicals have been specified (OECD 1992, 1995; Verhaar *et al.* 1996). If there were insufficient toxicity data to derive either HR or MRTVs for a non-polar narcotic, then the QSARs presented in Table 3 were used to estimate the chronic toxicity to 19 species. Only these QSARs were used as: they predict chronic toxicity; their quality had been rigorously assessed (Van Leeuwen *et al.* 1992); and they were recommended for use by the OECD (1995) and the Dutch (Van de Plassche *et al.* 1993). The limitations and strengths of QSARs are discussed in Warne (1998).

In order to use the QSARs the logarithm of the octanol-water partition coefficient (log Kow) for each non-polar narcotic was substituted into the QSARs. The units of the resulting data were converted to µg/L, and then

the data were inserted into the Burr Type III (BT III) SD method (described later in this paper) in order to calculate the TVs. In this method only one toxicity value is used to represent each species. Therefore, if there were more than one experimental toxicity value for a species the data were manipulated according to the rules set out in the section of this paper describing the BT III SD method. Toxicity values predicted by the QSARs were then combined with manipulated experimental values, except when they were for the same species. In such cases, the manipulated experimental data were used in preference to the QSAR predicted values. The combined toxicity data were then entered into the BT III SD method. The result was then divided by an AF of 10 in order to obtain the TV. This is done to account for the fact that the data are estimates of chronic toxicity. Greater information on the BT III SD method is provided later in this paper. Despite having toxicity data for at least 19 species for each chemical, only LR (interim) TVs were calculated, as the bulk of the data were estimates of chronic toxicity.

The Calculation of Acute to Chronic Ratios

Acute to chronic ratios (ACRs) are the ratio of the acute toxicity to the chronic toxicity data for a particular chemical and were calculated using the following formula:

$$ACR = \text{acute toxicity/chronic toxicity} \quad (3)$$

The acute and chronic data did not have to have the same measure of toxicity or endpoint but, they must be for the same species, and have been presented in the same paper or at least determined in the same laboratory. ACRs were either calculated directly from the toxicity data collated for this project or were obtained from a USEPA compilation (Thursby, *pers. comm.*).

The ACRs were used in the framework to provide estimates of chronic toxicity when there were only acute toxicity data available. They were used by both the BT III SD and the ANZECC & ARMCANZ AF methods to calculate TVs. Details of how they were used is provided in the sections of this paper describing those two methods. There are however, some limitations to the use of ACRs, which are discussed by Warne (1998).

Conversion of chronic toxicity data to chronic NOEC values

Generally, once the metal toxicity data had been screened, limited data remained to derive TVs. Most of the remaining data were chronic measures of toxicity (eg. LC50 and LOEC) other than the chronic NOEC data required to derive HR TVs. In order to overcome this problem, the chronic non-NOEC data were converted to chronic NOEC values using a series of conversion factors. The chronic LC50 or EC50, LOEC and the maximum acceptable toxicant concentration (MATC) (where the MATC is the geometric mean of the NOEC and LOEC) data were divided by 5, 2.5 and 2 respectively. A similar procedure was used by the Dutch (eg. Van de Plassche *et al.* 1993), except that they used larger conversion factors. The conversion factors used were the expert opinions of the author and Dr John Chapman (NSW EPA) and were based on examining the data collated to derive the TVs.

This procedure was only used in the derivation of the TVs for metals and metalloids. It was not necessary for the organic and inorganic chemicals as the chronic data for such chemicals were predominantly NOECs.

Correcting toxicity data for water hardness

The toxicity of cadmium, chromium (III), copper, lead, nickel and zinc is affected by water hardness (ie. the aqueous concentration of Ca and Mg ions). Therefore, this modifying factor was considered when the TVs for these metals were derived. Prior to calculating the TVs for these metals, all the toxicity data were modified to a standard water hardness of 30 mg $CaCO_3$/L (Markich *et al.* 2001). The hardness corrected metal toxicity data were then used to derive TVs using either the BT III SD or the ANZECC & ARMCANZ AF methods that are described later in this paper.

DETERMINING THE GRADE OF TV THAT CAN BE DERIVED

The new ANZECC & ARMCANZ (2000) framework is hierarchical and the highest possible grade of TV that could be derived, was. A number of rules, presented below, govern this hierarchical framework. These rules were only applied to those data that passed the screening process, described earlier.

HR TVs can be derived using two different types of chronic toxicity data and by two different methods. The preferred data to use was multiple species (MS) toxicity data followed by laboratory based single species (SS) toxicity data. Thus, if the MS data met the quality and minimum data requirements then they were used in preference to the SS data. Similarly, provided the data met the rules set out below, the BT III SD method was used in preference to the ANZECC & ARMCANZ AF method. The following rules applied to the MS data:

1. If the data met the minimum data requirements (Tables 4 and 5) of the BT III SD method, then that method was used to derive the TV. If the data did not meet the minimum data requirements of the BT III method then rule 2 was applied.
2. If the data met the minimum data requirements (Table 6) of the ANZECC & ARMCANZ AF method, then that method was used to derive the TV. If the data did not meet the minimum data requirements then rule 3 was applied.
3. If a MS HR TV could not be derived, it was determined whether a SS HR TV could be derived.

The same rules were applied to the derivation of a SS HR TV, with the exception that rule 3 was modified so that if a SS HR TV could not be derived then the data were examined in order to determine whether a MR TV could be derived.

Rules 1 and 2 for the MS data were also applied to the derivation of MR TVs. Rule 3 was, however, modified so that if a MR TV could not be derived, then the data were examined to determine whether a LR (interim) TV could be derived.

When deriving LR (interim) TVs the following rules were applied:

1. If the chemical was a non-polar narcotic (as defined by the OECD 1992, 1995; Verhaar *et al.* 1996) then the QSAR derived data as well as any manipulated experimental toxicity data were used by the BT III SD method to derive the TV.
2. If the chemical was not a non-polar narcotic and met the minimum data requirements of the ANZECC & ARMCANZ AF method (Table 6) then that method was used to derive the TV.
3. If a LR (interim) TV could not be derived then a LR (ECL) TV was derived using the ANZECC & ARMCANZ AF method (Table 6).

METHODS USED TO DERIVE TRIGGER VALUES

Two different methods were used to derive the TVs. These were the Burr Type III statistical distribution (BT III SD) and the ANZECC & ARMCANZ AF methods. Details of these methods are provided below. In addition, information is provided on the use of ACRs, which were used by both methods to derive TVs.

The Burr Type III Statistical Distribution Method

Background

Warne (1998) reviewed the three SD methods (Stephan *et al.* 1985; Aldenberg and Slob 1993; Wagner and Løkke 1991) used by regulatory authorities to derive WQGs. He recommended that the Aldenberg and Slob (A&S) method be used to derive the ANZECC & ARMCANZ WQGs. This method was used to derive the TVs in the earlier Draft ANZECC & ARMCANZ guidelines (1999).

The A&S method has several assumptions, the validity of which were discussed in Warne (1998). The most important assumption is that the sensitivity of species to toxicants has a log-logistic distribution. This

Table 4. *The minimum data required by the Burr Type III statistical distribution method for the three grades of Trigger Values (based on Aldenberg and Slob 1993).*

Level of Trigger Value	Minimum data requirements
HR	requires chronic NOEC toxicity data for five different species that belong to at least four different taxonomic groups (see Table 5).
MR	requires acute toxicity data for five different species that belong to at least four different taxonomic groups (see Table 5).
Interim (for non-polar narcotic chemicals only)	requires nineteen estimates of chronic toxicity derived by QSARs (see Table 3).

Table 5. *Types of taxonomically different organisms and the major subdivisions of organisms these belong to.*

Major subdivisions of organisms	Types of organisms that are considered as being taxonomically different
Fish	Fish
Invertebrates	Crustaceans, insects, molluscs, annelids, echinoderms, rotifers, hydra
Plants	Green algae, blue algae, red algae, macrophytes
Others	Blue-green algae (cyanobacteria), amphibians, bacteria (excluding *Vibrio fischeri*[1]), protozoans, coral, fungi

[1] This species was excluded as the endpoint is biochemical (refer to Table 2).

distribution is very similar to the log-normal, however it has fewer individuals near the median and has more individuals in the tails (Warne 1996). During the derivation of the TVs in the earlier Draft guidelines (ANZECC & ARMCANZ 1999) it was found that while the toxicity data for many chemicals met the minimum data requirements of the A&S method, the method could not be used as the data did not have a log-logistic distribution. In such cases the A&S method could not be used and the less preferred AF method had to be used to derive the TVs.

There is no theoretical reason why the distribution of toxicant sensitivity should be log-logistic (eg. Forbes and Forbes 1993). In fact, Aldenberg and Slob (1993) stated that the log-logistic distribution was chosen because it has 'some nice mathematical features that make certain calculations relatively easy'. Shao (2000) noted that the log-logistic distribution belonged to a family of distributions called Burr Type III (BT III). The variety of shapes that BT III distributions can have is large (Shao 2000). Thus, attempting to fit a BT III distribution to any given toxicity data set has a greater probability of success, than attempting to fit only the log-logistic distribution.

The BT III SD method uses a maximum likelihood method to determine which particular statistical distribution best fits the toxicity data for a particular chemical. This method is guaranteed to fit a statistical distribution to the toxicity data at least as well as the A&S method because the log-logistic distribution is a BT III distribution (Shao 2000). Therefore, if another distribution could not be found that fitted the data better than the log-logistic, then the log-logistic by default, fitted the data the best. Greater detail of the BT III method is provided in Shao (2000).

An additional difference between the BT III and A&S SD methods is the term used to describe the values that are calculated. The A&S method calculates the concentration that should be hazardous (HCx) to 'x' percentage of species in an ecosystem. Whereas, ANZECC & ARMCANZ (2000) refer to the concentration that should protect 'x' percentage of species (PCx). Thus, PCx = HC100-x.

The BT III SD method uses a maximum likelihood method to determine which particular statistical distribution best fits the cumulative frequency plot of toxicity data for a chemical. The maximum likelihood method also estimates the parameters that mathematically describe the selected distribution. Because the equation that describes the selected distribution is known, it is very simple to calculate the concentration that should theoretically protect any chosen percentage of species. To do this the cumulative frequency that corresponds to the percentage of species to be protected is entered into the equation for the distribution that best fitted the toxicity data. Thus, the 5th percentile of the selected distribution becomes the concentration that if not exceeded will protect 95% of species and the 10th percentile will protect 90% of species.

As PC values were derived using a sample of species in the environment to be protected rather than all species then, depending on the species that comprise a particular sample, a range of different estimates of PC values for the same chemical could be obtained (Figure 1). Aldenberg and Slob (1993) overcame this problem by developing two confidence limits, 95% and 50%. These indicate the degree of certainty that the calculated trigger value will protect the selected percentage of species. Thus, a PC95 50% value means that there is a 50% certainty that the concentration will protect at least 95% of species in an ecosystem, or alternatively, this can be expressed as, 50% of PC95 50% values will protect at least 95% of species. However, in the ANZECC & ARMCANZ (2000) framework only the 50% confidence limit was used as it was considered more statistically robust than the 95% (Fox 1999).

Strictly speaking the PC values calculated by the BT III SD method and reported in the Guidelines (ANZECC & ARMCANZ 2000) do not have confidence intervals. This was not viewed as necessary, because if there is a large sample size, the chosen percentile will approximate the median of estimates of the PC value. Thus, the 5th percentile should equal the PC95 50%. Despite PC values with confidence intervals (eg. PC 95 95%) not being used in the Guidelines (ANZECC & ARMCANZ 2000) the BurrliOZ software can calculate confidence intervals of any magnitude for any percentage of species to be protected (eg. PC5 30%, PC80 70%). This is done using a bootstrap technique that randomly selects data from the toxicity dataset, to provide 501 estimates of the concentration that should protect the desired percentage of species. The percentile of these estimates corresponding to the chosen confidence limit is then calculated. Thus, the 5th percentile of 501 estimates of the PC95 would become the PC95 95%.

It should be noted that the stated level of protection of a TV (eg. 95% of species with 50% certainty) is theoretical and may not occur in reality. This could occur for a number of reasons, which include:

1. the fact that only 50% of the TVs will protect 95% of species - so the percentage of species protected could be higher or lower;
2. that the sensitivity of the remainder of the species to the toxicant may not be the same as that used to derive the TV;

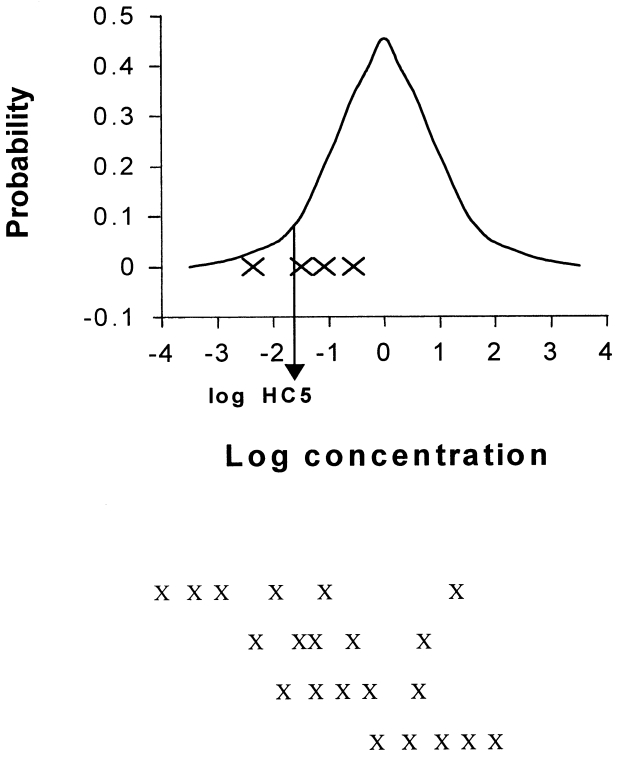

Figure 1. *An illustration of how PC95 estimates are distributed around the actual PC95 value for all species (modified from Warne, 1998). The curve represents the distribution of toxicity for a chemical on all species, with the actual PC95 value indicated. The 'X's in the graph represent log PC95 estimates from four sub-samples of the toxicity data as shown below the figure (data sets 1-4).*

3. the assumptions of the statistical model, used to calculate the TVs, may not be appropriate or valid; and
4. that the toxicity in the field and in the laboratory may be different due to a number of parameters including pH, organic matter, or suspended particulate matter.

Use of the Burr Type III Statistical Distribution Method

Before the TVs were calculated it was necessary to determine which grade of TV could be derived and by which method. This was done using the rules stated earlier in this paper and the minimum data requirements of the BT III SD method which are stated in Tables 4 and 5. Once this had been determined the data that were identified as extraneous were not used in any further calculations. For example, if a HR TV could be derived then only chronic data were used in subsequent calculations - the extraneous acute data were not used.

The BT III SD method can be used to derive HR and MR TVs for all chemicals that meet the minimum data requirements (Tables 4 and 5) and LR (interim) TVs for non-polar narcotics (refer to the section on QSARs). The method can not be used to derive LR (interim) TVs for chemicals other than non-polar narcotics, nor LR (ECL) TVs for any chemical. This is because the minimum data requirements of these types of TVs (Table 6) do not meet the minimum data requirements of the BT III SD method (Tables 4 and 5). The BT III SD method can use toxicity data with any measure of toxicity as long as only one measure is used for each calculation (Aldenberg and Slob, 1993). Despite this, TVs were only derived using chronic NOECs for organic chemicals and estimates of chronic NOECs for metals (refer to the section on converting chronic data to chronic NOECs) and non-polar narcotics (refer to the section on QSARs).

Only one toxicity datum is used in the BT III SD method to represent the sensitivity of each species. However, as there were usually multiple toxicity data for each species, the data generally required some manipulation. The rules governing the manipulations were adopted from Van de Plassche *et al.* (1993) and are stated below:

1. if there was only one toxicity datum, that was taken to represent the species;
2. if there were several toxicity values for the same endpoint, the geometric mean of the values was calculated and was taken to represent the species; and
3. if there were several toxicity values for different endpoints, the endpoint with the lowest geometric mean was taken to represent the species.

Once a single datum point was obtained for each species for which there were toxicity data, the values were plotted in a frequency versus toxicity histogram, in order to identify the chemicals that had a bi-modal toxicant sensitivity distribution. This was done because, while the BT III SD method can model such data, the range of the toxicity values would be very large, causing the concentration that corresponds to high PC values (eg PC 95 or the 5th percentile) to be unrealistically low. Pesticides were one group of chemicals that frequently had bi-modal distributions.

For chemicals with a bi-modal distribution all the data of the less sensitive group of species were removed from the calculations. The data for the more sensitive group of organisms were then entered in the BurrliOZ (Campbell *et al.* 2000) computer program, which does the BT III calculations. By using the BT III method on the more sensitive group of organisms the data could validly be modelled and the range of the toxicity data would be reduced, giving more environmentally realistic TVs. If however, the data did not have a

bi-modal distribution then the data that represented each of the species were entered into the BurrliOZ program (Campbell *et al.* 2000) and the TV calculated.

As stated earlier, the PC95 50% should protect 95% of species in an ecosystem with 50% certainty. However, the duration of the toxicant exposure that the organisms are protected from depends on the type of toxicity data used to derive the PC95 50% values. If the PC95 was calculated using chronic data, then the PC95 50% should provide the stated level of protection from chronic exposures. If however, the PC95 50% was based on acute toxicity data as is the case when MR TVs are being derived, the organisms will only be protected from acute exposure to toxicants. In the latter case, this is not consistent with the level of protection required by ANZECC (1992) and the PC95 50% value was modified so that it provided the same degree of protection to chronic exposures. This was done by dividing the PC95 50% value by either the acute to chronic AF (ie. 10) or an ACR. If however, the metal was an essential element (ie. boron, copper, iron, manganese, molybdenum, nickel, selenium and zinc) the acute to chronic AF was reduced to 2. This was done because, while an overabundance of these chemicals cause toxic effects, too low a concentration can also cause toxic effects due to deficiency.

An ACR was used in preference to an AF. Rules used in the application of ACRs to MR PC95 50% values were:

1. if there was only one ACR for a particular chemical, that ACR was used irrespective of the species or type of organism it was derived from; and
2. if there was more than one ACR, the geometric mean of all the ACR values was used.

In addition to the PC95 50% (the usual TV), concentrations that offered different levels of protection (i.e. PC99 50%, PC90 50%, and PC80 50%) were calculated using the BurrliOZ software (Campbell *et al.* 2000). The lower levels of protection were calculated because they can be used where it has been scientifically proven that the site is degraded. This concept is discussed in more detail in McAlpine and Humphrey (2001). The higher level of protection (PC 99 50%) was calculated for those chemicals that had the potential to bioaccumulate (ie. log Kow > 4 and/or BCF > 10 000). This higher level of protection is an attempt to indirectly account for the fact that these chemicals bioaccumulate and therefore could cause toxic effects to those organisms that eat other organisms that have been exposed to the chemicals (ie. secondary poisoning). This is further considered during the implementation and site-specific phases of the WQGs (Chapman *et al.* 2001).

THE ANZECC & ARMCANZ ASSESSMENT FACTOR METHOD

Background

This method (Table 6) is a modification of the Canadian (CCREM 1987; CCME 1991) AF method that was used previously by ANZECC (1992). The modifications were:

1. the type of TVs calculated and the introduction of the environmental concern level (ECL);
2. the type of toxicity data used; and
3. the minimum data requirements.

The Canadian (CCME 1991) and previous ANZECC water quality guidelines had only two levels of guidelines. These were the guidelines and interim guidelines. The ANZECC & ARMCANZ (2000) guidelines have three grades of TVs. These are the HR TVs, which equate with the guidelines of Canada (CCME 1991) and ANZECC (1992), the MR TVs which have no direct equivalent, and LR TVs which equate with the interim guidelines of Canada and ANZECC (1992). In addition, a new class of LR TV was established, the environmental concern level (ECL) (OECD 1992), which has no equivalent in the Canadian and ANZECC (1992) guidelines. The ECL can be calculated when there is as little as one datum. However, because of the very limited quantity of data used to derive these TVs, they are classed as having low reliability.

The AF method used by Canada (CCME 1991) and Australia (ANZECC 1992) used acute EC50 and LC50 data as well as chronic LOEC toxicity data to derive guidelines. Warne (1998) in contrast, argued that NOEC data are more suitable than LOEC data for deriving TVs that aim to protect organisms from harm. Hence, the ANZECC & ARMCANZ AF method uses the same acute data as Canada (CCME 1991) and ANZECC (1992) but chronic NOEC toxicity data. The acceptable types of toxicity data for deriving TVs are presented in Table 6.

The BT III SD method and the CCME AF method originally had different minimum data requirements. The data requirements of the CCME (1991) AF method were modified so that they were more consistent with the requirements of the BT III SD method as part of the modifications to form the ANZECC & ARMCANZ (2000) water quality guidelines. The rationale for these changes is provided in Warne (1998).

The magnitude of the AFs is governed by the number of extrapolations being made to convert the toxicity data to TVs, with each extrapolation having an AF of ten. Assessment factors of ten were used for each extrapolation despite considerable reservations about their validity, because it was felt that there was

Table 6. The type of toxicity data, the minimum data requirements and the magnitude of the assessment factors used to derive the trigger values (TVs) using the ANZECC & ARMCANZ AF method.

Type of Trigger Value	Type of Toxicity Data	Assessment Factor[a]	Minimum data requirements
HR TV	1. Chronic MS[b] NOEC	1. 10	1. chronic NOEC data from at least three MS studies that meet the quality requirements stated in the text.
	2. Chronic SS[c] NOEC	2. 10	2. chronic SS NOEC toxicity data for five different species. One for a fish, two for different taxonomic groups of invertebrates (see Table 5), one for a plant (Table 5) and one for any of the above groups or a different taxonomic group of organisms (Table 5).
MR TV	Acute SS LC/EC50	10 x ACR or 100	acute SS LC/EC50 toxicity data for five different species. One for a fish, two for different taxonomic groups of invertebrates (see Table 5), one for a plant (Table 5) and one for any of the above groups or a different taxonomic group of organisms (Table 5).
LR (Interim) TV	1. Chronic SS NOEC and acute SS LC/EC50	1. 10 x ACR or 100	1. three SS toxicity data, which can be acute or chronic. One for a fish, one for an invertebrate and one for an algae.
	2. Chronic SS NOEC	2. 20	2. three chronic NOEC toxicity data. One for a fish, one for an invertebrate and one for an algae.
	3. Chronic QSAR estimates	3. 10	3. QSAR estimates of chronic NOEC values for nineteen species.
LR (ECL) TV	Any toxicity data	1000	A toxicity datum.
		200	A chronic toxicity datum which is the lowest toxicity value.

[a] If the chemical is an essential element then one of the AFs is reduced to 2. [b] MS is multi-species. [c] SS is single species

insufficient evidence to derive new values (Warne 1998). The three extrapolations that are made in the ANZECC & ARMCANZ AF method are from laboratory to field, acute to chronic, and few to many species.

Only one extrapolation (ie. 10) is used to derive HR TVs, but the particular extrapolation used depends on the toxicity data. If chronic multiple species (MS) NOEC data were used the few to many species extrapolation was used due to the relative simplicity of the test systems compared to ecosystems. If however, chronic single species (SS) NOEC data were used then the laboratory to field extrapolation was used.

The laboratory to field and the acute to chronic extrapolations were applied to derive MR and the LR (interim) TVs from acute toxicity data. Thus, both these TVs were derived using an AF of 100. The same AF was used for the two different grades of TV for two reasons. First, it was not felt that a larger AF was warranted for the interim TV. Second, the lower confidence that the LR TV would offer the stated level of protection is indicated by the name.

When LR (interim) and LR (ECL) TVs were derived using chronic toxicity data they also used the laboratory to field and acute to chronic extrapolations. However, because there was some, although a limited number of, chronic data the acute to chronic extrapolation was reduced from 10 to 2. Thus, the LR (interim) and LR (ECL) TVs were obtained by dividing the lowest toxicity value by 20 and 200 respectively (Table 6). However, when LR (ECL) TVs were derived using acute

toxicity data then all three assessment factors were used - thus the lowest datum was divided by 1000.

Use of the ANZECC & ARMCANZ Assessment Factor Method

Before the TVs were calculated it had to be determined which grade of TV could be derived and by which method. This was done using the rules stated earlier in this paper and the minimum data requirements of the ANZECC & ARMCANZ AF method stated in Table 6. The ANZECC & ARMCANZ AF method can be used to derive HR, MR and both classes of LR TVs.

The minimum data requirements of the ANZECC & ARMCANZ AF method (Table 6) had to be met in order for a particular level of TV to be derived. However, there may be individual cases where expert judgment can be used to vary the data requirements. For example, it is well known that some insecticides have extremely low toxicity to plants - in such cases it could be argued that toxicity data for a plant species is not required and there is no decrease in the grade of the TV. For example, there might be toxicity data for fish and crustaceans, which would normally mean that only a LR (ECL) TV could be derived. However, if the chemical was an insecticide, the requirement for alga in order to derive an LR (interim) TV could be dropped and a LR (interim) TV could be derived.

Having determined the method to use and the grade of TV to be calculated the TVs were calculated by multiplying the lowest toxicity value for the chemical in question, by the appropriate AF (Table 6). The use of an ACR to calculate MR and LR (interim) TVs was preferred to the use of the acute to chronic AF. In such a case, the TV was calculated as the lowest toxicity value for the chemical in question, multiplied by the ACR and the laboratory to field AF (Table 6).

The rules for applying ACRs in the ANZECC & ARMCANZ (2000) AF method were:

1. if there was only one ACR for a particular chemical, then that ACR was used irrespective of the species or type of organism it was derived from; and
2. if there was more than one ACR, then the geometric mean of the ACR values (ACRgm) for each taxonomic group (as defined in Table 5) was calculated and then
   2a. if there was an ACRgm for the same taxonomic group that was used to derive the TV, then that ACRgm was used; or
   2b. if there was an ACRgm for only a taxonomic group other than that used to derive the TV, then that ACRgm was used; or
   2c. if there were ACRgm values for several taxonomic groups other than that used to derive the TV, then the geometric mean of all the ACR values was used.

If the chemical was an essential element (Boron, Copper, Iron, Manganese, Nickel, Nitrate, Selenium and Zinc) then one of the AFs was decreased from 10 to 2. Thus, the AFs were 200, 20 and 2 rather than 1000, 100 and 10 (Table 6). The reason for this has been presented in the section on the Burr Type III SD method.

EVALUATION OF THE DERIVED TRIGGER VALUES

Once the TVs had been derived they were evaluated by comparing the TV with the raw toxicity data used to derive them. The aim was to determine whether any species, for which toxicity data were available, would suffer toxic effects if exposed to the TV. If any of the following conditions were met then the TV was considered to provide inadequate protection:

1. If a HR TV based on a limited data base (< 10 data) was greater than the geometric mean of experimental chronic NOEC data for any important species. The species can be important on the basis of commerce, rarity, or ecological significance.
2. If a HR TV based on a large data set (> 10 data) has more than 5% of species with geometric means greater than the TV.
3. If a MR TV was above any experimental chronic NOEC datum. Exceptions were made if the geometric mean for the species was considerably greater than the TV and/or the value was considered an outlier after either examining the paper or if the value was more than two orders of magnitude different from the remainder of the data for that species.
4. If the TV was greater than one third of any acute toxicity data. Exceptions were made if the lowest datum was considered an outlier.

In such cases a number of means of modifying the TV were implemented. For HR TVs the HC99 50% was adopted as the TV. If adequate protection was still not provided and the chronic dataset that was used was small, then a MR TV was derived and it was evaluated. For MR TVs, if an ACR was used to derive the TV it was examined and if possible a more appropriate ACR was used in the calculation. If there was no more appropriate ACR then the default ACR of ten was used in the calculation. If the resultant PC95 50% value still did not meet the criteria then the PC99 50% was calculated firstly using an ACR or alternatively using the default AF.

Low reliability (interim) TVs were not validated as they were obtained by dividing the lowest toxicity value by an AF and therefore they will automatically offer adequate protection.

The TVs for metals were also checked against background concentrations to ensure that unrealistically low TVs (lower than the background) were not derived. The background data used was that presented in the Australian and New Zealand water quality guidelines (ANZECC & ARMCANZ 2000). None of the TVs was lower than background levels. This should also be considered during the derivation of site-specific TVs (Chapman 2001).

FACTORS NOT ACCOUNTED FOR IN THE DERIVATION OF THE TRIGGER VALUES

Organisms in aquatic environments are usually exposed simultaneously to multiple toxicants in the form of mixtures. It was, however, not possible to account for the toxicity of mixtures when deriving TVs as the toxicity may differ for each mixture. Thus, the toxicity of mixtures needs to be taken into account when site-specific assessments are conducted (Chapman 2001).

ACKNOWLEDGEMENTS

The author would particularly like to thank Dr John Chapman for his invaluable assistance in the development of the derivation scheme and calculation of the TVs. Thanks also to Dr David Leece, Dr Graeme Batley, Dr Quanxi Shao, Dr Chris Hickey, Mr Kevin McAlpine, Dr Scott Markich, Dr Rick van Dam, Ms Helen Davies, Mr Martin Krough, Mr Rick Krassoi and Professor Barry Hart for discussions during the development of the framework; Ms Zada Lipman and Ms Jeni Emblem for providing literature and expert comment on the Precautionary Principle; and Mr Meenatchi Sunderam, Ms Rebecca Rose, Dr Fleur Pablo, Ms Annette Schifko, Dr Scott Markich and Dr Chris Hickey for their contributions to deriving the trigger values; Dr Eddy Campbell, Dr Mark Palmer, Dr Quanxi Shao, and Ms Dale Wilson, for developing the BurrliOZ software; and Dr Rick Van Dam, Dr Chris Hickey and Dr Chris Humphrey for discussions during the development of BurrliOZ. The New South Wales Environment Protection Authority, the Environmental Research Institute of the Supervising Scientist, and the CSIRO Centre for Advanced Analytical Chemistry and Mathematical and Information Sciences financed this work.

REFERENCES

Aldenberg T and Slob W. 1993. Confidence limits for hazardous concentrations based on logistically distributed NOEC toxicity data. *Ecotoxicol. Environ. Saf.* **25**, 48-63.

ANZECC (Australian and New Zealand Environment and Conservation Council) **1992**. *Australian water quality guidelines for fresh and marine waters.* ANZECC, Canberra, Australia.

ANZECC and ARMCANZ (Australian and New Zealand Environment and Conservation Council and Agriculture and Resource Management Council of Australia and New Zealand) **1994**. *National water quality management strategy. Policies and principles. A reference document.* ANZECC & ARMCANZ, Canberra, Australia.

ANZECC & ARMCANZ (Australian and New Zealand Environment and Conservation Council and Agriculture and Resource Management Council of Australia and New Zealand) **1999**. *National water quality management strategy, Draft Australian and New Zealand guidelines for fresh and marine water quality*. ANZECC & ARMCANZ, Canberra, Australia.

ANZECC & ARMCANZ (Australian and New Zealand Environment and Conservation Council and Agriculture and Resource Management Council of Australia and New Zealand) **2000**. *National water quality management strategy, Australian and New Zealand Guidelines for fresh and marine water quality*. ANZECC & ARMCANZ, Canberra, Australia.

Campbell E, Palmer MJ, Shao Q and Wilson D. 2000. BurrliOZ: A Computer Program for the Estimation of the Trigger Values for the ANZECC & ARMCANZ water quality guidelines. In *National water quality management strategy, Australian and New Zealand guidelines for fresh and marine water quality*. ANZECC & ARMCANZ, Canberra, Australia.

CCREM. (Canadian Council of Resource and Environment Ministers) **1987**. *Canadian water quality guidelines.* Canadian Council of Resource and Environment Ministers, Inland Water Directorate. Environment Canada, Ottawa, Canada.

CCME. (Canadian Council of Ministers of the Environment) **1991**. Appendix X - A protocol for the derivation of water quality guidelines for the protection of aquatic life. In *Canadian water quality guidelines.* Canadian Council of Resource and Environment Ministers, Inland Water Directorate. Environment Canada, Ottawa, Canada.

Chapman JC. 2001. The revised Australian and New Zealand water quality guidelines for toxicants: Approach to their derivation and application. *Australas. J. Ecotox.* **7**, *95-108*.

Chapman JC, Warne MStJ and Patra RWR. 2001. Case studies for applying the Australian and New Zealand toxicant guidelines. *Australas. J. Ecotox.* **7**, *157-174*.

Forbes TL and Forbes VE. 1993. A critique of the use of distribution-based extrapolation models in ecotoxicology. *Functional Ecology* **7**, 249-254.

Fox DR. 1999. Setting water quality guidelines: A statisticians perspective. *SETAC News*, May 1999, 17-18.

Hansch C, Leo A and Hoekman D. 1995. *Exploring QSAR. Hydrophobic, electronic and steric constraints.* American Chemical Society, Washington DC, USA.

Hickey CW and Pyle E. 2001. Derivation of water quality guidelines for heavy metals using a risk-based methodology: A site-specific approach for New Zealand. *Australas. J. Ecotox.* **7**, *137-156*.

Mackay D, Shiu WY and Ma KC. 1992a. Illustrated handbook of physical-chemical properties and environmental fate for organic chemicals. Volume I – Monoaromatic hydrocarbons, chlorobenzenes and PCBs. Lewis Publishers, Michigan, USA.

Mackay D, Shiu WY and Ma KC. 1992b. Illustrated handbook of physical-chemical properties and environmental fate for organic chemicals. Volume II – Polynuclear aromatic hydrocarbons, polycyclic dioxins and dibenzofurans. Lewis Publishers, Michigan, USA.

Mackay D, Shiu WY and Ma KC. 1993. Illustrated handbook of physical-chemical properties and environmental fate for organic chemicals. Volume III – Volatile organic chemicals. Lewis Publishers, Michigan, USA.

Mackay D, Shiu WY and Ma KC. 1995. Illustrated handbook of physical-chemical properties and environmental fate for organic chemicals. Volume IV – Oxygen, nitrogen and sulfur containing compounds. Lewis Publishers, Michigan, USA.

Markich SJ, Brown PL, Batley GE, Apte SC and Stauber JL. 2001. Incorporating metal speciation and bioavailability into water quality guidelines for protecting aquatic ecosystems. *Australas. J. Ecotox.* **7**, 109-122.

Markich SJ, Warne MStJ, Westbury A-M and Roberts CJ. *In press*. A compilation of data on the toxicity of chemicals to Australasian species. Part 3: Metals. *Australas. J. Ecotox.*

McAlpine KW and Humphrey CL. 2001. An overview of the new guidelines for Australian and New Zealand aquatic ecosystems. *Australas. J. Ecotox.* **7,** 43-52.

OECD. (Organisation for Economic Co-operation and Development). **1992**. *Report of the OECD workshop on extrapolation of laboratory aquatic toxicity data to the real environment.* OECD Environment Monographs No. 59. OECD, Paris, France.

OECD. (Organisation for Economic Co-operation and Development). **1995**. *Guidance document for aquatic effects assessment.* OECD Environment Monographs No. 92. OECD, Paris, France.

Petersen LS and Pedersen F. 1995. *Water quality criteria for selected priority substances*. Danish Ministry of the Environment and Energy, Danish Environmental Protection Agency, Copenhagen, Denmark.

RIVM. (National Institute of Public Health and Environmental Protection) **1995**. *Quality assurance document for deriving environmental quality objectives (INS and I-Values)*. Report No. ACT/H/003. RIVM, Bilthoven, The Netherlands.

Roux DJ, Jooste SHJ and MacKay HM. 1996. Substance-specific water quality criteria for the protection of South African freshwater ecosystems: Methods for the derivation and initial results for some inorganic toxic substances. *S. African J. Sci.* **92**, 198-206.

Shao Q. 2000. Estimation for hazardous concentrations based on NOEC data: An alternative approach. *Environmetrics* **11**, 583-595.

Shiu WY, Ma KC, Varhanickova D and Mackay D. 1994. Chlorophenols and alkylphenols: A review and correlation of environmentally relevant properties and fate in an evaluative environment. *Chemosphere* **29**, 1155-1224.

Stephan CE, Mount DI, Hansen DJ, Gentile JH, Chapman GA and Brungs WA. 1985. *Guidelines for deriving numerical national water quality criteria for the protection of aquatic organisms and their uses.* USEPA Report No. PB-85-227049. USEPA, Washington DC, USA.

Sunderam RIM, Warne MStJ, Chapman J, Rose R, Hawkins J, and Pablo F. 2000. The ANZECC & ARMCANZ toxicant water quality guideline database. ANZECC & ARMCANZ, Canberra, Australia.

Tomlin C. 1994. *The pesticide manual. Incorporating the agrochemicals handbook*. British Crop Protection Council and Royal Society of Chemistry, Oxford, England.

USEPA. (United States Environmental Protection Agency). **1986**. *Quality criteria for water.* EPA Report No. EPA 440/5 86-001. US Department of Commerce, Virginia, USA.

USEPA. (United States Environmental Protection Agency) **1994.** *AQUIRE (Aquatic toxicity information retrieval).* Office of Research and Development. National Health and Environmental Effects Research Laboratory, Mid-Continental Ecology Division, Duluth, Minnesota, USA.

Van de Plassche EJ, Polder MD and Canton JH. 1993. *Derivation of maximum permissible concentrations for several volatile compounds for water and soil.* Report No. 679101 008. National Institute of Public Health and Environment Protection, Bilthoven, The Netherlands.

Van Leeuwen CJ, Van der Zandt PTJ, Aldenberg T, Verhaar HJM and Hermens JLM. 1992. Application of QSARs, extrapolation and equilibrium partitioning in aquatic effects assessment. 1. Narcotic industrial pollutants. *Environ. Toxicol. Chem.* **11**, 267-282.

Verhaar HJM, Ramos EU and Hermens JLM. 1996. Classifying environmental pollutants. 2: Separation of class 1 (baseline toxicity) and class 2 (polar narcosis) type compounds based on chemical descriptors. *J. Chemometrics* **16,** 109-116.

Verscheuren K. 1983. *Handbook of environmental data on organic chemicals.* Van Nostrand Reinhold, New York, USA.

Wagner C and Løkke H. 1991. Estimation of ecotoxicology protection levels from NOEC toxicity data. *Water Research* **25**, 1237-1242.

Warne, MStJ, Connell, DW and Hawker, DW. 1991. Comparison of the critical concentration and critical volume hypotheses of non-specific toxicity for individual compounds. *Toxicology,* **66**, 187-195.

Warne MStJ, Westbury A-M and Sunderam RIM. 1998. A compilation of data on the toxicity of chemicals to Australasian species. Part 1: Pesticides. *Australas. J. Ecotox.* **4**, 93-144.

Warne MStJ and Westbury A-M. 1999. A compilation of data on the toxicity of chemicals to Australasian species. Part 2: Organic Chemicals. *Australas. J. Ecotox***. 5**, 21-86.

Warne, MStJ. 1996. The theory and practise of extrapolation techniques to derive environmental quality criteria: The Dutch approach. In *Proceedings of the third national workshop on the health risk assessment and management of contaminated sites.* Contaminated Sites Monograph Series No 5. South Australian Health Commission, Adelaide, South Australia, pp 402-416.

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.

Weast RC 1987. *CRC Handbook of chemistry and physics.* CRC Press, Boca Raton, Florida, USA.