supervising scientist report **135**

Critical review of methods to derive water quality guidelines for toxicants and a proposal for a new framework

Michael StJ Warne

Environment Australia
Department of the Environment

supervising scientist

Michael StJ Warne – Ecotoxicology Section, Environment Protection Authority of New South Wales, Centre for Ecotoxicology, University of Technology Sydney, Westbourne St, Gore Hill, NSW 2065, Australia.

This report should be cited as follows:

Michael StJ Warne 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.

The Supervising Scientist is part of Environment Australia, the environmental program of the Commonwealth Department of Environment and Heritage.

Supervising Scientist
GPO Box 787, Canberra ACT 2601 Australia

ISSN 1325-1554

ISBN 0 642 24338 7

Printed in Darwin by NTUniprint.

Contents

Figures

Tables

Acknowledgments

The author would like to thank Dr John Chapman, Dr David Leece, Dr Chris Hickey, Professor Barry Hart, Mr Kevin McAlpine, Dr Rick van Dam, Ms Helen Davies, Mr Martin Krough and Mr Rick Krassoi for their critical comments on the manuscript. The author would also like to thank Ms Zada Lipman and Ms Jeni Emblem for providing literature and expert comment on the Precautionary Principle.

Preamble

Late in 1995 ERISS was given the task of managing the review of the ANZECC water quality guidelines. In September 1996 ERISS asked if the Ecotoxicology Section of the NSW EPA was willing to revise the water quality guidelines for toxicants. The tasks to be completed in this review included:

1. Review methodologies at an international level for calculating guideline values for the protection of aquatic ecosystems;
2. Use real data to compare results from the more appropriate methods and recommend a methodology for ANZECC guidelines;Collate and assess relevant ecotoxicological data from Australia, New Zealand and overseas for freshwater, estuarine, marine and more generally tropical environments;
3. Conduct a review of the relative sensitivity of Australian and New Zealand aquatic life forms to toxicants compared to overseas species;
4. Calculate guideline values for all chemicals using the accepted methodology giving preference to Australian and New Zealand data;
5. Determine and prioritise chemicals/subject areas that require additional studies to redress information deficiencies in future reviews;
6. Discuss the philosophy being the use of water quality guidelines and the approach used to derive the values.

This document addresses the first two tasks in the review of ANZECC water quality guidelines for toxicants.

Executive summary

The aims of this document were to review and critically assess the available methodologies for deriving water quality guidelines (WQGs) and to substitute real toxicity data into these methods in order to recommend which method or methods were the most suitable for use by ANZECC.

There are two principle approaches to determining WQGs. The original method called the assessment factor method divided the lowest toxicity value by an assessment factor, the magnitude of which was based on the number, character and quality of the available toxicity data. The more data, and the more realistic they were, the lower the magnitude of the assessment factor. Typical assessment factors used are 10, 100 and 1000. The aim of such methods is to protect all species from lifetime exposures to toxicants. This type of approach is used by a variety of countries including Australia, New Zealand, USA, Canada, Denmark, The Netherlands and South Africa and the OECD has recommended it. A new approach—statistical extrapolation methods—has been developed since 1984. They use toxicity for all species that are available and fit a particular distribution to the data and from this calculate the concentration that should protect any percentage of species. This type of approach has also been adopted by many countries—in fact all the above mentioned countries except Australia, New Zealand and Canada use both the assessment factor and statistical extrapolation techniques. The USA, The Netherlands, OECD and South Africa use the statistical extrapolation techniques in preference to the AF methods unless there is insufficient data whereas Denmark prefers the AF method to the statistical extrapolation techniques.

Of the various versions of the AF method and the statistical extrapolation method the current ANZECC AF method and the Aldenberg and Slob method used by The Netherlands were chosen for detailed analysis. The critical evaluation of these two methods revealed that both had limitations. The major weaknesses of the assessment factor method are the arbitrary nature of the AFs, the questionable validity of the acute to chronic ratios, the assessment factors are too small to provide protection to all species and the method is not transparent ie. estimating the level of protection provided is not possible. The major weaknesses of the Aldenberg and Slob method are the limited number of WQGs that can be derived by this method, the questionable suitability of using no observed effect (NOEC) data, its relative complexity, and the validity of several of the assumptions may be compromised.

Comparison of the two methods in terms of the Precautionary Principle revealed that both methods had a number of precautionary elements. Despite this neither method was, *in toto*, precautionary because they both failed to account for the toxicity of mixtures, accumulation of toxicant in the animal tissue, and transfer of chemicals between the various compartments of the environment.

The comparison of 40 WQGs derived using real toxicity data by both methods indicated that the Aldenberg and Slob method derived WQGs that were significantly lower and offered a significantly greater degree of protection than the AF method. However, the WQGs derived by the Aldenberg and Slob method were not consistently lower than those derived by the AF method.

It was therefore concluded that it was not possible to state with any degree of confidence which of methods (Aldenberg and Slob or AF) was better.

The framework suggested for the derivation of WQGs is based upon this conclusion—it uses both methods (ie. the current ANZECC AF and the Aldenberg and Slob methods) to

determine an estimate of the WQG for a chemical and then chooses the lower value as the WQG. Depending on the quality, quantity and type of toxicity data available a number of different methods are used to derive different types of WQGs (see Figure 1). There is a different level of confidence in each type of WQG that they provide the desired degree of environmental protection.

It is also proposed that three types of WQGs be derived based on the type, quality and quantity of the toxicity data available. For chemicals for which there is adequate suitable data Level I and Level II WQGs would be derived. There is a higher degree of confidence that Level 1 WQGs deliver the desired degree of environmental protection than Level II WQGs. However, in both cases there is sufficient confidence that the WQG should deliver the desired degree of protection. When the quality, quantity and type of toxicity data is not adequate then there is a low degree of confidence that the resulting WQG will provide the desired degree of protection. Such WQGs are termed ‘interim’ WQGs.

In the section comparing the precautionary natures of the AF and Aldenberg and Slob methods it was highlighted that a major limitation was that the toxicity of mixtures, accumulation of toxicant in the animal tissue, and transfer of chemicals between the various compartments of the environment were not considered. It was felt that given the current level of knowledge it was only possible to incorporate the toxicity of mixtures into a new method for deriving and using WQGs. Thus it is recommended that when water-resource managers use the WQGs the toxicity of mixtures be considered.

1 Introduction

Water quality guidelines (WQGs) are commonly used by environmental managers seeking to protect the aquatic environment (USEPA 1986, 1994, CCREM 1991, ANZECC 1992, Van de Plassche et al 1993, OECD 1995, Roux et al 1996, Petersen & Pedersen 1995). The majority of the current Australian and New Zealand WQGs for toxicants were derived using an assessment factor (AF) approach (ANZECC 1992) which was based on the Canadian AF method (CCREM 1991). The AF method has been the main method used internationally to derive WQGs. However, in the last decade, new methods using a risk-based approach have been developed (Stephan et al 1985, Kooijman 1987, Van Straalen & Denneman 1989, Wagner & Løkke 1991, Aldenberg & Slob 1993). These methods are statistically based and have been adopted by Denmark, South Africa, The Netherlands and USA to derive their national WQGs. A third method for deriving water quality guidelines is to use toxicity data from multiple species toxicity tests. While this method is incorporated into The Netherlands (Van de Plassche et al 1993) and OECD (1995) frameworks for deriving WQGs, it does not appear to have been used yet.

New ANZECC WQGs are to be derived and rather than simply adopting or modifying overseas values as was largely the case previously (ANZECC 1992). This exercise will provide the opportunity to develop expertise in deriving WQGs and will ensure that greater consideration is given to Australian and New Zealand species and water conditions. The objectives of this paper were to:

1. Critically assess the suitability of various types of toxicity data and methods for estimating toxicity;
2. Critically evaluate methods for deriving water quality guidelines;
3. Compare the methods for deriving water quality guidelines in terms of how well they adhere to the Precautionary Principle;
4. Compare water quality guidelines derived by the various methods using the same toxicity data;
5. Recommend a framework to be used in the forthcoming review of the ANZECC water quality guidelines.

2 Review of methods to derive water quality guidelines

2.1 Assessment of types of toxicity data for use in deriving water quality guidelines

There are four key characteristics of toxicity data that are relevant to deriving WQGs. These are the ecological relevance of the experimental systems (ie, single species or multiple species toxicity tests), the duration of exposure to toxicant (ie, acute or chronic exposures), the ecological relevance of the biological endpoint measured, and the magnitude of the biological endpoint measured (ie, EC50 and LC50, NOEC and LOEC, or LC5 and EC5). These characteristics will be examined below.

2.1.1 Single species versus multiple species toxicity data

The test systems used to generate single species (SS) toxicity data are simplistic. Generally, individuals of a single species with fairly uniform characteristics have been bred in the

laboratory and then exposed to a toxicant in water. In many cases, the tests are static with no renewal of the water. The aim of these experiments is to standardise as many environmental factors as possible in order to increase the reproducibility of the experiments. This standardisation, however, has the unfortunate effect of reducing the environmental realism of the test systems and the resulting toxicity data.

The great strength of SS tests and the resulting data is that they are very simple to conduct, relatively rapid and inexpensive. The information is also easy to interpret, as there are few, if any, modifying factors that need to be considered. Another advantage of SS toxicity data is that they are far more abundant for a wider variety of chemicals, organisms and biological endpoints than multiple species (MS) toxicity data. This is important as it means that many WQGs can be derived using this type of data.

The methods used to obtain SS toxicity data, and hence the toxicity data themselves, have been criticised because: the routes of exposure are environmentally unrealistic and too simple (eg exposure to a constant concentration of the toxicant); the experiments do not take into account the variation in environmental conditions that occurs in the field; they do not account for the variation in wild populations over time; they only examine species specific responses; they do not account for species interactions nor the possibility of indirect effects nor possible bioaccumulation effects (eg Ward & Jacoby 1995, Graney et al 1995).

A series of studies were conducted on freshwater rivers and streams by the USEPA to resolve whether single species exposed to effluents *in situ* were good predictors of the effects on aquatic communities (Marcus & McDonald 1992). The results indicated that SS tests were not useful in predicting the magnitude of any given toxic effect at the community level (Marcus & McDonald 1992, Parkhurst 1994). However, there was a relationship between toxic effects on single species and community effects at a coarser level (ie significant effects on single species meant it was likely that measurable detrimental effects would occur at the community level) (Marcus & McDonald 1992, Parkhurst 1994).

Some argue that due to the many limitations SS toxicity data are of little or no use in setting WQGs and the management of the environment (eg Underwood 1995, Cairns 1995). Whereas, Mount (1994) argues that SS toxicity tests are extremely useful and despite their limitations will continue to be the main form of toxicity test conducted. Similarly, Sprague (1995) argued that while many of the criticisms of single species tests are valid they are not particularly relevant to anti-pollution work as 'we get 90% of the answer from a small range of single-species tests.... The important thing is to take action, rather than wait for a 98% answer.' Others (Chapman 1995c, McPherson 1995) argue that single species tests should only be one of the means of assessing the effect of pollutants on ecosystems.

The use of more complex and environmentally realistic test systems such as MS tests can overcome all the limitations of single species toxicity tests. Multiple species tests systems are generally microcosms, mesocosms or field studies in which numerous species are exposed to one chemical. Microcosm and mesocosm studies are simplified aquatic ecosystems with mesocosms being larger and more biologically complex than microcosms. A lot of work has been conducted with these tests systems in the last ten years and some very elaborate systems have been established (eg Graney et al 1995). While less mesocosm research has been conducted in Australia and New Zealand there are currently five mesocosms operating: a marine mesocosm and four artificial stream mesocosms (Krassoi pers comm).

The major strength of MS toxicity data is that they are more environmentally realistic and therefore fewer assumptions have to be made in order to derive WQGs than is the case for SS data. However, there are some weaknesses in the experiments that provide these data

(La Point & Perry 1989, Guckert 1993). Two weaknesses are related to the cost of establishing and maintaining such experimental systems. Firstly, there is often insufficient replication of the treatments thus decreasing the ability to statistically reject hypotheses and the ability to determine if treatments are significantly different. Secondly, there are often insufficient treatments, thus, proper dose-response relationships can not be determined and causality may not be able to be established. An unrelated problem is, that generally, vast quantities of data are produced from such experiments and it is often difficult to attribute any toxicological or environmental meaning to the data.

The design of MS toxicity tests is crucial to the quality of the resulting data. Badly designed tests may result in data that are no more environmentally realistic than SS toxicity data and may not permit sound conclusions to be drawn. For these reasons the OECD (1995) has established a series of criteria by which the quality of MS tests can be determined. These criteria will be examined in detail later in this report.

Despite the advantages of MS tests and toxicity data, it is highly unlikely in the near future that these methods will become the predominant form of toxicity tests due to the cost and complexity. For example, the OECD (1992a) provides guidance on when such tests should be conducted and the USEPA has withdrawn funding for assessing pesticide impacts in mesocosms (Bradbury pers comm). Another disadvantage of such methods is the general lack of such toxicity data compared with that available for single species. Thus, these methods could only derive WQGs for a limited number of chemicals.

Multiple species toxicity data from well designed microcosms, mesocosms and field studies are better suited to derive WQGs than SS toxicity data. However, when such data are not available then single species toxicity data can and should be used to derive WQGs as they allow some prediction of effects at higher levels of organisation.

2.1.2 Acute versus chronic toxicity data

There do not appear to be any universally accepted definitions of acute and chronic data in aquatic toxicology, unlike the situation in human and mammalian toxicology. Rand et al (1995) defined acute tests as those in which the test organism is exposed for no more than four days. Chronic refers to experiments in which the organisms are exposed for periods representing a significant but undefined proportion of their life span. Thus, the duration necessary to be chronic is species dependent—with short-lived organisms requiring shorter exposure to be classified as chronic than longer-lived organisms.

It is well established that the aqueous concentration required to exert a toxic effect decreases as the duration of exposure increases. This relationship is asymptotic—ie, it approaches but never reaches zero as the time of exposure increases. This, however, does not imply the concentration of toxicant at the target site is inversely related to the exposure duration. In fact, the target site concentrations should theoretically be relatively constant (eg Abernethy et al 1988, Pawlisz & Peters 1993) and in practice they are (eg Abernethy et al 1988, Warne et al 1991, Pawlisz & Peters 1993). The time dependence of toxicity with aqueous concentrations and time independence with target site concentrations is due, at least partly, to the kinetics of uptake, metabolism and removal (depuration) (Connell 1990).

The general aim of WQGs is to protect organisms from lifetime exposures to toxicants, although the USEPA has a system that protects aquatic organisms from short-term as well as life-time exposures. The use of acute toxicity data is therefore not as appropriate for deriving WQGs as chronic toxicity data. If acute data are used then some means must be used to convert the data to chronic data or to account for the difference in toxicity between the two types of data. Methods for doing this will be discussed later in the review.

2.1.3 The ecological relevance of the biological endpoints of toxicity data

A wide variety of biological endpoints are measured in toxicity tests. These can be subdivided into a number of groups: behavioural, biochemical and 'others'. The other endpoints include mortality, reproductive impairment, hatchability, immobilisation and inhibition of growth. Behavioural endpoints include mobility, motility, burial rate, ventilation rates, swimming rate, phototactic responses, filtering and feeding rates. Biochemical endpoints include induction and modified activity of a range of enzymes, DNA changes, histopathological lesions and damage and immune system dysfunction.

While the debate over the environmental relevance of these three types of endpoints is not completely resolved, the majority of ecotoxicologists would agree that the ecological significance of biochemical and behavioural endpoints is doubtful (eg Holdway 1996, McCarty & Munkittrick 1996). The OECD (1992a) concurs, stating that survival, growth and reproduction endpoints have direct relevance for ecosystems and should be given more weight than other endpoints. At this stage it is difficult to attribute ecological relevance to all behavioural endpoints. This may be, at least in part, due to the limited research that has addressed this issue. An example of a behavioural endpoint that is ecologically relevant is the thoracic appendage beating rate of cladocera which Jones et al (1991) found was correlated with feeding rates and population survival. With such limited research it is impossible to state categorically whether behavioural endpoints are ecologically relevant or not.

It is therefore concluded, at this stage, that only toxicity data that measure survival (this includes immobilisation and photosynthesis), growth and reproduction are appropriate for deriving WQGs. This recommendation is not totally precautionary. However, it is scientifically defensible, whereas deriving WQGs based on biochemical and/or behavioural endpoints would be hard to justify except on the basis that an effect had occurred.

2.1.4 LC50 type data versus NOEC type data

The median lethal concentration (LC50) and the median effect concentration (EC50) are usually determined by the Probit (Bliss 1934a,b) or Spearman-Karber methods (Hamilton et al 1977). The reason for determining the median value is that the variance is least at the median and therefore the concentration that causes this toxic effect can be most accurately calculated. It was widely felt that this type of endpoint was inappropriate for protecting the environment. Therefore a new type of endpoint that is calculated using hypothesis testing statistical techniques and which measure a much lower biological effect was developed.

These new measures of toxicity are the no observed effect concentration (NOEC) or level (NOEL) or no observed adverse effect concentrations (NOAEC) and the lowest observed effect concentration (LOEC) or level (LOEL) or lowest observed adverse effect concentrations (LOAEC). The NOEC is the highest concentration used in a toxicity test that produces an effect that is statistically not significantly different ($p < 0.05$) from the controls. The LOEC is the lowest concentration used in a toxicity test that has an effect that is statistically different ($p < 0.05$) from the controls and that has all higher concentrations also exerting a significant effect at least as great as the LOEC (OECD 1996). These measures of toxicity are determined using hypothesis testing statistical techniques such as analysis of variance (ANOVA), Tukey's, Steel's Many One Rank, Wilcoxon Rank Sum, t-test with Bonferroni Adjustment and Dunnett's test.

There has been discussion in the literature over the relative merits of NOEC and LOEC toxicity data. Critics such as Hoekstra and Van Ewijk (1993), Noppert et al (1994) and Chapman et al (1996) feel that it is not appropriate to use such data for regulatory purposes. They would prefer to use toxicity values that correspond to a fixed biological effect (eg an

LC5 or EC10) that would be calculated using regression analysis. Some reasons for this are that NOEC and LOEC data depend on: the concentrations used in the toxicity tests; the sample size; the accuracy of determining the toxicity; and the statistical level of significance that is chosen (Hoekstra & Van Ewijk 1993, Noppert et al 1994, Chapman et al 1996). Also, NOEC values are not the toxicant concentrations that cause no biological effect, rather they are the highest concentrations tested that cause toxic effects that are not statistically different from the control. This is clearly illustrated by the work of Moore and Caux (1997) who examined 181 pesticide toxicity data sets and found that 77% of the NOECs exerted biological effects of between 10 to 30%. The USEPA (1991) and Hoekstra and Van Ewijk (1993) obtained similar results.

There are however, problems with the proposal of determining LC5 type data. The principal one is that there is considerable error in the estimation of values that are located in the tails of distributions. However, the error is greatly reduced if the estimate of toxicity is determined by interpolating between actual data rather than extrapolating from data (Stephan & Rogers 1985, Moore & Caux 1997). In order to overcome this problem toxicity tests would require more treatments at lower concentrations, larger numbers of replicates and more test organisms per treatment. A more important limitation is that there are virtually no such data currently available, thereby limiting the number of WQGs that could be derived using such data. Despite the limitations of NOEC and LOEC data, no regulatory body has recommended the cessation of their use. It is therefore recommended that NOEC data be used to derive the new set of ANZECC WQGs but that their use be phased out as LC5 type data become available.

3 Assessment of methods for estimating chronic toxicity data

Both the assessment factor and statistical extrapolation methods require certain minimum amounts of toxicity data in order to derive WQGs. When there are insufficient data, WQGs can not be derived by these methods. This lack of chronic toxicity data has been overcome by two main methods that estimate chronic toxicity. The first method is quantitative structure-activity relationships (QSARs) which derive chronic toxicity data based only on various physicochemical properties of the chemicals. The second approach uses acute toxicity data that are then converted by several means to chronic toxicity data.

3.1 Quantitative structure-activity relationships

Quantitative structure-activity relationships (QSARs) can be used to derive WQGs when there is very limited or no toxicity data. The QSARs are relatively simple models that relate the biological activity of chemicals to physicochemical properties or molecular descriptors of the chemicals. The variety of biological activities that have been modelled include toxicity, bioconcentration, inhibition of enzyme reactions, mutagenesis, carcinogenesis, antibacterial properties and biodegradation (eg Nirmalakhandan & Speece 1988, Hansch & Leo 1995). Properties that have been used in QSARs to model toxicity belong to four major types: partitioning properties; electronic properties; steric (shape and volume) properties; and miscellaneous properties (table 1). In fact over two hundred different properties have been used to predict toxicity (Van der Waterbeemd & Testa 1987).

Table 1 Examples of physicochemical properties used in QSARs

Type of physicochemical property	Physicochemical properties
Partition properties	Octanol-water partition coefficient, aqueous solubility, lipid-water partition coefficient.
Electronic properties	Polarisability, number of valence electrons, electron charge density, dipole moment, hydrogen bonding ability, acid dissociation constant.
Steric properties	Number of carbon atoms, molecular weight, and various measures of volume and surface area.
Miscellaneous properties	Boiling point, melting point, molecular weight, connectivity indices, vapour pressure.

The least toxic group of chemicals exerts their toxicity by the non-polar narcosis mode of action. Such chemicals partition into the lipid membranes of nerve cells and generally disrupt, by some as yet unknown means, the functioning of the nerve. The toxicity of such chemicals can be modelled using QSARs based only on the logarithm of the octanol-water partition coefficient (log Kow) as illustrated in the following equation (Könemann 1981):

$$\log 1/LC50 = 0.87 \log Kow - 4.87 \qquad r^2 = 0.98 \qquad (1)$$

QSARs for chemicals with more complex modes of action require more physicochemical properties in order to model toxicity successfully and often take the form:

$$\log \text{toxicity} = \pm a\,(\text{partition term}) \pm b\,(\text{electronic term}) \pm c\,(\text{steric term}) \pm d \qquad (2)$$

A large number of QSARs have been published for a wide variety of chemicals and different mechanisms of action to numerous species (eg USEPA 1988, Hermens 1989, Donkin 1994). However, there have been only five QSARs derived based upon Australian organisms. These have modelled the toxicity of narcotic agents to mixed marine bacteria (Warne et al 1989a, Warne et al 1990), the crab *Portunus pelagicus* (Mortimer & Connell 1994, 1995) and the cladoceran *Ceriodaphnia* cf. *dubia* (Rose et al 1998).

3.1.1 Using QSARs to derive WQGs

The potential role for QSARs in deriving water quality guidelines is that they can provide estimates of the toxicity of chemicals for which there is either no or very little toxicity data. These estimates can then be used by the various methods for deriving WQGs or the equivalent (Van de Plassche et al 1993, OECD 1995). For example, the Dutch use QSARs (see Appendix 1) to estimate the toxicity of non-polar narcotic chemicals irrespective of the amount of experimental toxicity data available. These estimates are then substituted into the A&S method (Van de Plassche et al 1993) to derive WQGs. In contrast the OECD (1995) uses QSARs to estimate the toxicity of non-polar narcotic agents and polar narcotic agents. However, the OECD only uses QSARs when there is no experimental toxicity data. The use of these QSARs to predict the toxicity of narcotic chemicals is significant because they comprise approximately 60% of industrial chemicals (Veith et al 1983).

It is important that the quality of QSARs that are proposed for deriving WQGs be thoroughly assessed prior to their use and that they are of high quality. General prerequisites for QSARs to be of high quality are that they cover a wide range of log Kow values (or whichever physicochemical property is used in the model) and that they are based on a reasonable number of chemicals. These requirements ensure that the QSARs are representative of the chemicals they are attempting to model. There are two more specific measures of the quality of QSARs. The first is the quality with which the QSAR models the toxicity data used to

derive it, which is usually measured by the coefficient of determination (r^2). Generally QSARs with r^2 values greater than 0.8 are viewed as having sufficient quality. The second is the accuracy with which the QSARs predict the toxicity of chemicals not used to derive them (ie predictive quality). The predictive quality can be measured by the ratio of the toxicity predicted by the QSAR and the experimentally determined toxicity (Nabholz et al 1993). If the QSAR estimates the toxicity perfectly the ratio will equal one, while, if the ratio is less than or greater than one then the QSAR either overestimates or underestimates the actual toxicity respectively.

The QSARs developed by the USEPA have had their predictive quality assessed using the ratio method (Nabholz et al 1993). The acceptable ratios were taken to be between 0.1 to 10, which corresponds to the one order of magnitude error generally accepted for QSARs. This study of 920 individual QSAR estimates revealed that the ratio was between 0.1 and 10 in 85% of cases, was less than 0.1 for 9% and greater than 10 for 6% of the estimates. Unfortunately, this assessment only provides information on the overall predictive quality of the QSARs developed by the USEPA rather than information on the quality of individual QSARs (Nabholz et al 1993).

The QSARs used by The Netherlands (Van de Plassche et al 1993) and the OECD (1995) do not appear to have had their predictive quality assessed. They were, however, selected using a set of criteria that should ensure their general suitability and quality. Further, The Netherlands (Van de Plassche et al 1993) and the OECD (1995) must have considered these QSARs were of sufficient quality as they both recommend their use.

As mentioned earlier there are QSARs for many of groups of chemicals besides the non-polar narcotics. However, factors limit the usefulness of these other QSARs to derive WQGs. Firstly, even when rules are available to guide the choice of the appropriate QSAR they are not foolproof and an inappropriate QSAR may be used to predict the toxicity of a chemical. Secondly, there are fewer QSARs for these other mechanisms of action and most of them model acute toxicity rather than the chronic toxicity, which is preferred for deriving WQGs. It is therefore highly unlikely that a similar set of QSARs to that developed for non-polar narcotics could be compiled for chemicals with other mechanisms of action apart from polar narcotics.

Toxicity data derived from QSARs are only estimates of the actual toxicity and can have quite large errors. It could be argued that it is inappropriate to use QSARs in deriving WQGs. However, this review argues that a WQG derived using a large number of QSAR estimated toxicity data would be at least as likely to protect aquatic ecosystems as a WQG value based on a very limited number of experimental data points.

The main weaknesses or limitations of the QSARs are:

1. There are relatively large numbers of high quality chronic QSARs only for non-polar narcotics;
2. They are species and biological endpoint specific (ie each QSAR can only model a specific endpoint to a specific organism);
3. They are limited to either chemicals of similar structural characteristics or to chemicals with the same mode of action;
4. They are only valid within the range of the physicochemical properties of the chemicals used to derive the QSAR;
5. The choice of the appropriate QSAR to predict the toxicity of a chemical can be confusing and requires chemical knowledge;

6. The errors between the QSAR estimated toxicity and the measured toxicity of a chemical can be quite large.

The main strengths of QSARs for non-polar narcotic chemicals are:

1. They are simple models that predict toxicity using the logarithm of the octanol-water partition coefficient (log Kow);
2. They provide an estimate of the toxicity of chemicals and therefore permit the derivation of WQGs for chemicals with little or no toxicity data.

3.2 Methods for converting acute toxicity data to chronic data

There are four methods for converting acute toxicity data to chronic toxicity data. These are the generic acute to chronic ratio (Mount & Stephan 1967), the chemical specific acute to chronic ratio (Kenaga 1979) and two statistical extrapolation methods (eg Mayer et al 1994, Lee et al 1995, Sun et al 1995). The problems and limitations associated with the first two methods will be discussed in detail later in this review. Therefore this section will only examine the two statistical methods.

The statistical methods were developed because of the general lack of chronic data and the perceived problems with the acute to chronic ratio method (Mayer 1990). The simpler of the two methods is called the two step linear regression method as it uses linear regression analysis in two phases. In the first phase the probit of toxicity data for each exposure time is plotted against the logarithm of concentration and regressed. The LC0 or EC0 value for each time interval is calculated by substituting zero into the regression equation. In the second phase the LC0 or EC0 values for each exposure time are plotted against the reciprocal of the exposure time and regressed. The resulting equation is used to calculate the LC0 at an 'indefinite exposure' (Mayer 1990).

The second method uses multifactor probit analysis (MPA) to determine equations that describe the three dimensional shape of the plot of mortality versus concentration and time. This method requires more data than the two step linear regression method but allows the concentration that exerts a specific toxicological effect at a given exposure period to be calculated (Mayer et al 1994).

The predictive capabilities of both methods were assessed by Mayer et al (1994) in a study of the toxicity of 18 chemicals to 7 species. The LC0 values derived by the two step linear regression method were in good agreement with published maximum acceptable toxicant concentrations (MATC, ie the geometric mean of NOEC and LOEC values). Eighty-three per cent of the LC0 values were within a factor of 2 of published MATC values and all the LC0 values were within a factor of 3.2. Lee et al (1995) compared the LC0 values with published MATC values from 22 chronic studies and found that in 70% of cases the LC0 values, derived by the MPA method, were within a factor of 2 of the MATC values. It thus appears that both methods can convert acute lethality data to chronic lethality data that are in close agreement with publicised MATC data.

A potential problem for both these methods is that they assume that the plot of probit toxicity values versus the logarithm of concentration is linear. However, such a plot is typically sigmoidal (Gelber et al 1985) with the deviations from linearity occurring as the biological effects approach 0 and 100%. As the methods are attempting to estimate the concentration of a chemical that causes no effect then a linear regression model has the potential to incorrectly estimate the LC0. While this may lead to inaccurate estimation of the LC0 it does not invalidate the methodology.

The weaknesses of the methods developed by Mayer and co-workers are that they:

1. Require toxicity data from studies with at least five different concentrations and observations taken from at least four different intervals (eg 24 h, 48 h, 72 h, and 96 h);
2. Can only use acute lethality data to estimate chronic lethality and growth effects;
3. Assume the plot of probit toxicity values versus the logarithm of concentration is linear;
4. Have only been validated using fish.

The strengths of the methods are:

1. The LC0 values closely approximate MATC values for both lethality and growth;
2. There is no inter-species extrapolation as with the acute to chronic ratio method;
3. The two-step method is simple;
4. There are few assumptions;
5. There is a computer software program available (Mayer et al 1992).

4 Assessment of methods to derive water quality criteria

There are two main approaches to developing WQGs: the assessment factor methods and the statistical extrapolation methods. These methods will be examined individually.

4.1 Assessment factor methods

In these methods, the lowest reported toxicity value is divided by a constant that is variously called an assessment (AF), uncertainty, application or safety factor. The magnitude of the AF is governed by the perceived 'quality' of the toxicity data; ie the more environmentally realistic the toxicity data the smaller the AF and vice-versa. This approach for deriving WQGs was first proposed by Hart et al (1945) and was adopted from methods used in human health to derive average daily intakes (Cotruvo 1988, Calabrese & Baldwin 1993). The approach is typified by the method developed by the USEPA (1984b) and subsequently modified by Canada (CCREM 1991) and the OECD (1992a). The Canadian AF method is used by Australia and New Zealand (ANZECC 1992). However, the USEPA (1986), OECD (1995), The Netherlands (eg Van de Plassche et al 1993), Denmark (Bro-Rasmussen et al 1994) and South Africa (Roux et al 1996) all now use a statistical extrapolation method in preference to an AF method which is only used when there is insufficient data.

The magnitude of the various AFs, the type and magnitude of the extrapolations that are inherently assumed by the AFs used in the modified USEPA (OECD 1992a) and CCREM (1991) methods are presented in table 2.

The field to laboratory extrapolation has a value of ten (USEPA 1986, CCREM 1991, OECD 1995) and it accounts for the supposition that laboratory studies are likely to underestimate the toxicity in the field. Proposed reasons for this underestimation include: laboratory tests are conducted on animals that are robust and easily bred/maintained in the laboratory rather than 'sensitive' species; non-lethal endpoints are often difficult to detect but may be more sensitive indicators of the onset of detrimental effects; life stages not tested in the laboratory may be more sensitive to toxicants (Hart 1996) and all the limitations associated with SS toxicity tests that were discussed earlier. However, it is also possible for laboratory based experiments to overestimate the toxicity in field situations. This can arise because laboratory experiments

only use bioavailable forms of toxicants while in the field it is likely that only a small percentage of the chemicals are present in bioavailable forms.

Table 2 The assessment factors, types and magnitudes of the extrapolations used in the modified USEPA and CCREM methods

Available toxicity data	Type of extrapolation	Modified USEPA method[a]	CCREM method[b]
Chronic NOEC (for the USEPA) or LOEC (for ANZECC)	Field to laboratory	10	10
Acute LC50 or EC50	Field to laboratory & acute to chronic	100 (10 x 10)	ACR or 20 or 100[c]
Acute LC50 or EC50 for one of two species	Field to laboratory & acute to chronic & interspecies	1000 (10 x 10 x10)	ACR or 20 or 100[d]

a It is assumed toxicity data are available for at least an algae, a crustacean and a fish (OECD 1992a).

b Assumes that toxicity data are available for at least three species of fish of which two must be chronic; two invertebrates, one of which should be planktonic; and a freshwater vascular plant or algae (CCREM 1991).

c An AF of 50 is used for non-persistent chemicals while 100 is used for persistent chemicals when no ACR is available (Chapman 1995a).

d Where data are not sufficient to meet the requirements set in b then interim WQGs are derived (ANZECC 1992). It should be noted that although this method is available no interim WQGs have been derived by ANZECC.

The acute to chronic extrapolation is extensively used to derive WQGs because the vast majority of toxicity data are acute whereas chronic data are preferred for environmental protection. The CCREM method (CCREM 1991), like the original USEPA method (USEPA 1986), uses an acute to chronic ratio (ACR) derived from another species for the same chemical in preference to a generic ACR. When a chemical specific ACR is not available then CCREM (1991) and the USEPA (1986) use a generic ACR. CCREM (1991) uses an ACR of 2 or 10 depending on the environmental persistence of the chemical. While the modified (OECD 1995) and unmodified USEPA (1986) methods use one generic ACR of 10.

The USEPA (1986) and OECD (1995) apply an interspecies AF of 10 when the minimum toxicity data set requirements for the derivation of WQGs are not met. This is used because there is an increased amount of uncertainty in deriving WQGs from such a small sample size. In contrast, the CCREM (1991) method and hence the method used by ANZECC (1992) do not use an additional assessment factor to account for the increased uncertainty associated with using limited toxicity data.

4.1.1 Criticisms

Criticisms of the assessment factor approach revolve around the scientific validity of assessment factors, the type of toxicity data that should be used, the magnitude of the assessment factors, and whether or not the method is consistent with a risk framework and the principle of ecologically sustainable development (ESD).

4.1.1.1 Scientific validity

Numerous authors have acknowledged that AFs are arbitrary, have no theoretical scientific basis and are purely empirical (Hart 1974, Nicholson 1984, Kooijman 1987, Okkerman et al 1991, OECD 1992a, Schudoma 1994, Rand et al 1995, OECD 1995). Goldberg (1975) asserted that using assessment factors was tantamount to admitting that information essential for risk assessments was lacking. Nicholson (1984) considered that:

> the greatest difficulty in determining criteria from acute toxicity data is the choice of applications factors. There is little scientific basis for application factors except that they are the result of careful judgement… There is little evidence, in most cases, that the arbitrary value chosen is

indeed the best choice, ie whether a particular value for an application factor will provide 'adequate' protection and whether a less (or more) stringent value would be more appropriate.

The fact that there is no universally accepted magnitude for assessment factors (as seen in table 2) confirms their arbitrary nature.

Hart et al (1995) state that a major criticism of the Canadian and ANZECC approach is that 'the concept…is not based on whether the value actually protects 100% (of species) or even an estimate of that'. There is always error associated with any measurement yet there is no estimate of the error involved with the derivation of WQGs using the AF method. Thus there is uncertainty in the degree of protection that is offered and there is no idea of the margins of error involved in WQGs derived using the AF method.

Australia, Canada, New Zealand, South Africa and the USA use ACRs, the use of which has been criticised for a number of reasons. Firstly, the use of an ACR implies that the mechanism of action under acute and chronic exposures is the same but there is conflicting evidence on this point. Baird et al (1990) showed that some chemicals have different mechanisms of action for acute and chronic effects. Whereas Mayer and Ellersieck (1986) and Mayer et al (1994) found some chemicals had the same mechanism of action and that chronic effect levels could be predicted from acute data. Secondly, the ratio is obtained for one species and then applied to another, thus error can be introduced due to interspecies extrapolation (Calabrese & Baldwin 1993). Thirdly, the magnitude of the ACR can vary markedly with the biological endpoints that are being compared in order to derive the ACR and the degree of protection that is desired. The latter two points will be discussed in greater detail in following sections.

4.1.1.2 Type of data used

The AF method used by ANZECC (1992) uses LOEC data whereas the modified USEPA method uses NOEC data (OECD 1995). The use of LOEC data is unusual because as implied by the name, this is the lowest concentration at which toxicological effects can be statistically detected. NOEC and LOEC values are different. The size of this difference varies according to the arbitrary concentrations used in the toxicity test but should be less than three fold assuming that the toxicity tests conformed with OECD protocols. In order for LOEC data to be extrapolated to chronic field situations and thus offer the same degree of protection as NOEC derived WQGs, this difference between NOEC and LOEC data should be taken into account. Thus, the CCREM field to laboratory AFs should be larger (ie 15–30) than that used in the modified USEPA method. The fact that the USEPA and CCREM methods use the same AF (ie 10) is inappropriate and means that the CCREM method will offer less protection than the modified USEPA method. When the previous points are combined with the fact that NOEC and LOEC values are generally determined simultaneously NOEC data should be used whenever it is available in preference to LOEC data.

4.1.1.3 Conformity with other principles

The assessment factor method of deriving WQGs ignores all other data except the lowest and is therefore an example of the worst known case type of approach. Such a procedure is at odds with a risk-based approach, which requires an array of data in order to derive estimates of the probability of certain toxicological events occurring.

Risk based concepts and procedures are central to many of the more recently adopted scientific, social and political paradigms. Ecologically sustainable development (ESD) is the philosophical basis for the ANZECC water quality guidelines (ANZECC 1992) and as such the goal of the WQGs is to 'protect biological diversity…and maintain ecological processes and systems'. However, all human activity impacts on the environment and so it is impossible to protect and preserve all species. This fact was recognised by the Biodiversity Working

Party (1991) which stated that 'all development is likely to cause some loss of genetic component of biodiversity'. This implicitly implies that a certain level of environmental degradation, loss of genetic material and loss of species is acceptable as long as the integrity of the systems is maintained. By aiming to protect all species, the ANZECC WQGs fail to recognise that human activity always impacts on the environment and suggests that there are threshold concentrations for xenobiotic compounds below which no toxic effects will occur. There is no evidence in the literature to support the concept of toxicity thresholds (ie a concentration of a toxicant exists at which there is no toxicological effect on the test organisms) (Okkerman et al 1993, Emans et al 1993). Further, risk assessment theory does not support such a concept (Hart et al 1995). Therefore, the stated level of protection offered by the ANZECC WQGs should be reconsidered.

4.1.1.4 Magnitude of assessment factors

Another major criticism of the AF method is the magnitude of the various AFs. The validity of the magnitude of the current AF values will be discussed individually below.

The acute to chronic ratio

The AF used to convert from acute to chronic data can be either an acute to chronic ratio (ACR) or a generic AF. The CCREM method that is also used by ANZECC (1992) uses an acute to chronic ratio and it is inherently assumed that this extrapolates acute LC50 type data to chronic LOEC data. It has been shown that the magnitude of the ACR varies markedly with the biological endpoints used to derive the ACR (Suter et al 1987, Calabrese & Baldwin 1993). For example, Suter et al (1987) obtained ACR values of 31, 28 and 89 when comparing acute LC50 data with chronic EC25 (hatchability, adult mortality, and larval mortality) data respectively. It is pertinent to note that the ACR values reported by Suter et al (1987) are much larger than the generic values of 2 and 10 used by CCREM (1991) and ANZECC (1992) when chemical specific ACRs are not available.

The magnitude of ACRs also varies with the degree of protection desired (Calabrese & Baldwin 1993, Suter et al 1987, Kenaga 1982). For example, Kenaga (1982), in a study of the acute and chronic toxicity of 84 chemicals (which included inorganics, organics and pesticides) to nine species of fish and two invertebrates, derived 135 ACR values which ranged from 1 to 18 100 with a mean value of 10. The generic value used by the USEPA (1984b), the OECD (1992a) and the Dutch (Van de Plassche et al 1993) is numerically equal to the mean ACR obtained by Kenaga (1982). The Canadian (CCREM 1991) method which is used by Australia and New Zealand (ANZECC 1992) uses generic values of 2 and 10 for non-persistent and persistent chemicals respectively. Yet, the use of an ACR value of 10 would underestimate the chronic toxicity of 57% of the chemicals used by Kenaga (1982) for all tested species, 64% of chemicals to the fathead minnow and 47% of the chemicals to *Daphnia magna* (table 3). The use of an ACR of 2 would be even less protective. In order to increase the degree of protection a higher ACR would be required.

The use of two different generic ACR values by ANZECC and Canada, depending on the persistence of the chemical, is a worthy attempt to subdivide chemicals into smaller groups for which more pertinent ACR values could be derived. However, such an approach is not scientifically sound as Kenaga (1982) found that there was no association between the magnitude of the ACR and the persistence of the chemical. In fact there was no association between the magnitude of the ACR and a range of environmentally important properties of chemicals including bioconcentration factor and the octanol-water partition coefficient (Kenaga 1982).

Table 3 The distribution of 135 acute to chronic ratios for 84 chemicals derived by Kenaga (1982)

ACR range	Percentage of chemicals having ACR values in the given range		
	All species	Fathead minnow	*Daphnia magna*
1–9	43	36	52.8
10–99	43.7	50	33.3
100–999	11.9	10	13.9
1000–9999	0.7	2	0
10 000–99 999	0.7	2	0

In another study, Suter and Rosen (1988) found the ratio between acute LC50 values and maximum acceptable toxicant concentrations (MATC, which is the geometric mean between the LOEC and NOEC) for sheepshead minnow and a crustacean were 95% of the time less than 18.6 and 7.9 respectively. Slooff et al (1983) in a more extensive study found that for 95% of 164 chemicals the ACR was equal to or less than 25.6. While Kuhn et al (1989a,b) determined the acute to chronic ratio for 73 chemicals and found it had a mean value of 140 with a range from 2 to 3000.

It is apparent from the above studies that the ACR values of individual species vary markedly for different chemicals and therefore it is highly doubtful that the use of generic ACRs is valid. In fact, Hart (1996) recently cautioned against the use of generic ACR values when chemical specific ACR values were not available. Further, even if a generic ACR is to be used, the literature revealed that an ACR of 2 and 10 (used to derive some of the current ANZECC water quality guidelines) would lead to significant under protection of species from a range of chemicals. To overcome this Calabrese and Baldwin (1993) recommend a generic ACR value of 50.

The interspecies assessment factor

The interspecies AF is used to account for the toxicity data only being available for a very limited number of species that most probably do not represent the full range of sensitivities. The value used for this AF by the USEPA and OECD is 10. In contrast, the CCREM (1991) and ANZECC (1992) methods do not have any interspecies AF. The interspecies AF of 10 was adopted from the human health methods (Cotruvo 1988). A number of studies have determined the interspecies AF between humans and commonly used experimental mammals (ie dogs, monkeys, rabbits, guinea pigs, rats, hamsters and mice). Altman and Dittmer (1962) obtained interspecies AFs of up to 14.5. Evans et al (1944) found values between 2.5 and 152. Hayes (1967) obtained values of between 1.9 and 100 for acute data and 0.58 to 9.4 for chronic data and Krasovskij (1976) obtained differences between the most sensitive laboratory species and man of 1.5 to 3.4. The geometric means of all these studies were below 13. Thus, Dourson and Stara (1983) and Johannsen (1990) concluded that the available data supported an interspecies AF of ten. The relevance of this evidence to support the use of an interspecies AF of 10 in deriving WQGs is questionable for several reasons. Firstly, the aim of human health measures is not to protect all humans—they permit a percentage of the population to suffer potential effects whereas the ANZECC WQGs aim to protect all aquatic life forms. Secondly, the geometric mean of the interspecies AFs was used to validate the use of an interspecies AF of 10. However, using a mean value will, in nearly all cases, not protect the more sensitive species. If all species are to be protected then interspecies AFs based on the mean should not be used, rather the maximum recorded interspecies AF should be used. Thirdly, there were only a limited number of species used to derive these interspecies AFs (ie 8

in Altmann & Dittmer 1962) and they were all mammals. In aquatic environments there are many more species and more widely different species that require protection. Both of these factors would greatly increase the probability that larger interspecies AFs will be obtained for aquatic ecosystems.

A number of studies have determined interspecies AFs for a variety of aquatic organisms (eg Kenaga 1978, Kimerle et al 1983, Le Blanc 1984). The findings of Kimerle in a study of 82 chemicals are summarised in table 4. These findings have been used to support an interspecies AF of 10. Interspecies AF values should be the difference in toxicity between the most and least sensitive species. Whereas, in this study the *Daphnia* and rainbow trout are known to be sensitive to a wide range of pollutants and the values are derived by comparing these to the most sensitive species. Therefore, the values cited by Kimerle et al (1983) are not interspecies AFs and the real interspecies AF would be larger than indicated.

Table 4 The distribution of differences between four species (*Daphnia* sp, fathead minnow, bluegill and rainbow trout) and the most sensitive species for 82 chemicals

Organism	Percentage of chemicals for which the magnitude to the most sensitive species is	
	10	100
Daphnia sp	76%	93%
Fathead minnow	47%	74%
Bluegill	52%	70%
Rainbow trout	74%	91%

The most extensive study on interspecies AFs was conducted by Slooff et al (1986). It examined the acute toxicity of 15 different chemicals to 35 species of freshwater organisms. Figure 1 presents the distribution of interspecies AFs calculated in the study. If 95% of the species tested by Slooff et al (1986) were to be protected then an interspecies AF of approximately 1000 would be necessary (Calabrese & Baldwin 1993). Assuming the Slooff et al study holds true in general, then an interspecies AF of ten would underestimate the sensitivity of approximately 90% of species.

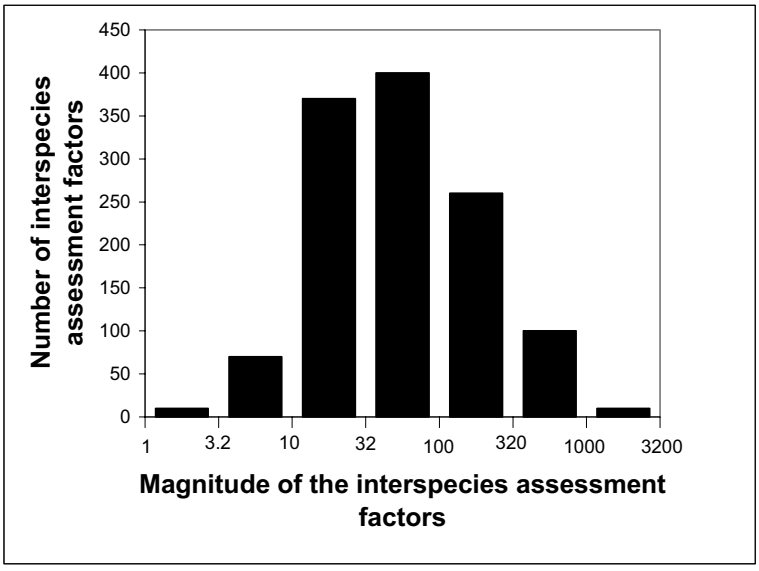


Figure 1 The distribution of interspecies assessment factors derived by Slooff et al (1986)

A potential improvement on the use of generic interspecies AF values is the use of AFs based on the phylogenetic relatedness of the species being compared. A number of studies have shown that there is a positive relationship between the phylogenetic relatedness of species and the magnitude of the interspecies AF (Le Blanc 1984, Barnthouse et al 1990, Calabrese & Baldwin 1993) (table 5).

Table 5 The interspecies assessment factors for a range of different interspecies extrapolations that should protect 95 and 99% of the species

Type of interspecies extrapolation	Interspecies assessment factors to protect a theoretical percentage of all species	
	95%	99%
Species within genus	10.0	16.3
Genera within family	11.7	16.9
Families within order	99.5	145.0
Orders within class	64.8	87.5
Classes within phyla	1000[1]	

[1] Obtained from Slooff et al (1986);all the other interspecies AF values were obtained from Calabrese and Baldwin (1993). Calabrese and Baldwin (1993) recommended the use of this value.

The current system for deriving WQGs, used by USEPA, CCREM and ANZECC relies on data from a very limited range of organisms, typically fish, crustaceans and algae. Yet the aim is to protect all aquatic life forms which will involve extrapolations between different kingdoms and therefore interspecies AFs of greater than 145 (Calabrese & Baldwin 1993) will be required to protect 99% of species and higher values to protect all species.

Given the above, the fact that the method used by ANZECC (1992) does not use an interspecies AF is a major failing. It is most probable that this method will derive WQGs that do not provide adequate protection to the environment.

4.1.2 Does the assessment factor method used by ANZECC provide the desired level of protection?

The stated aim of the ANZECC WQGs is 'to protect all forms of aquatic life and all aspects of the aquatic life cycle… The intention is to protect all life forms during indefinite exposure to the water' (CCREM 1991 cited in ANZECC 1992).

The preceding text highlighted that the AFs used for interspecies and acute to chronic extrapolations are too small to protect all species, which is the aim of the ANZECC WQGs. Only one study (Napier 1992) has addressed the issue of whether or not the ANZECC WQGs deliver the stated level of protection. This study assessed the biological recovery of a creek downstream of several abandoned mines. It found that several species not present in the creek were present in adjoining tributaries that acted as control sites, despite the concentrations in the creek being below the ANZECC guidelines. This indicated that some species had been eliminated and thus the criteria did not provide the stated level of protection. There are two possible causes of this lack of protection: the assessment factors used are too small and/or that the WQGs are based only on the direct effects of individual chemicals. From the study by Napier, it is not possible to determine which is the dominant factor, although the latter definitely contributes as the river was subject to a mixture of metal pollutants. However, even if this is the reason for the failure of the ANZECC WQGs to protect all the species, simultaneous exposure to multiple toxicants is normal in the environment. The failure to address the toxicity of mixtures is a major failing of the current ANZECC WQGs which will be addressed later in this review.

4.1.3 Strengths and weaknesses

The strengths of the assessment factor methods are:

1. They are very simple to use;
2. They are easily understood;
3. The magnitude of the AFs can easily by modified to reflect new toxicological findings (Johannsen 1990). However, while the data strongly indicate that the AFs should be larger than the current values there were insufficient data to modify them.

The weaknesses of the assessment factor methods are:

1. The assessment factors have no theoretical basis, they are purely empirical;
2. There is debate over the scientific validity of acute to chronic ratios;
3. The current generic assessment factors have been shown to be too small;
4. From the one field study it does not appear that the ANZECC WQGs protect all aquatic life forms;
5. The method is not transparent as it does not state the degree of protection provided by an AF of a certain magnitude and thus does not permit informed decisions and debate over the level of protection to occur.

4.2 Statistical extrapolation methods

An alternative approach to the assessment factor method has been developed over the past ten years. The methods differ from the assessment factor approach because they are statistically based, offer a different degree of protection, and a measure of the associated uncertainty. Briefly, all these methods use toxicity data obtained from tests on individual species and fit a statistical distribution to the data to derive an aqueous concentration that should protect 95% of the species in the environment. However, all the statistical extrapolation methods can calculate different levels of protection.

There are three different extrapolation techniques: the Stephan et al (1985) method; the Aldenberg and Slob (1993) method (A&S) which is an enhancement of the Kooijman (1987) and Van Straalen and Denneman (1989) methods; and the Wagner and Løkke (W&L) method (1991). The techniques vary in the data they use and the assumed distribution of species sensitivity to toxicants. Of the three extrapolation techniques this review will only examine the A&S method. This was done for a number of reasons. Firstly, the OECD examined both the A&S and the W&L methods and recommended the former. The A&S method has subsequently been adopted by The Netherlands and the OECD. Secondly, far more validation work has been conducted on the A&S than the W&L method. Thirdly, the Stephan et al method was excluded because it: assumes there is a threshold toxicity value below which no detrimental effects will occur and the scientific literature and risk assessment theory does not support such an concept (Okkerman et al 1991, OECD 1992a, Emans et al 1993, Pedersen et al 1994, NZ Ministry of the Environment 1996); uses an arbitrary assessment factor of two without any justification (Hart et al 1995, NZ Ministry of the Environment 1996); assumes that ecosystems can tolerate high concentrations for short periods of time; the two WQ criteria system is not practical requiring extensive sampling over prolonged period of time (NZ Ministry of the Environment 1996); and it has extensive data requirements ie acute toxicity data from species belonging to at least eight different taxonomic groups and chronic toxicity data for species belonging to at least three different taxonomic groups (Delos 1995). These limitations are discussed in more detail in Warne (1996) and Hart et al (1995). Another

pertinent reason for not considering the Stephan et al method is that the USEPA itself recognises that the method requires updating (Delos 1995). It is, therefore, surprising to note that South Africa is currently using the Stephan et al method (1985) combined with the USEPA assessment factor method to derive WQG for freshwater (Roux et al 1996).

4.2.1 The Aldenberg and Slob method

The aim of this method is to determine concentrations of toxicants that should protect any chosen percentage of species in a compartment of the environment. Typically, the concentration that should protect 95% of the species is calculated and this is termed the concentration hazardous to 5% of the species (HC5). The method assumes that the sensitivities of species to toxicants has a logistic distribution (fig 2). The first step in the method is therefore, to test how well the data fit the logistic distribution using the Kolmogorov-Smirnov test (D'Agostino & Stephens 1986). Providing the data are not significantly different from a logistic distribution they can be used (Emans et al 1993).

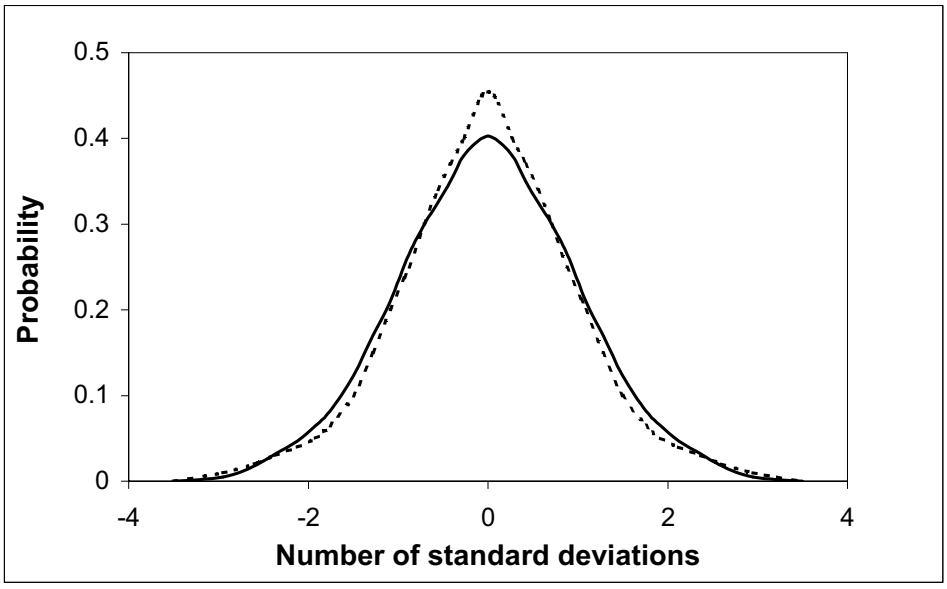


Figure 2 Distribution curves of the sensitivities of species. The solid line represents the normal distribution and the dashed is the logistic distribution.

Assuming the distribution of sensitivities of all species in an ecosystem is known and has a logistic distribution, the HC5 is calculated using the equation:

$$\log HC5 = \mu - (K\beta \cdot \sigma) \quad (3)$$

where μ is the mean toxicity of all the species in the ecosystem (eg moles/L, mg/L); $K\beta$ is an adjustment factor that equals 1.62 for a logistic distribution of sensitivities and when toxicity data are available for all species; and σ is the standard deviation of the toxicity values for all species in the ecosystem.

Equation 1 assumes that toxicity data is available for all species in an ecosystem. Therefore, it calculates the true (ecosystem) HC5 value, as all the parameters are known. There is no chemical for which toxicity data are available for all species of an ecosystem. Equation 1 must therefore be modified to equation 2 in order to account for the data being a sample of all toxicity data for an ecosystem.

$$\log HC5 = x - (K\beta \cdot s) \quad (4)$$

where x is the mean toxicity value for all the species in the sample, Kβ is an adjustment factor, and s is the standard deviation of the toxicity values for all species in the sample.

As the sample mean and standard deviation values of equation 2 are estimates of the true (ecosystem) mean and standard deviation values then the calculated HC5 value is an estimate of the true (ecosystem) HC5 value. There is error associated with the estimated HC5 value and thus if a number of different samples (fig 3), containing the same number of species, were used to derive HC5 values, each data set would produce its own estimate of the true HC5. Thus, there would be a distribution of the estimated HC5 values around the true HC5 value (fig 3).

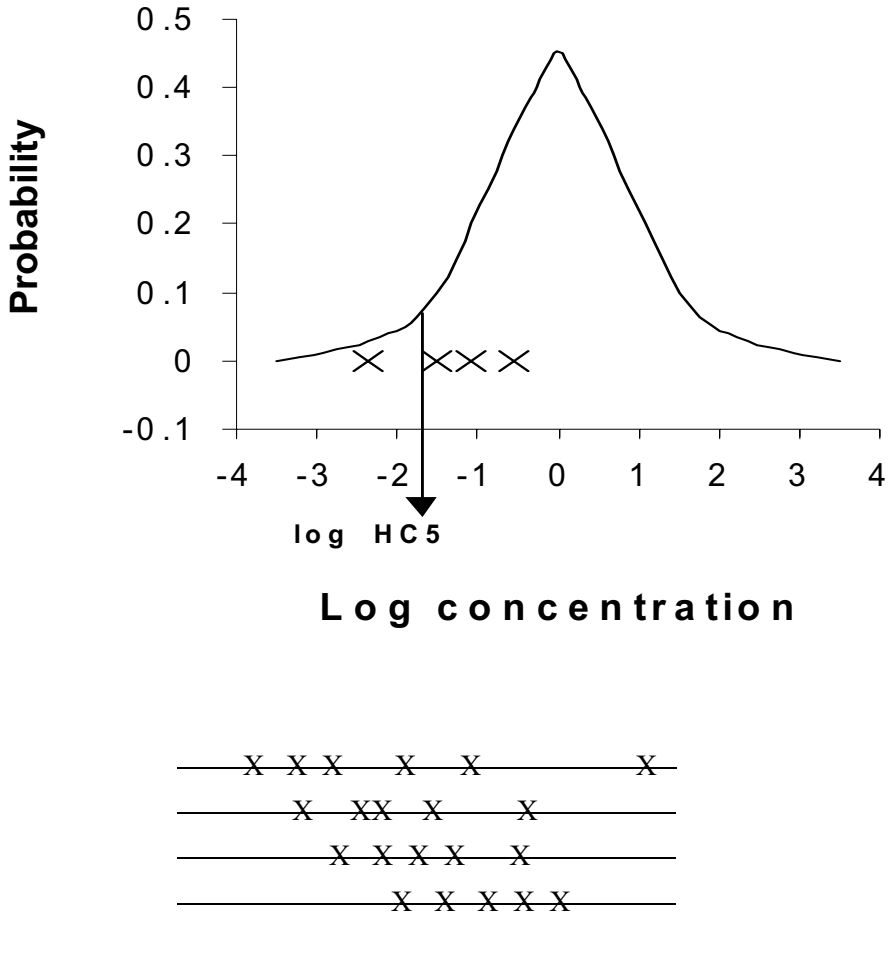


Figure 3 The curve represents the distribution of toxicity for a chemical on all species, with the actual HC5 value indicated. The 'X's represent log HC5 estimates from four sub-samples of the toxicity data as shown below the figure. This illustrates how HC5 estimates are distributed around the actual HC5 value for all species.

As there is error associated with any HC5 value derived from a sample it is desirable that the HC5 value protects the desired percentage of species in the vast majority of cases. The term Kβ adjusts the HC5 value to ensure that it has a certain probability of over and underestimating the true (ecosystem) value. The β term denotes this probability and is either 50% or 95%. The Kβ

values were calculated using Monte Carlo simulation techniques to ensure the HC5 values would over and underestimate the true (ecosystem) HC5 value by the stated amount (ie 50% over and underestimation for HC5 values with 50% confidence limits and 5% overestimation and 95% underestimation for HC5 values with 95% confidence limits).

Principally for these reasons, Aldenberg and Slob (1993) advocated the use of HC5 values with 95% confidence limits as the 'safe' concentration of toxicant to the ecosystem (ie a concentration sufficiently low that it should maintain ecosystem form and function). They also recommended the HC5 values with 50% confidence limits as the best estimate of the 'hazardous concentration' respectively (ie a lower degree of protection that was chosen by the Dutch policy makers as a compromise between environmental protection and cost).

The A&S extrapolation method does not utilise any assessment factors in the normal sense of the word. However, the magnitude of the HC5 value depends on the standard deviation(s) of the sample data sets used and the Kβ value. The more variable the toxicity data the larger the value of s and the smaller the HC5 values will be.

The method can also be used in a reverse manner to determine the number of species that will not be protected (q) by a concentration (C) of a toxicant in the environment. This is done by fitting a logistic curve to the data using the equation:

$$q = 100\{1-[1 + \exp(A)]^{-1}\} \quad (5)$$

where A is calculated by

$$A = [\pi^2 (x_m - \ln C)] \div (3 s_m K\beta) \quad (6)$$

Where x_m is the mean toxicity of all species in the sample, C is the concentration of toxicant in the environment, s_m is the standard deviation of the natural logarithm of the NOEC values and Kβ is an adjustment factor.

Equation 5 could be useful in ecological risk assessments, site-specific investigations and in assessing the potential benefit of various remediation strategies. As the tails of species sensitivity distributions are large, a relatively large increase or decrease in chemical concentration may be permitted before significantly altering the level of protection being provided. Transferring the chemical concentration in the environment to the corresponding degree of protection will facilitate decision-making concerning the relative risk of a particular site or particular pollution event.

4.2.1.1 Criticisms

The Aldenberg and Slob method makes a number of assumptions the validity of which will be discussed in turn.

Assumption One: The ecosystem is sufficiently protected by protecting a given percentage of the species comprising that particular ecosystem.

Forbes and Forbes (1993), Smith and Cairns (1993) and Schudoma (1994) stated that there was no evidence to support the concept that protecting 95% of species would maintain ecosystem function. Forbes and Forbes (1993) further state that 'it seems reasonable to suggest that there are species whose removal would affect ecosystem processes' however, they themselves admit there is no evidence to support such a view. The work of Emans et al (1993) and Okkerman et al (1993) showed that for the data available, protecting 95% of the species would protect all the species in the mesocosms studied. This result does not prove that protecting 95% of the species maintains ecosystem function but it does support the concept and shows that it provides an environmentally realistic level of protection.

Calabrese and Baldwin (1993) stated that it would be difficult to reconcile the concepts of keystone species and WQGs designed to only protect 95% of species. The same problem would apply to commercially important or threatened species. However, this potential problem can easily be overcome. When toxicity data for such species are available they can be included in the derivation of WQGs and therefore it can be readily known whether or not the resulting WQG will protect these species. If they do not protect the species then the level of protection can be altered so they are protected.

Assumption Two: The distribution of species sensitivities in ecosystems is closely approximated by the log-logistic distribution.

The choice of distribution of species sensitivity has been criticised on the basis that species sensitivity to toxicants may not be logistically distributed. However, Kooijman (1987) found the distribution of species sensitivity for fourteen different chemicals confirmed the assumption. Also, the method is only used on data that passes the Kolmogorov-Smirnov test (D'Agostino & Stephens 1986) and therefore has a distribution not statistically different from the logistic. However, the power of the Kolmogorov-Smirnov test decreases with the number of data used. Therefore, with decreased amounts of data, there is an increased probability of incorrectly accepting that the data fits the log-logistic distribution. The use of the logistic distribution is further strengthened because it has been shown that the type of distribution used in the various statistical extrapolation methods has little effect on the magnitude of the WQGs (OECD 1992a).

Chemicals such as pesticides that are species specific may have a bimodal distribution (Aldenberg pers comm). The use of such bimodal data in the A&S method would lead to a very large standard deviation and hence unrealistically low HC5 values. This can be overcome by only using the A&S method on the more sensitive group of organisms thus minimising the standard deviation and giving more environmentally relevant WQGs. While such a modification may protect more than 95% of species using the data for all species would lead to an even greater degree of protection.

Assumption Three: The distribution of sensitivities of laboratory animals closely approximates the distribution of sensitivity of species from natural ecosystems.

Smith and Cairns (1993) cited theory and the work of Baird et al (1990) to suggest that laboratory species may have a narrower range of sensitivity than field species. The work of Baird et al (1990) showed that different clones of a species have different sensitivities to cadmium. However, these differences between clones do not necessarily translate into differences between field and laboratory species. In fact, Smith and Cairns (1993) admitted there were no experimental data to support their claim. Contrary to the suggestion of Smith and Cairns (1993), Emans et al (1993) based on their experimental data found no evidence to support the notion that the range of sensitivity of species was different in the laboratory and in the field.

Assumption Four: Interactions between species living in ecosystems can be ignored.

This assumption is inherently made by all methods that use single species toxicity data to derive WQGs. There is evidence supporting the use of single species toxicity data to derive WQGs (Marcus & McDonald 1992, Emans et al 1993, Okkerman et al 1993, Parkhurst 1994). However, multiple species toxicity data that take interspecies interactions into account are preferred (USEPA 1984a, Van de Plassche et al 1993, OECD 1992a, 1995). This issue is examined in more detail in section 2.1.1.

Assumption Five: Toxicity data are derived from independently conducted tests on randomly selected species from the ecosystem.

Criticisms of the second part of this assumption do have merit. Toxicity tests have concentrated on a relatively limited number of well known species and this is likely to continue in the future though to a lesser extent as more tests are developed. This problem was highlighted by Seegert et al (1985) who examined the National (USA) toxicity database for NOEC toxicity data on 21 priority pollutants to 99 species. They found that only 6% of freshwater fish species of continental USA had been tested; 40% of the data came from 11 test species; and 25% of the data came from just two families of organisms *Salmonidae* and *Daphnidae*. The situation was even worse for LC50 data. The Aldenberg and Slob method attempts to account for working with a sample by calculating 50 and 95% confidence limits for the HC5 value. However, it should be pointed out that this only accounts for statistical uncertainty and that the ability to test the predictions of the model is limited by the lack of available toxicity data.

Assumption Six: Species will not be exposed to mixtures of chemicals.

Criticisms of this assumption are valid. However, this criticism is valid for all methods used to derive WQGs, including the assessment factor approach. How this issue can be resolved will be discussed later in this review (section 8.4).

Assumption Seven and Eight: (7) The toxicological endpoint used is appropriate for measuring safety; (8) The method yields environmental quality guidelines that are conservative by nature.

These hypotheses are closely related and can be best dealt with by discussing them simultaneously. Smith and Cairns (1993) do not favour the use of NOEC data from single species tests and suggest that multiple species tests (field studies, mesocosms and microcosms) would provide much more useful data from which to derive WQGs. While there is merit in this there are insufficient data of this type and therefore WQGs could only be derived for a limited number of chemicals. This means that single species toxicity data will be required to derive WQGs for the vast majority of chemicals .

The conservative nature and appropriateness of using NOEC data in the A&S method can be determined by comparison with WQGs derived using multiple species data and the assessment factor method. Water quality guidelines derived from multiple species toxicity data are equal to the lowest recorded NOEC for the species in the test system (section 4.3). Emans et al (1993) conducted an extensive literature search for toxicity data and then compared the lowest multiple species NOEC values with HC5 values derived using the A&S method. They found the HC5 values with 50% confidence limits were always lower or equal to the multiple species NOEC values while the HC5 values with 95% confidence limits were much lower than the multiple species NOEC values. A similar finding was obtained by Okkerman et al (1993). There are in fact, disadvantages to the multiple species toxicity tests, which were examined in detail in section 2.1.1. It is possible that NOEC and LOEC type data may be replaced in the foreseeable future by EC5, EC10 or other similar low effect levels (refer to section 2.1.4). However, the A&S method is sufficiently robust that should another measure of toxicity replace the NOEC it will not affect the performance of the method.

Forbes and Forbes (1993) compared HC5 values with 95% confidence limits and those derived using the modified USEPA method for eight compounds both of which had been published by the OECD (1992a). They found there were no significant differences ($p = 0.116$) between the values and concluded the A&S method was not inherently conservative. However, it is a small

sample on which to base any conclusion. The work of Emans et al (1993), Okkerman et al (1993) and Schudoma (1994) all support the validity of assumption eight.

A comparison of WQGs derived using the A&S method and the modified USEPA AF method for forty chemicals was conducted as part of this review. The data used for this comparison were NOEC data. Strictly, LOEC data should have been used to derive the ANZECC WQGs (ANZECC 1992), however, LOEC data were not readily available. The comparison of the WQGs, though, was not hindered by the use of NOEC data as ANZECC WQGs derived using LOEC data would be up to three times larger than those derived using NOEC data. Thus, if there were significant differences between the A&S and ANZECC derived WQGs based on NOEC data, the differences would be even larger if LOEC data were used. The WQGs were compared using the same statistical test as Forbes and Forbes (1993) and it was concluded that the HC5 95% values were significantly smaller (ie more protective) than the AF values used by ANZECC. This is discussed in more detail in section 6 of this review.

Other criticisms

A weakness of the A&S method (and indeed of all statistical extrapolation methods) is that it is very sensitive to the data when there are a limited number of data. With small data sets there is increased uncertainty about how well the sample reflects the population. Therefore, larger $K\beta$ values are used (Appendix 2) and the WQG values generally become smaller. A study by the Danish EPA (Pedersen et al 1994) indicated that WQGs derived using data sets containing less than 5 values were very dependent on the spread of the values, whereas for data sets containing five or more values this effect was markedly reduced. The Danish therefore recommended that this method should not be used on small data sets.

Another limitation of the A&S method is that it uses NOEC data of which there is a limited amount. There is vastly more LC50 and EC50 data available. The lack of NOEC data therefore limits the numbers of chemicals for which WQGs can be derived using this method. A pragmatic resolution to this problem is to combine the A&S method with a method of converting acute to chronic date. This could be done by using either of the methods developed by Mayer and co-workers (eg Mayer et al 1986, 1994), acute to chronic ratios or generic AFs. In this way an acute LC50 or EC50 data could be converted to chronic NOEC data which would then be used in the Aldenberg and Slob method. This modification of the A&S method was proposed by Dr John Chapman. Dr Van de Plassche of the National Institute for Public Health and Environmental Protection in The Netherlands (pers comm) had only one reservation to the new proposal which was that it was a hybrid and therefore there was no clear distinction between the A&S method and the AF methods.

Despite the above criticisms and limitations of the A&S method, both Okkerman et al (1991, 1993) and Emans et al (1993) concluded the A&S method was better than the modified USEPA AF method (1984b) as they felt it had a more sound scientific basis. Calabrese and Baldwin (1993) also recommended the Van Straalen and Denneman approach (and hence the A&S method) over the assessment factor method as it represented ‘a theoretically sound approach for the derivation of chemical specific ecosystem MATCs’ and that statistical extrapolation methods in general ‘offer important advantages’ over the other approaches. Since Calabrese and Baldwin came to their conclusion, the work of Emans et al (1993), Okkerman et al (1993) and Schudoma (1994) has provided further evidence to support the A&S method. Similarly, the USEPA uses the Stephan et al (1985) statistical extrapolation method in preference to assessment factor methods (OECD 1995). More recently, Hart et al (1995) also recommended that New Zealand use the A&S method in preference to the assessment factor method and the OECD (1995) recommended the method to derive

maximum acceptable concentrations (MATCs) which are the equivalent of the WQGs of Australia and New Zealand.

4.2.1.2 Does the Aldenberg and Slob method provide the desired level of protection?

No studies have been conducted with the aim of determining what percentage of species is protected by the HC5 values derived by the A&S method. Okkerman et al (1993) and Emans et al (1993), however, compared the HC5 values derived using single species toxicity data with WQGs derived using the lowest NOEC from multiple species (MS) tests for 13 and 18 chemicals respectively. Both studies found that the HC5 50% values for all chemicals were equal or lower than the MS NOEC values and that the HC5 95% values were much lower than the MS NOEC values.

It is generally agreed that toxicity data from multiple species tests (ie microcosms, mesocosms and field studies) are more environmentally realistic than single species (SS) data (see section 2.1.1). However, the lowest MS NOEC value may not protect all species. In fact, this is unlikely due to the small number of species in these tests compared with real ecosystems. As the HC5 values currently available are smaller than the MS NOEC data for the same chemical, they offer an environmentally realistic degree of protection ie a level of protection that accounts for all the direct and indirect effects of the toxicant within that multiple species test system.

4.2.1.3 Strengths and weaknesses

The method has a number of strengths:

1. It uses single species toxicity data which comprise the vast majority of toxicity data available and are much cheaper and more rapid to obtain than multiple species toxicity data;
2. It uses toxicity data for all species which are available, thus conforming to risk assessment principles, rather than using only the lowest toxicity value as in the assessment factor method;
3. It has a sound statistical basis providing the assumptions of the method are met;
4. The HC5 values with either 95% or 50% confidence limits are equal to or less than criteria derived using multiple species tests;
5. It is a flexible method, it can use any measure of toxicity and can calculate HC values to protect any chosen percentage of species except 0% and 100%;
6. Points 4 and 5 mean that the method is transparent and it allows both the level of protection and uncertainty in the chosen level of protection to be chosen. The approach also enables a more informed debate to occur over the level of protection to be offered.
7. It can be used in the reverse manner to determine what level of protection (ie percentage of species) is offered when a certain concentration of a chemical occurs in the environment. This should be useful in ecological risk assessments and site-specific investigations.
8. Several aspects of the methodology have been validated.

The weaknesses of the method include:

1. The data requirements may limit the number of WQGs that can be derived;
2. It uses NOEC data and there is currently debate in the literature over the suitability of such data for regulatory purposes;
3. It is more complex to understand how the WQGs are derived than with the assessment factor method;
4. Several of the assumptions of the method may be compromised.

4.3 Multiple species test method

This approach uses toxicity data from microcosm, mesocosm and field experiments. Environmental concern levels and WQGs derived by the OECD (1995) and The Netherlands (Emans et al 1993, Okkerman et al 1993, Van de Plassche pers comm) using this method are equal to the NOEC value for the most sensitive species tested.

While multiple species (MS) tests are more environmentally realistic than single species tests (see section 2.1.1) they are still relatively simple systems that contain a limited number of species and different types of organisms. It is unlikely that the ecosystems in micro and mesocosms will be as complex, contain as many species or have all the possible food chains and nutrient pathways as the real environment. Therefore, it is uncertain that a WQG equal to the lowest measured NOEC from a MS test will provide sufficient protection to the environment. When the Precautionary Principle (see section 5.1) is considered, then the WQG should be less than the lowest measured NOEC. A WQG lower than the lowest NOEC could be obtained by entering the NOEC data from all species tested into both the A&S and AF methods.

5 Assessment of the methods in terms of the Precautionary Principle

5.1 The Precautionary Principle

The Precautionary Principle was devised by German bureaucrats in 1965 as a means of allowing policy decisions to be made when there was not adequate scientific information. It was termed 'Vorsorgeprinzip' or the foresight principle (Kettle 1993). When there is uncertainty concerning the available scientific evidence, science cannot answer questions of how to deal with environmental issues. Under such circumstances the Precautionary Principle may be implemented, as it allows the necessary policy decisions to be made. Thus, the Precautionary Principle can not be scientific and 'accepts a non-scientific basis for decision making' (Cameron 1993).

The principle is widely used at all levels of government particularly in the international sphere, eg the Ministerial Declaration of the Second International Conference on the Protection of the North Sea, the Baltic Sea Declaration, Treaty of Rome, the Rio Declaration on Environment and Development, the UN Framework Convention on Climate Control and the London amendment of the Montreal Protocol (Cameron 1993). In a local context, the principle has been included in the Australian Intergovernmental Agreement on the Environment signed by the federal and all state governments (Lipman pers comm). It has also been incorporated into the Australian National Water Quality Management Strategy (ANZECC & ARMCANZ 1994), the Australian National Strategy for Ecologically

Sustainable Development (ESD Steering Committee of the Department of Prime Minister and Cabinet 1992), and the New Zealand Resource Management Act (New Zealand Ministry of the Environment 1996).

The most commonly used wording for the Precautionary Principle and that adopted in this review is:

> that if there are threats of serious or irreversible environmental damage, lack of full scientific certainty should not be used as a reason for postponing measures to prevent environmental degradation (section 6 (2) (a) *NSW POEA Act 1991*).

The level of risk of environmental damage that is permitted before implementation of the Precautionary Principle is subject to interpretation. Some argue that inaction is unjustified if there is any risk to the environment. However, all human activity, even in the simplest forms of society, entails environmental impacts. Thus, we are forced, if we accept the need for human survival, to accept certain levels of risk to the environment. To do otherwise 'introduces a Utopian element into the Precautionary Principle which can not be sustained' (Cameron 1993). Hence, it is generally accepted that risks of non-negligible or serious and irreversible damage are required before inaction to protect the environment is unjustified (eg UNCED 1992). Some countries have reduced the extent of precaution in the Precautionary Principle by increasing the degree of certainty required before implementing the principle. For example, the French require a high degree of certainty of environmental damage. The British only invoke the Precautionary Principle after damage has been demonstrated (Cameron 1993) and it therefore would more correctly be named the 'hindsight principle'.

Opponents of the Precautionary Principle argue that the implementation of the principle may impose an unnecessary burden on industry and therefore indirectly on society. They argue that being overly cautious and erring on the side of safety increases costs to industry that must meet these criteria. The Precautionary Principle overtly admits that this may occur; however, it inherently assumes that it is better to overestimate the potential dangers to the environment than it is to underestimate them.

5.2 The degree of precaution in the methods for deriving water quality guidelines

The AF method used by ANZECC (1992) basically takes the most sensitive toxicity value (ie the worst case) and divides this by an assessment factor. The selection of the assessment factor used to derive WQGs is precautionary in nature ie the greater the uncertainty of the relevance of the laboratory data to the field the larger the assessment factor used to derive the WQG. However, the assessment factors contain little scientific basis except that they are the result of careful empirical judgement (Hart 1974, 1982, Sprague 1976, Skidmore & Firth 1983, Nicholson 1984, Calow 1995). The aim of the ANZECC WQGs to protect all aquatic life forms is definitely precautionary. However, the work presented earlier in this review showed that the ANZECC WQGs examined did not appear to protect all aquatic life forms. Therefore it is apparent that, in some cases at least, the assessment factors applied were not precautionary.

The A&S (1993) method assumes that the species sensitivity has a log-logistic distribution whereas many, but not all, biological characteristics have a normal or Gaussian distribution (Wagner & Løkke 1991). Using the same data the use of the log-logistic distribution will lead to the calculation of WQGs with lower concentrations than those calculated by other extrapolation techniques that assume a log-normal or log-triangular distribution (Aldenberg & Slob 1993). Thus the method is more precautionary than other statistically based extrapolation methods. The A&S method is also precautionary in two aspects of the

calculation of WQGs (equation 4). Firstly, the larger the variation in toxicity data the higher the standard deviation and the lower the subsequent WQG and vice versa. Secondly, as the number of species for which toxicity data are available decreases, the uncertainty of the sample being a true representation of all species increases. Thus, the adjustment factor ($K\beta$) increases (see Appendix 2) and the resulting WQG is reduced and vice versa.

Both methods for deriving WQGs contain a number of elements that are precautionary in nature. However, it is not possible to quantify the degree of precaution and therefore it is not possible to determine which method is more precautionary. Both methods, however, fail to consider a number of phenomena that are known to occur for at least some toxicants. These include:

- The combined toxicity of mixtures;
- Bioaccumulation or 'secondary poisoning';
- The environmental fate of the chemicals (ie degradation and transport between different compartments of the environment).

As both the methods omit these phenomena they can not be precautionary.

6 Comparison of water quality guidelines derived using the modified USEPA assessment factor and the Aldenberg and Slob methods

A comparison of WQGs derived using the modified USEPA assessment factor (AF) method and the Aldenberg and Slob (A&S) method for eight chemicals was conducted by the OECD (1992a) (table 6).

Table 6 Water quality guidelines derived using the modified USEPA assessment factor method and the Aldenberg and Slob method (modified from OECD 1992a)

Chemical name[1]	Modified USEPA method (lowest NOEC/10)	Aldenberg and Slob method	
		HC5 50%	HC5 95%
K_2CrO_7	0.01	0.071	0.010
NaBr	1.0	4.1	0.21
TPBS	0.032	0.27	0.042
2,4-DCA	0.0032	0.064	0.0082
p-NT	0.032	0.45	0.10
DNOC	0.0032	0.031	0.0028
Pentachlorophenol	0.00032	0.0043	0.00033
Dimethoate	0.0032	0.018	0.00036

1 TPBS is tetrapropylene benzene sulfonate; 2,4-DCA is 2,4-dichloraniline; p-NT is para nitro-toluene; and DNOC is dinitro-orthocresol.

The AF method used is very similar to the method used by ANZECC (table 2) the key difference being that the USEPA method uses NOEC data rather than LOEC data. Statistical comparisons of these WQGs, using the non-parametric Wilcoxon signed ranks test, showed that the HC5 50% values derived using the A&S method were significantly larger than WQGs derived using the AF approach. In contrast, there were no significant differences between HC5 95% values and the WQGs derived using the AF method. Thus, the assessment factor

method WQGs evaluated were more precautionary than the HC5 50% values derived by the A&S method, but the assessment factor WQGs and the HC5 95% were equally precautionary. However, these results should be interpreted cautiously because the analysis was based on only eight chemicals which may not be representative of all chemicals. In comparison, Schudoma (1994) used Monte Carlo simulation and found that in all 1000 cases he examined the HC5 values, irrespective of their confidence limit, were lower than values calculated by the modified USEPA method.

Water quality guidelines for forty chemicals including twenty organics, eight metals, three inorganics and nine pesticides were derived using both the modified USEPA AF method and the A&S method as part of this review (table 7). It should be noted that these values may not be the final WQGs that are derived using either method as other data may be available. The NOEC toxicity data used was obtained from a series of publications by the Dutch National Institute for Public Health and Environmental Protection (Hesse et al 1990, Kalf et al 1995, Van de Plassche & De Bruijn 1992, Van de Plassche et al 1993, Van de Plassche 1994) (Appendix 3). The WQGs were compared using the Wilcoxon signed rank test. This analysis revealed the HC5 50% and HC5 95% values were significantly larger ($p < 0.05$) and smaller ($p < 0.001$) respectively, than the modified USEPA AF values (table 7).

Table 7 Water quality guidelines for forty chemicals derived using the modified USEPA assessment factor and Aldenberg and Slob methods

Chemical type	Chemical name[a]	Water quality guidelines (mg/L)		
		Aldenberg & Slob method		Modified USEPA method (lowest NOEC/10)
		HC5 50%	HC5 95%	
Organic	1,1,1-trichloroethane[c]	5.0×10^{-1}	1.6×10^{-3}	1.3×10^{-1}
	1,1,2-trichloroethane[c]	3.5×10^{0}	2.0×10^{-1}	1.0×10^{0}
	1,2-dichloroethane[c]	8.4×10^{0}	7.2×10^{-1}	1.1×10^{0}
	2,4-dichloroaniline[b,c]	6.4×10^{-2}	8.2×10^{-3}	3.2×10^{-3}
	3-chloropropene[c]	7.2×10^{-1}	3.0×10^{-2}	3.2×10^{-1}
	Benzo(a)anthracene[c]	2.9×10^{-3}	8.5×10^{-4}	4.0×10^{-4}
	Benzo(a)pyrene[c]	2.0×10^{-3}	6.7×10^{-4}	2.7×10^{-4}
	Benzo(ghi)pyrlene[c]	2.7×10^{-4}	1.3×10^{-5}	5.0×10^{-5}
	Chrysene[c]	3.4×10^{-3}	1.2×10^{-3}	4.4×10^{-4}
	Dinitro-orthocresol[b,c]	3.1×10^{-2}	2.8×10^{-3}	3.0×10^{-3}
	Ethylbenzene[c]	2.6×10^{-1}	5.7×10^{-4}	1.0×10^{-1}
	Fluoranthene[c]	6.9×10^{-3}	1.5×10^{-3}	1.6×10^{-3}
	Naphthalene[c]	9.9×10^{-2}	2.1×10^{-2}	2.1×10^{-3}
	ρ-nitrotoluene[b,c]	4.5×10^{-1}	1.0×10^{-1}	3.2×10^{-2}
	Pentachlorophenol[c]	3.2×10^{-3}	5.3×10^{-4}	3.2×10^{-4}
	Pentachlorophenol[b,c]	4.3×10^{-3}	3.3×10^{-4}	1.0×10^{-4}
	Phenanthrene[c]	2.5×10^{-2}	6.7×10^{-3}	3.2×10^{-3}
	TPBS[b,c]	2.7×10^{-1}	4.2×10^{-2}	3.2×10^{-2}
	Trichloroethane[c]	2.1×10^{0}	2.0×10^{-2}	5.8×10^{-1}
	Trichloromethane[c]	5.6×10^{0}	1.3×10^{-1}	9.7×10^{-1}

Table 7 cont.

Chemical type	Chemical name[a]	Water quality guidelines (mg/L)		
		Aldenberg & Slob method		Modified USEPA method (lowest NOEC/10)
		HC5 50%	HC5 95%	
Metal	Berylium[c]	1.6×10^{-4}	2.8×10^{-6}	6.8×10^{-5}
	Cadmium[c]	1.8×10^{-4}	3.5×10^{-5}	8.5×10^{-5}
	Cadmium[d,e]	2.7×10^{-1}	1.6×10^{-2}	7.5×10^{-2}
	Cobalt[c]	2.6×10^{-3}	8.2×10^{-6}	5.0×10^{-4}
	Copper[c]	3.9×10^{-3}	2.2×10^{-3}	3.0×10^{-4}
	Copper[d,e]	6.2×10^{0}	1.8×10^{-1}	1.3×10^{0}
	Mercury[c]	2.0×10^{-5}	9.3×10^{-7}	2.0×10^{-6}
	Selenium[c]	3.5×10^{-3}	5.1×10^{-4}	9.0×10^{-5}
Inorganic	Potassium dichromate[b,c]	7.1×10^{-2}	1.0×10^{-2}	1.0×10^{-2}
	Sodium arsenate[c]	6.0×10^{-3}	2.0×10^{-5}	1.0×10^{-3}
	Sodium bromide[b,c]	4.1×10^{-3}	2.1×10^{-4}	1.0×10^{-3}
Pesticide	α-HCH[c]	2.2×10^{-5}	2.0×10^{-5}	9.0×10^{-4}
	β-HCH[c]	1.7×10^{-2}	8.3×10^{-4}	2.7×10^{-3}
	Carbofuran[d,e]	2.0×10^{-1}	3.5×10^{-3}	1.7×10^{-1}
	Chlorpyrifos[d,e]	5.2×10^{-1}	3.6×10^{-2}	4.6×10^{-2}
	Dieldrin[c]	8.3×10^{-5}	1.1×10^{-6}	1.2×10^{-5}
	Dimethoate[b,c]	1.8×10^{-2}	3.6×10^{-4}	1.0×10^{-3}
	Endrin[c]	3.1×10^{-6}	2.0×10^{-8}	3.0×10^{-6}
	γ-HCH[c]	8.6×10^{-5}	6.1×10^{-5}	2.2×10^{-4}
	Thiram[d,e]	2.5×10^{-1}	1.6×10^{-2}	3.8×10^{-2}

a the raw data used to derive these values are presented in Appendix 3 except for the chemicals with a 'b'

b these values were obtained from OECD (1992a) and the raw data is not readily available

c toxicity data is for aquatic species

d toxicity data is for soil species

e the units are not mg/L but rather mg/kg

It would have been preferable to compare WQGs derived using the A&S method and the AF method used by ANZECC (1992), however, the LOEC data required by the latter method were not readily available. When the modified USEPA derived values were multiplied by three in order to estimate the values that would be derived by the AF method used by ANZECC then the HC5 50% values were still significantly larger and the HC5 95% values were still smaller.

Table 8 presents the WQGs as two different ratios. The first is when the HC5 95% value is less than the modified USEPA value and assumes that the HC5 95% is equal to 1. The second is when the modified USEPA value is less than the HC5 95% and assumes that the modified USEPA value is 1. This revealed that 67.5% and 70% of the HC5 95% values respectively were less than and less than or equal to the modified USEPA values. However, when it was taken into account that LOEC data can be up to 3 times larger than NOEC data then 87.5% of the HC5 95% values would be lower than or equal to the ANZECC AF derived values.

Table 8 The ratio of water quality guidelines derived using the ANZECC assessment factor and Aldenberg and Slob methods

Chemical type	Chemical name[a]	Ratio of toxicity when HC5 95 < ANZECC (HC5 95% CI = 1)		Ratio of toxicity when HC5 95% > ANZECC (ANZECC = 1)	
		A&S method HC5 95% CI	Modified USEPA method (lowest NOEC/10)	A&S method HC5 95% CI	Modified USEPA method (lowest NOEC/10)
Organic	1,1,1-trichloroethane[c]	1	79.2		
	1,1,2-trichloroethane[c]	1	4.8		
	1,2-dichloroethane[c]	1	1.5		
	2,4-dichloroaniline[b,c]			2.6	1
	3-chloropropene[c]	1	10.6		
	Benzo(a)anthracene[c]			2.1	1
	Benzo(a)pyrene[c]			2.5	1
	Benzo(ghi)pyrlene[c]	1	3.7		
	Chrysene[c]	1	3.7		
	Dinitro-orthocresol[b,c]	1	1.1		
	Ethylbenzene[c]	1	176.2		
	Fluoranthene[c]	1	1.07		
	Naphthalene[c]			10.3	1
	ρ-nitrotoluene[b,c]			3.1	1
	Pentachlorophenol[c]			1.7	1
	Pentachlorophenol[b,c]			1.03	1
	Phenanthrene [c]			2.1	1
	TPBS[b,c]			1.3	1
	Trichloroethane[c]	1	28.9		
	Trichloromethane[c]	1	7.7		
Metal	Beryllium[c]	1	24.4		
	Cadmium[c]	1	2.4		
	Cadmium[d,e]	1	4.8		
	Cobalt[c]	1	61.3		
	Copper[c]			7.4	1
	Copper[de]	1	7.1		
	Mercury[c]	1	2.2		
	Selenium[c]			5.6	
Inorganic	Potassium dichromate[b,c]	1	1		
	Sodium arsenate[c]	1	49.6		
	Sodium bromide[b,c]	1	5.0		
Pesticide	∝-HCH[c]	1	45.9		
	β-HCH[c]			30.3	1
	Carbofuran[d,e]	1	48.6		
	Chlorpyrifos[d,e]	1	1.2		
	Dieldrin[c]	1	11.2		

Table 8 cont

Chemical type	Chemical name[a]	Ratio of toxicity when HC5 95 < ANZECC (HC5 95% CI = 1)		Ratio of toxicity when HC5 95% > ANZECC (ANZECC = 1)	
		A&S method HC5 95% CI	Modified USEPA method (lowest NOEC/10)	A&S method HC5 95% CI	Modified USEPA method (lowest NOEC/10)
	Dimethoate[b,c]	1	10.0		
	Endrin[c]	1	148.3		
	γ-HCH[c]	1	3.6		
	Thiram[d,e]	1	2.3		

a - The raw data used to derive these values are presented in Appendix 3 except for the chemicals with a 'b'.
b - These values were obtained from OECD (1992a) and the raw data is not readily available.
c - Toxicity data is for aquatic species.
d - Toxicity data is for soil species.
e - The units are not mg/L but rather mg/kg.

For the organics, metals, inorganics and pesticides the HC5 95% values were lower than the modified USEPA values for 55%, 87.5%, 100% and 89% of the chemicals respectively. Allowing for LOEC values being up to three times the NOEC values indicates that 90% of the organic chemicals would have HC5 95% values lower than the ANZECC values. For the other groups of chemicals no change would occur. Therefore the HC5 95% values derived by the A&S method provide a greater degree of protection to the environment in the majority of cases than is provided by the modified USEPA method and would be provided by the AF method used by ANZECC.

While the above analysis revealed that the HC5 95% values as a group were significantly lower than the modified USEPA AF derived values this does not necessarily mean that they will provide a higher degree of protection. This is because the tails of the logistic and normal distributions are very long and large decreases in concentration may be required in order to increase the degree of biological protection (refer to fig 2). To resolve this issue the values presented in table 7 were substituted into equations 5 and 6 (section 4.2.1) in order to calculate the percentage of species that should theoretically be protected (table 9). This could only be done for those chemicals for which the original toxicity data the WQGs were based on were available (ie 32 chemicals).

Both methods provided a high level of protection, with the lowest level being greater than 89% (table 9). A Wilcoxon signed ranks test of the data indicated that the HC5 95% should protect a statistically significant ($p < 0.05$) greater percentage of species than the USEPA AF derived values. Again, based on the fact that LOEC data is larger than NOEC data, it can be inferred that the HC5 95% values would provide a statistically greater degree of protection than ANZECC WQGs derived using LOEC data. However, when the HC5 95% values were examined individually they did not always offer a higher degree of protection than the modified USEPA AF derived values.

The values derived above (table 7) and an estimate of the equivalent ANZECC values, which assumed that the LOEC values were three times the NOEC values, were compared with the current ANZECC WQGs. However, it should be noted that the values derived above are not necessarily the values that will be derived if these methods are used as other toxicity may be available and this could modify the resulting WQG.

Table 9 The percentage of species that should theoretically be protected by the WQGs presented in table 7. These values obtained by entering the WQGs into the Aldenberg and Slob method and back calculating to the level of protection.

Chemical type	Chemical name[a]	% of species that should be protected	
		HC5 95	Modified USEPA
Organics	1,1,1-trichloroethane[c]	99.770	93.770
	1,1,2-trichloroethane[c]	98.976	90.937
	1,2-dichloroethane[c]	97.239	94.980
	3-chloropropene[c]	99.768	94.013
	Benzo(a)anthracene[c]	99.999	99.999
	Benzo(a)pyrene[c]	99.999	99.999
	Benzo(ghi)pyrlene[c]	99.999	99.998
	Chrysene[c]	99.999	99.999
	Ethylbenzene[c]	99.825	93.136
	Fluoranthene[c]	99.993	99.992
	Naphthalene[c]	99.869	99.996
	Pentachlorophenol[c]	99.999	99.999
	Chenanthrene [c]	99.982	99.995
	Trichloroethane[c]	99.546	89.976
	Trichloromethane[c]	99.063	90.663
Metals	Beryllium[c]	99.996	99.917
	Cadmium[c]	99.999	99.999
	Cadmium[d,e]	99.815	98.659
	Cobalt[c]	99.977	99.466
	Copper[c]	100.00	100.00
	Copper[de]	98.792	89.368
	Mercury[c]	99.999	99.999
	Selenium[c]	99.987	99.999
Inorganics	Sodium arsenate[c]	99.995	99.911
Pesticides	∝-HCH[c]	99.999	99.962
	β-HCH[c]	99.999	99.999
	Carbofuran[d,e]	99.887	92.706
	Chlorpyrifos[d,e]	99.666	99.557
	Dieldrin[c]	99.999	99.998
	Endrin[c]	99.999	99.988
	γ-HCH[c]	99.999	99.999
	Thiram[d,e]	99.999	99.096
Statistics	Mean	99.754	97.690
	95% CL	99.554–99.955	96.37–99.01

a the raw data used to derive these values are presented in Appendix 3 except for the chemicals with a 'b'

b these values were obtained from OECD (1992a) and the raw data is not readily available

c toxicity data is for aquatic species

d toxicity data is for soil species

e the units are not mg/L but rather mg/kg

This comparison was possible for only 11 chemicals (table 10) and revealed that for chromium there was no difference between any of the values except the estimated ANZECC value was higher. However, for all the other chemicals except pentachlorophenol and dieldrin, both the HC5 95% and the modified USEPA AF derived values were smaller than the corresponding current ANZECC values. When the current ANZECC WQG for pentachlorophenol was examined in detail it was not possible to obtain the same value using the data from which it was supposedly derived. It is most probable that a typographical error occurred in generating the ANZECC WQG and the correct value should be 0.5 µg/L which is the CCREM (1991) value. When a value of 0.5 µg/L is used, then once again the HC5 95% and AF method values are both lower than the current ANZECC value. The estimated ANZECC values were in most cases less than the current ANZECC values, the exceptions being chromium, pentachlorophenol, ethylbenzene and dieldrin. However, it is difficult to draw conclusions when comparing the current ANZECC WQGs with the estimated ANZECC values, as they are only estimates. Nevertheless, these findings are significant because they indicate that irrespective of the method (either AF or Aldenberg and Slob) used to derive the new ANZECC WQGs, there are going to be reasonably large decreases for some chemicals and increases for other chemicals.

Table 10 A comparison of current ANZECC water quality guidelines with those developed in this study using the modified USEPA method, the Aldenberg and Slob method and an estimate of the ANZECC values based on equivalent LOEC data

Chemical name	Water Quality Guidelines (µg/L)			
	Current ANZECC WQGs[1]	HC5 95%[2]	Modified USEPA AF[3]	Estimated ANZECC (3 x mod USEPA)
Arsenic	50	0.0216	1	3
Beryllium	4	0.0027	0.068	0.204
Cadmium	0.2–2	0.035	0.085	0.255
Chromium	10	10	10	30
Copper	2–5	2.22	0.3	0.9
Mercury	0.1	0.0009	0.002	0.006
Selenium	5	0.507	0.09	0.27
Pentachlorophenol	0.05[4]	0.43	0.21	0.63
Ethylbenzene	140	0.567	100	300
Dieldrin	0.002	0.00107	0.012	0.036
Endrin	3	0.00002	0.003	0.009

1 The assessment factor method currently used by ANZECC (1992)

2 The Aldenberg and Slob (1993) method

3 The modified USEPA assessment factor method (OECD 1992a)

4 This value could not be verified using the information from which it was said to be derived. The most likely value is 0.5 ug/L (CCREM 1991).

7 Summary of the comparison of the assessment factor, Aldenberg and Slob, and multiple species methods

The scientific assessment of the assessment factor (AF), Aldenberg and Slob (A&S) and multiple species (MS) methods indicated that all suffered from a number of limitations. The MS method provides toxicity data that are more environmentally realistic due to the biological complexity of the test system. The lack of replication and treatments may compromise the scientific rigour of some MS tests. However, MS tests that have sufficient

replication and treatments offer the best means of deriving WQGs. There is no theoretical basis for AFs (ie, they are empirical) and the magnitude of AFs has been shown to significantly underestimate the degree of variability in extrapolating between species and between acute to chronic toxic effects. A number of assumptions made by the A&S method may be compromised and are the subject of ongoing scientific debate. As the methods were so different and all had inadequacies it was difficult to determine which method was more valid than the other.

Despite this the A&S method does offer several advantages over the AF method:

- it conforms to risk assessment principles;
- it permits the choice of both the level of protection and the level of uncertainty in the chosen level of protection;
- the concept of the degree of protection is easily understood and should enable more informed debate;
- the method can be used in the reverse manner for ecological risk assessment.

Comparison of the methods in terms of the Precautionary Principle indicated that while both methods failed to take into account the toxicity of mixtures, bioaccumulation and inter-compartmental transfer, the A&S method contained more precautionary elements. However, concluding one method was more precautionary than the other was not possible because of the inability to quantify the precautionary nature of each element. The comparison of the WQGs derived using both methods for forty compounds indicated that the HC5 95% values were significantly lower and provided a significantly greater degree of protection than those derived by the modified USEPA AF method and, by implication, the ANZECC AF method. These results indicate that the HC5 95% values of the A&S method were more precautionary than the ANZECC AF derived values for between 70% and 87.5% of the chemicals examined. However, the HC5 95% values were not consistently lower than those derived using the modified USEPA assessment factor method nor did they consistently provide a greater degree of protection. It is therefore not possible to state with any degree of confidence whether either the A&S or AF method is better than the other.

8 Methods to increase the precautionary nature and relevance of water quality guidelines to ANZECC

As mentioned previously, the current methods for determining WQGs are not fully precautionary as they do not incorporate the toxicity of mixtures, bioaccumulation and inter-compartmental transfer, all of which are known to occur. The relevance of the WQGs to Australia and New Zealand could be increased by considering such variables as whether or not overseas toxicity data should be used and factors that are known to affect the toxicity of chemicals in local waters. How each of these factors could be incorporated into the derivation of WQGs is discussed below.

8.1 The use of Australian and New Zealand toxicity data

Aquatic toxicity data for Australian and New Zealand species and conditions is relatively limited. Thus the derivation of most water quality guidelines for Australia and New Zealand will have to rely predominantly on overseas data. It is therefore important to know whether overseas toxicity test protocols and toxicity data are relevant to Australia and New Zealand. The relative sensitivity of Australian and overseas aquatic species to toxicants has been examined but the results have

been contradictory. Skidmore and Firth (1983) felt that for aquatic biota as a whole, there were no real differences. Similarly, Markich and Camilleri (1997) found there was no difference in the sensitivity of tropical Australian and temperate USA fish species to copper. Johnston et al (1990) used OECD test protocols and compared the toxicity of Australian and overseas fish, cladocera and algae and concluded that there were only differences between the cladoceran species. When differences did occur the Australian cladoceran, *Ceriodaphnia* cf. *dubia*, was less sensitive than the overseas species. Rose et al (1998) found differences in the sensitivity of the Australian *Ceriodaphnia* cf. *dubia* and the Northern Hemisphere *Daphnia magna* to a range of narcotic chemicals. However, in contrast to the findings of Johnston et al (1990), Rose et al found the Australian species was approximately three times more sensitive than the overseas species. Sunderam et al (1992, 1994) measured lethal effects of fish to endosulfan and found that some overseas species were more sensitive than the tested Australian species. In contrast, Davies et al (1994) compared the sub-lethal sensitivity of fish to a range of pesticides and found the Australian species were more sensitive. Mulhall (1997) compared the toxicity of a range of polar narcotics to *Ceriodaphnia* cf *dubia* and *Daphnia magna* and found there was no overall pattern—for some chemicals *C.* cf. *dubia* was more sensitive, for some it was less sensitive or as sensitive. Due to the lack of any comprehensive comparison the relative sensitivity of Australian and overseas aquatic species remains unclear. To the authors knowledge no similar comparisons have been made between the sensitivity of New Zealand and overseas species. It must therefore, be assumed for the current review of WQGs that there is no difference in sensitivity.

The NSW EPA is currently establishing an Australasian ecotoxicology database for both aquatic and terrestrial data. The database currently holds approximately 2500 entries with data for approximately 170 chemicals and effluents and 250 organisms (Warne pers comm). The database has two immediate uses. Firstly, it will act as a source for Australasian toxicity data for deriving WQGs. Secondly, it will enable a much more comprehensive comparison of the relative sensitivity of Australasian and Northern Hemisphere organisms.

Due to the lack of toxicity data to Australian and New Zealand species it would only be possible to derive WQGs for a very limited number of chemicals if they were to be based solely on this data. Therefore, it is recommended that Australian, New Zealand and overseas toxicity data be used to derive WQGs. In recommending this it is assumed, but this remains to be verified, that there is no difference in the sensitivity of local and overseas species to toxicants.

8.2 Factors that modify toxicity

It is well established that a number of physical and chemical factors can modify the toxicity of chemicals (Mance 1987, Rand et al 1995, Landis & Yu 1995). The physical factors that can affect toxicity include temperature and light intensity. Chemical factors that are known to affect toxicity include pH, Eh, salinity, water hardness, total and dissolved organic matter, suspended particulate matter and dissolved oxygen.

Which particular factor(s) will modify the toxicity of a particular chemical depends on the physicochemical properties of the chemical. For instance, the pH and Eh of the surrounding water generally affects the toxicity of metals. These two parameters largely control the ionic state of metals which inturn modifies lipid and water solubility and binding properties. For non-volatile, non-ionic organic chemicals these factors have little to no effect rather, they are affected by the presence of organic matter, suspended matter and sediment in the water. For volatile organic chemicals temperature is probably the most important characteristic affecting aquatic toxicity.

Research conducted both locally and overseas has investigated the affects of the above modifying factors (eg Patra et al 1995, 1996, Willis 1994, Mayer & Ellersieck 1986). Despite this there has only been extensive work and understanding developed on the affect of water hardness on the toxicity of metals. Attempting to account for all the modifying factors would be an enormous task and is beyond the capabilities of the current review given the time restraints. It is therefore proposed that only hardness be considered in deriving WQGs.

8.3 The implications of practical quantitation limits and background levels to deriving WQGs

Practical quantitation limits (PQLs) are the concentrations that should be detectable by laboratories conducting routine analyses (Plues pers comm). These can be quite different from detection limits, which vary for each instrument and may bear little relevance to normal concentrations of samples that can be detected in laboratories. Some existing water quality guidelines are less than the PQLs and below the background levels of chemicals. The occurrence of such WQGs has been criticised (McAlpine pers comm). The following sections will examine such occurrences and discuss the implications to deriving and implementing WQGs.

8.3.1 Practical quantitation limits

It is possible that WQGs may be below the practical quantitation limits (PQLs) for a particular chemical. The Australian Department of Land and Water Resources has criticised WQGs below the PQLs (McAlpine pers comm). However, the occurrence of such WQGs is really not a problem.

The implementation of WQGs below the detection limit is no different from that for any other chemical. If the aqueous concentration of the chemical is shown to exceed the WQG (ie it is measurable) then further investigation is required. One problem with WQGs that are below the PQLs is that they allow the WQG to be exceeded by a larger amount than would otherwise be permitted before further investigation is commenced. This arises because it only becomes apparent that a chemical exceeds the WQG when it exceeds or equals the PQL. Whereas for chemicals with WQGs above the PQL, the occurrence of an aqueous concentration exceeding the WQG warrants investigation. This problem is potentially detrimental to the environment because adequate environmental protection may not be provided.

The alternative to this is to derive WQGs that exceed or equal the PQLs. There are a number of problems with such an approach. Firstly, PQLs change over time therefore the WQG would constantly be changing and it would be difficult to ascertain what the WQG value is. Secondly, the WQGs are based on the best available toxicological knowledge and to increase the WQG and therefore decrease the degree of protection provided simply because the chemical can not be detected is scientifically indefensible. Thirdly, if this approach were adopted then the environment would be provided with different levels of protection for different chemicals. Finally such a proposal would also encourage the use of less sensitive analytical techniques. While the existing situation is not ideal—as it potentially allows aqueous concentrations to significantly exceed the WQGs before detection—it is still preferable to the alternative.

8.3.2 Background levels

Background levels are the aqueous concentrations of chemicals that are of natural origin and are not associated with any direct or indirect interference by humans. They are normally associated with metals, however, they can also exist for organic chemicals. Background levels

for chemicals vary considerably from one locality to another. For example, the background levels of metals in Australia vary over one to three orders of magnitude (ANZECC & NHMRC 1992).

It is often argued that the variation in background levels makes establishing national WQGs impossible or at best very difficult. The reason for this is that there are aquatic organisms in virtually all water bodies irrespective of the background levels. For instance, animals live in water bodies with high background levels of metals that may be considerably above the WQGs. Critics of national WQGs therefore argue that there are many instances or localities for which the national WQGs are not valid and that only site-specific investigations can provide adequate WQGs. However, while there is definitely some merit in these criticisms there are two fundamental problems with such a proposition. Firstly, the role of the ANZECC WQGs is to act as a screening process. If the WQGs are exceeded then further investigation is required which may eventually lead to site-specific investigations. Secondly, to have a system based only on site-specific investigations and site-specific WQGs is impractical and unnecessarily expensive.

As the purpose of the ANZECC WQGs is to act as a screening process there is no problem posed by WQGs that are below the background levels in certain localities. If this occurs it can be dealt with by considering the background level when the water manager implements the WQGs. Alternatively, it can be considered when follow up investigations are conducted.

8.4 Toxicity of mixtures

All current methods for deriving WQGs determine supposedly 'safe concentrations' for individual compounds, however, they do not take into account that other compounds may be present and also exerting toxic effects (Chapman pers comm, Vighi & Calamari 1996).

It is widely known that organisms living in the environment are not exposed to individual compounds but to complex mixtures that may contain hundreds or thousands of compounds (eg Könemann 1980, Hermens et al 1984, 1985, Broderius & Kahl 1985, Vighi & Calamari 1996, Van Leeuwen et al 1996, Grimme et al 1996, Ankley & Mount 1996). Mixtures of organic chemicals with non-specific mechanisms of action generally have additive toxicity (Hermens et al 1984, 1985) (see Glossary). However, this is not always the case with synergistic and antagonistic mixtures (see Glossary) being reported (Broderius & Kahl 1985, De Zwart & Slooff 1987, Warne et al 1989b). It also appears that the number of components in a mixture affects the toxicity of the mixture (McCarty & Mackay 1993, Warne & Hawker 1995). Research has also shown that concentrations as low as two per cent of the LC50 (Deneer et al 1988) and below the NOEC (Kraak et al 1994) contribute to the toxicity of mixtures.

A recent review (Ross 1996) examined the toxicity of 1048 mixtures which included metals, organics, pesticides and inorganics and combinations of these. The review found that the 95% confidence limit for the mean of all the mixtures overlapped additivity and hence it was unlikely that the mean toxicity of the mixtures was significantly different to additivity. It also found that in the vast majority of mixtures examined the duration of exposure (ie chronic or acute), the type of biological endpoint (ie lethal or sub-lethal), the type of chemicals in the mixture and type of organism tested did not affect the toxicity of the mixture (ie they were essentially all additive). The two exceptions to this were mixtures of metals that were antagonistic and mixtures of inorganics that were synergistic.

However, in attempting to incorporate the toxicity of mixtures into water quality guidelines it is not sufficient just to examine the mean toxicity of mixtures. Rather a decision needs to be made concerning what percentage of toxicant mixtures the organisms in the environment

should be protected from. This issue is similar to the issue of what percentage of species should be protected by WQGs.

Closer examination of the data in Ross revealed that despite the mean being essentially additive there were numerous mixtures, which were either synergistic or antagonistic. The magnitude of the deviation of a mixture from toxic additivity can be expressed as a percentage. For example, values of -3, 0 and 50% mean that the toxicity of the mixtures are 3% less than, equal to, and 50% greater than toxic additivity respectively. Table 11 shows the percentage deviation from toxic additivity for various cumulative percentages of all the 1048 mixtures.

The data presented in table 11 could be useful in deriving WQGs as they indicate how toxic mixtures are in terms of the percentage of mixtures the environment is to be protected from. [This data indicated that in order to protect the environment from 50% of mixtures it should be assumed that the toxicity of mixtures is additive. To protect the environment from 90% of the mixtures the toxicity should be assumed to be 114 % more toxic than toxic additivity. While to protect the environment from 100% of the mixtures, the toxicity should be assumed to be 3233% more than toxic additivity. Thus any WQG that aims to protect the environment from more than 52% of the mixtures should assume that the toxicity is greater than additive. However, this observation was predominantly based on toxicity data for mixtures with two or three components.

Table 11 The percentage of the 1048 mixtures examined by Ross (1996) that deviate from toxic additivity by a certain amount

Percentage of mixtures	Percentage deviation from toxic additivity[1]
50	-3
52	0
75	40
80	63
85	92
90	114
95	142
99	400
99.5	614
100	3233

[1] Positive value indicates that the mixtures are more toxic than toxic additivity while a negative value means the mixtures are less toxic than additivity.

The work by Ross (1996) also supported the funnel hypothesis of Warne and Hawker (1995) that predicted that mixtures with a small number of components may deviate markedly from additivity but that as the number of components increases the mixtures deviate less from additivity and become essentially additive. The review by Ross indicated that this appeared to be true irrespective of the type of chemicals in the mixtures.

The combination of these two observations leads to a series of estimates of how toxic mixtures should be assumed to be, based on the number of components in the mixture (table 12). However, as most mixtures that the environment is likely to be exposed to will contain numerous components (ie more than 20), it is probably sufficient to assume toxic additivity.

Table 12 The number of components in mixtures of toxicants and the mean percentage deviation from toxic additivity that the mixtures must be assumed to have in order to protect the environment from all mixtures with that number of components

Number of components	Mean % deviation from toxic additivity[1]
2	3233
3	317
4	128
5–9	234
10–15	121
16–20	65
21–50	8

[1] A positive value indicates that the mixtures are more toxic than toxic additivity while a negative value means the mixtures are less toxic than additivity.

Other workers have also concluded that mixtures of chemicals were largely additive (Pedersen et al 1994, Ankley & Mount 1996, Vighi & Calamari 1996) and that the concept of additivity needed to be included in the derivation of water quality guidelines (Calamari & Vighi 1992, Bro-Rasmussen et al 1994, Van Leeuwen et al 1996, Vighi & Calamari 1996, Grimme et al 1996). However, these studies were not as extensive as that of Ross.

Only the Dutch have any procedure to account for the toxicity of mixtures. They have two levels of WQGs, the maximum permissible concentration (MPC) (which is equivalent to ANZECC WQGs) and the target level (TL). The MPC values do not account for mixtures, only the target values do this. The toxicity of mixtures is accounted for by dividing the MPCs of individual chemicals by 100 (VROM 1994). The rationale for this is that there are at least 100 compounds in the Rhine River at any one time (VROM 1994). Thus they assume all compounds are present with 99 others and the joint toxicity is additive. However, it should be noted that the TLs represent the long-term goals for water quality in The Netherlands. They do not apply currently, so technically the Dutch also do not have WQGs that account for the toxicity of mixtures.

Toxic addition occurs in two forms: response addition and concentration addition (Plackett & Hewlett 1952) (see Appendix 5 for details of the differences). Of these only the concentration addition method is recommended for regulatory purposes. The response addition model is not recommended because it can not be used with WQGs derived using an AF method and requires the dose response relationship for each component of a mixture which is generally not reported in the literature.

Neither the AF method nor the A&S method can incorporate the toxicity of mixtures into the derivation of WQGs. However, the values derived by these methods could be modified to account for mixture toxicity using the following formula, which is based on the equation derived by Vighi and Calamari (1996)

$$TTM = \Sigma (C_i \div WQG_i) \quad (9)$$

where TTM is the total toxicity of the mixture, C_i is the concentration of 'i'th component in the mixture and WQG_i is the water quality guideline for the 'i'th compound. If TTM exceeds one, the mixture has exceeded the WQGs. However, if TTM is less than or equal to one the WQGs have not been exceeded. Further, if the aqueous concentration of any chemical in the mixture exceeds or equals its WQG then the water quality guidelines are automatically exceeded.

A problem with this approach is that it is not possible to identify all the chemicals in a water body and in such cases an estimate of the toxicity of the mixture using the above method would underestimate the actual toxicity. In fact it is highly unlikely that it will ever be feasible to identify all the chemicals in a water body. However, it is possible to identify a good proportion of those most likely to be exerting toxic effects. This could be done by determining which groups of chemicals may be released from sources and then analytically identifying and quantifying the chemicals belonging to these groups. Even though all chemicals will not be identified by such an approach it is probably sufficient to identify the major toxicants and consider the toxicity of mixtures of these chemicals.

Another means of overcoming this problem that could be useful in site specific investigations is whole effluent toxicity testing methods (WETT). These methods are used extensively in the USA (USEPA 1984a), England, Ireland, Sweden and Denmark (Pedersen et al 1994). In this approach, test organisms are directly exposed to the effluent after the mixing zone and in this way the measured toxicity is the combined toxic effect of all the components in the effluent. Using WETT methods the components of the effluent do not have to be identified in order to obtain an estimate of the toxicity (Pedersen et al 1994). Whole effluent toxicity testing techniques are not suitable for the generation of national water quality guidelines because they measure the toxicity of individual effluents that would not necessarily apply to effluents from similar industrial facilities let alone different industries. However, they do have a potentially significant role in the assessment of individual sites and in the derivation of site-specific effluent licenses.

Neither the assessment factor method used by ANZECC (1992) nor the Aldenberg and Slob method are precautionary or even preventive as they do not account for the toxicity of mixtures. It would be consistent with the Precautionary Principle to assume that the presence of all chemicals, at any concentration, can exert a toxic effect and that the toxicity of all mixtures of compounds is at least concentration additive.

8.5 Bioaccumulation of chemicals

The AF and statistical extrapolation methods only account for the direct effects of toxicants. They do not take into account indirect effects—such as the toxicological effects of chemicals obtained from food. Such indirect effects are generally termed secondary poisoning and are caused by three mechanisms: bioconcentration; bioaccumulation; and biomagnification (see Glossary). Bioconcentration and bioaccumulation are universally accepted to occur. However, there is still scientific debate over whether or not biomagnification in aquatic systems is a universal phenomenon (Thomann 1981, Thomann & Connolly 1984, Connell 1989, 1990, Gobas et al 1993, Galassi et al 1994). All three mechanisms can cause chemicals to concentrate in animal tissue at levels greater than in the surrounding water. Chemicals with the octanol-water partition coefficient (Kow) values or bioconcentration factors (BCF) greater than 1000 are assumed to have a significant potential for accumulating in animal tissue (Connell 1990, Worksafe Australia & NICNAS 1991, OECD 1995).

To maximise the precautionary nature of water quality criteria it would be best to take secondary poisoning into consideration. A possible way to do this is the method developed by Romijn et al (1993, 1994) and subsequently modified by Traas et al (1996) and Jongbloed et al (1996) to suit the terrestrial environment, or methods based on food web analysis (Thomann 1981, Nichols et al 1995). A major problem with the food web based methods is that they are very complex and require extensive data sets that are not available for the vast majority of chemicals.

Both these methods have been used (eg Van de Plassche 1994, USEPA 1994) to derive WQGs but there is currently no internationally recognised method for accounting for secondary poisoning (Bro-Rasmussen et al 1994). To facilitate international acceptance and use of such methods the Romijn et al method was presented in the OECD guidance document for aquatic effects assessment (OECD 1995).

Romijn's model determines the maximum acceptable risk level (MAR) which is essentially a WQG modified so that it protects fish from direct toxic effects and fish-eating organisms from secondary poisoning. This is achieved using the formula:

$$MAR = NOEC_{fish\text{-}eater} / BCF_{fish} \quad (10)$$

where the units for MAR are in mg/L, $NOEC_{fish\text{-}eater}$ is the no observed effect concentration for fish-eating species with units of mg/kg and BCF_{fish} is the bioconcentration factor for fish with units of kg/L.

This model assumes bioconcentration into fish and then biomagnification by species that eat fish. It is a highly simplified representation of the foodwebs that occur in the environment where biomagnification may happen three or four times from the lowest life forms to the highest predators. If a greater degree of protection is desired then it may be possible to modify equation 10 to include further biomagnification events.

The current ANZECC water quality guidelines only consider secondary poisoning for DDT, PCBs and mercury. These WQGs were adopted directly from the USEPA (1976) that were calculated using the formula:

$$WQG = LOEC_{fish\text{-}eater} / \text{geometric mean of } BCF_{fish} \times LC_{fish} \quad (11)$$

where BCF_{fish} is the bioconcentration factor for fish and LC_{fish} is the lipid content of the fish.

However, DDT, mercury and PCBs are by no means the only chemicals known to cause secondary poisoning let alone all those with the potential for secondary poisoning. Thus, while the WQGs may be precautionary for these three chemicals they are not preventive for the remainder of chemicals that are known to cause secondary poisoning and are not precautionary for those that may cause secondary poisoning. Secondary poisoning should be considered in deriving WQG using either the Romijn et al or the USEPA methods for all chemicals with Kowand BCF values greater than 1000.

8.6 Inter-compartmental transport of chemicals

A limitation of much environmental management is that it has been principally concerned with regulations specific to individual environmental compartments (eg air, soil, sediment, water and biota). This approach has been widely criticised as it tends to be uncoordinated and environmental problems are often simply shifted from one compartment of the environment to another (eg World Commission on the Environment and Development 1987, Reilly 1990, Lipman 1993, Escande 1994). This phenomenon has been termed problem shifting or problem displacement (Simonis 1993).

Another limitation with such a compartment specific approach is it fails to recognise that pollutants are transported between the various compartments of the environment. With compartment specific legislation it is possible that even though the environment quality criterion (EQC, the equivalent of WQGs but for each compartment of the environment) for one compartment is not exceeded the EQC in another compartment is exceeded due to inter-compartmental transfer. Compartment specific legislation is inherently liable to such occurrences and for this reason is currently being replaced by more holistic approaches

(Lipman 1993). A local example of this is the policy of the current NSW Government to consolidate the *Clean Air Act*, *Clean Waters Act* and *Noise Control Act* amongst others into one bill, the *Protection of the Environment (Operations) Bill* (Allan 1995).

One method for modifying compartment specific EQCs to a more holistic approach is called harmonisation. This methodology was developed and is used by the Dutch (Van de Meent & De Bruijn 1995) as part of their risk assessment system. However, they have also proposed using it to derive EQCs. This process modifies EQC values so that they account of inter-compartmental transport of contaminants. Harmonisation applies an environmental fate model to compartment specific EQC values to ensure all compartments of the environment are provided with at least the policy level of protection (ie for the Dutch it is protecting 95% of species with 50% confidence, HC5 50%). This inherently implies that in one or two compartments the level of protection may be greater than the policy level.

Harmonisation is achieved by assuming the chemical is released into one compartment at the EQC concentration, then uses the fate model to calculate the concentration in the other compartments. The calculated concentrations in each compartment are then compared with the corresponding EQC values. If the calculated concentrations are not less than the EQCs the EQC for the initial compartment is reduced. The fate modelling is then repeated until all the calculated concentrations are below the corresponding EQC values for the other compartments. The chemical is then added sequentially to the remaining compartments at the appropriate EQC and the process repeated until the EQCs for all compartments are modified so they offer at least the policy level of protection to all other compartments. The Danish EPA (Fredenslund et al 1995) evaluated the computerised version of the harmonisation process called SimpleBox and concluded that it was a useful tool for estimating predicted environmental concentrations (PECs) on a regional scale.

Currently, WQG values for Australia and New Zealand are derived without any consideration to inter-compartmental transfer. Thus, the current method for deriving EQC values is neither precautionary nor even preventive. Inter-compartmental transfer does occur, however whether it occurs for all chemicals is unknown. Therefore, in order for EQCs to be precautionary inter-compartmental transfer should be considered for all chemicals.

8.7 Recommendations

It is recommended that:

1. Only toxicity data with lethality, immobilisation, growth or reproduction endpoints be used;
2. Australian, New Zealand and overseas toxicity data be used;
3. The effect of hardness on the toxicity of metals be accounted for;
4. Practical quantitation limits and background levels not be considered in deriving WQGs;
5. The toxicity of mixtures be accounted for;
6. Bioaccumulation be taken into account for chemicals where the Kowand/or BCF are greater than 1000.

9 A possible future framework for deriving water quality guidelines

The above analysis of the multiple species (MS), Aldenberg and Slob (A&S) and assessment factor (AF) methods revealed that there were limitations to all methods. Despite this MS

toxicity data provides the best means of deriving WQGs. The comparison of WQGs derived using the A&S and AF methods revealed that neither method consistently yielded lower WQGs than the other. Based on these findings a pragmatic and precautionary framework for deriving WQGs is proposed.

9.1 Overview of the proposed framework

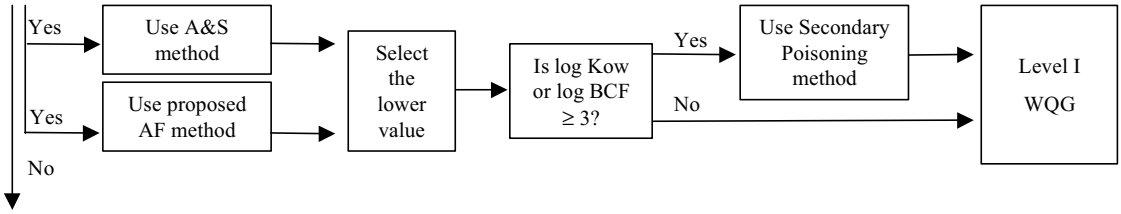
The proposed framework is hierarchical with three different levels of WQG being derived (fig 4). The different levels of WQGs reflect the quantity, type and representativeness of the available toxicity data and the degree of confidence that this data will derive WQGs able to provide adequate environmental protection. Level I and Level II WQGs are both derived using toxicity data in which there is sufficient confidence that the resulting WQG will provide adequate protection, with Level I offering a greater degree of confidence. The 'interim' WQGs are the third level and are derived using data in which there is insufficient confidence that the resulting WQG will provide adequate protection. WQGs are derived for the highest possible level for which there is adequate suitable toxicity data. If there is insufficient suitable data for a level I WQG then one continues down the hierarchy (fig 4) until the available data meets the minimum data requirements for a level of WQG.

The proposed framework uses both an AF and a statistical extrapolation method providing the data meets the minimum data requirements for both methods. The AF method recommended for use is a modified version of the CCREM AF method (1991) which will henceforth be referred to as the 'proposed AF method'. Details of this method will be provided in section 9.2. The statistical extrapolation method recommended for use is the A&S method. The proposed framework is slightly different for Level I WQGs compared with Level II and Interim WQGs. If, once the data have been screened, no data are omitted in calculating Level I WQGs, the A&S method should be used in preference to the 'proposed AF method'. However, if the data are not logistically distributed or do not meet the minimum data requirements of the A&S method then the 'proposed AF method' should be used. When data are omitted in calculating Level I WQGs then both the A&S and 'proposed AF method' should be used. The results of these two methods should then be compared with the lower value becoming the WQG for that chemical providing the Kow and/or BCF of the chemical are less than 1000. The framework recommended for Level II and Interim WQGs is essentially the same as the second method for deriving Level I WQGs. For chemicals that exert their toxicity by the narcotic mode of action, it is proposed that interim WQGs should be derived using the nineteen QSARs used by the Dutch (Appendix 1). The resulting estimates of chronic toxicity should then be used by both the A&S method and the 'proposed AF method' with the lower of the two estimates becoming the interim water quality guideline (fig 4).

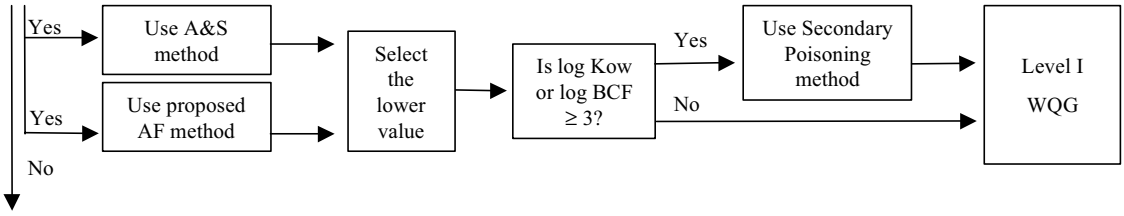
If the chemical has a Kow or BCF value greater than 1000 then it is recommended that one of the methods that accounts for secondary poisoning should be used (section 8.5). Which method should be used depends on the availability of toxicity and physicochemical data. The resulting value would then become the WQG. It is also recommended that the effect of water hardness on the toxicity of several metals (ie copper, zinc, cadmium, lead, chromium III and nickel) should be taken into account in deriving the WQGs for these chemicals. However, due to the limited number of chemicals this is necessary for, figure 4 does not include this modification.

Level I WQGs

1. Are there toxicity data from ≥ 3 multiple species toxicity tests?

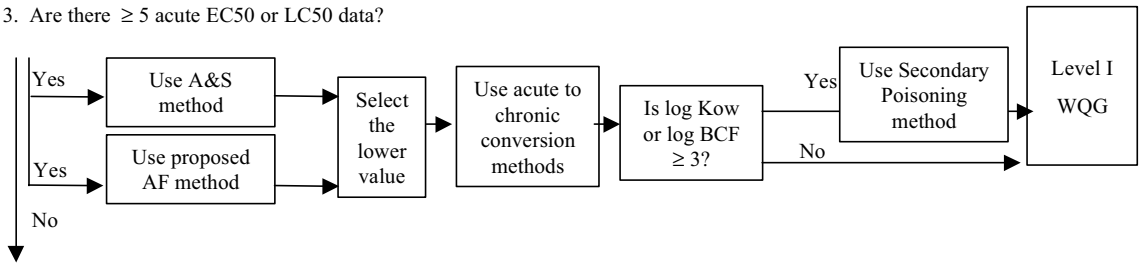


2. Are there ≥ 5 chronic NOEC data?



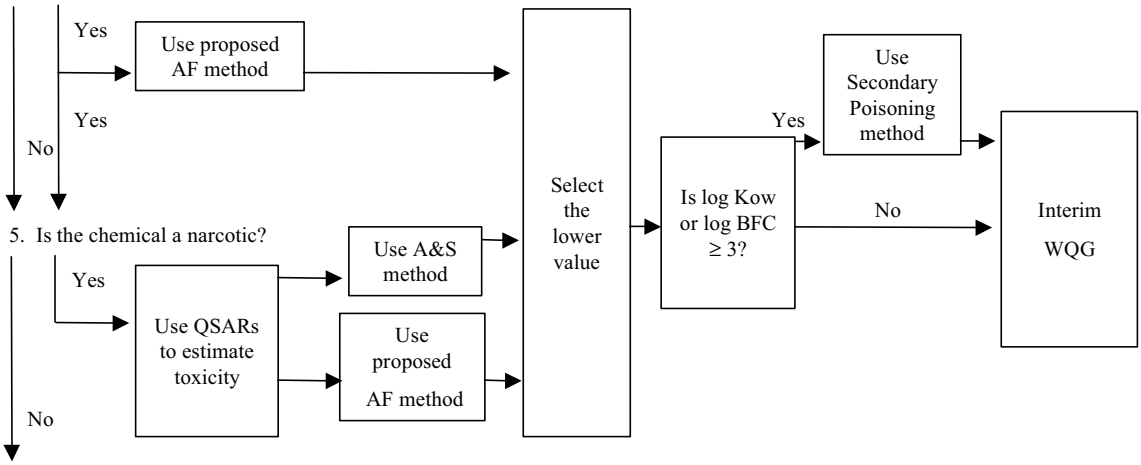
Level II WQGs

3. Are there ≥ 5 acute EC50 or LC50 data?



Interim WQGs

4. Are there ≥ 3 toxicity data*?



6. Can not derive a WQG, but can derive an environmental concern level (ECL).

*This can be all acute or all chronic data or a combination of both.

Figure 4 A schematic diagram of the proposed framework for deriving the three different levels of water quality guidelines

It is also proposed that the WQGs account for the toxicity of mixtures. This is, however, not possible to do when deriving the WQGs. Rather, it is recommended that it is done when the WQGs are implemented into water management decisions.

The proposed method is different in several ways to frameworks used overseas. Firstly, it does not give a clear preference to either an AF or statistical extrapolation method, except for Level I WQGs. In contrast the USEPA (1986), OECD (1992a) (see fig 5), Danish (Petersen & Pedersen 1995) and The Netherlands (Van de Plassche et al 1993)(see fig 6) methods all prefer statistical extrapolation techniques to the AF methods. However, the proposed approach is entirely consistent with the findings of this review. Secondly, it is proposed that the toxicity of mixtures should be taken into account which is not done in any of the reviewed frameworks. Thirdly, it is proposed that the guidelines should account for the potential for secondary poisoning. Only the USEPA and The Netherlands currently consider this in deriving WQGs.

It is felt that the proposed method will deliver the most scientifically rigorous and defensible WQGs given the current state of knowledge. However, as recognised by the OECD (1994) and Rensvik (1994) the final decision on the actual level of a WQG is a socio-political decision, which takes into account many factors other than just science.

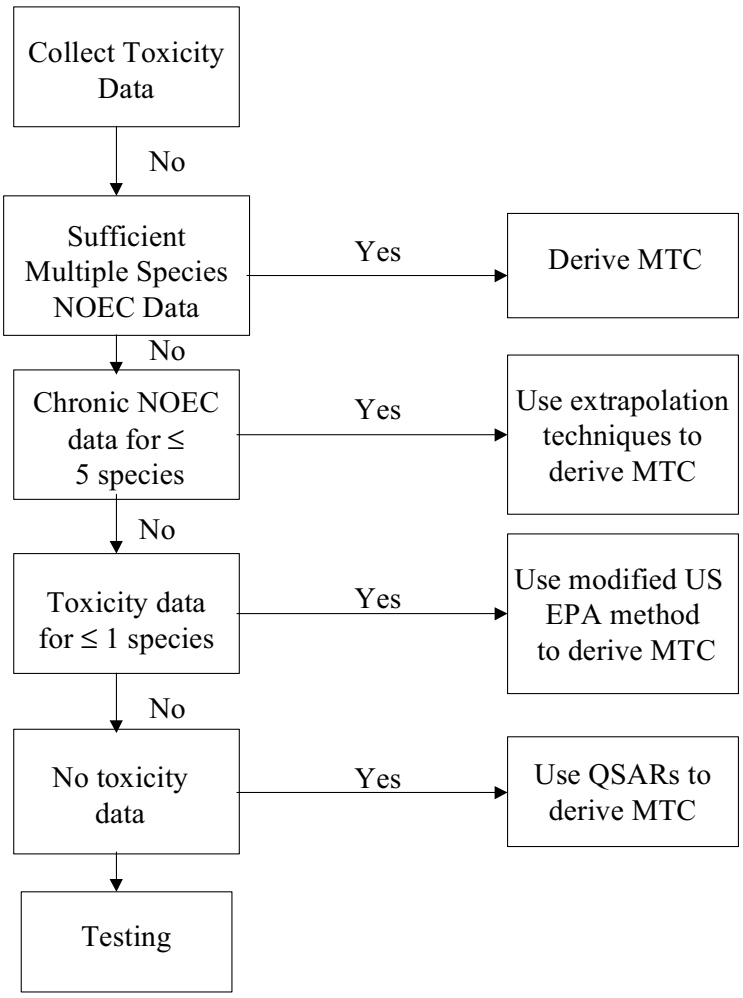
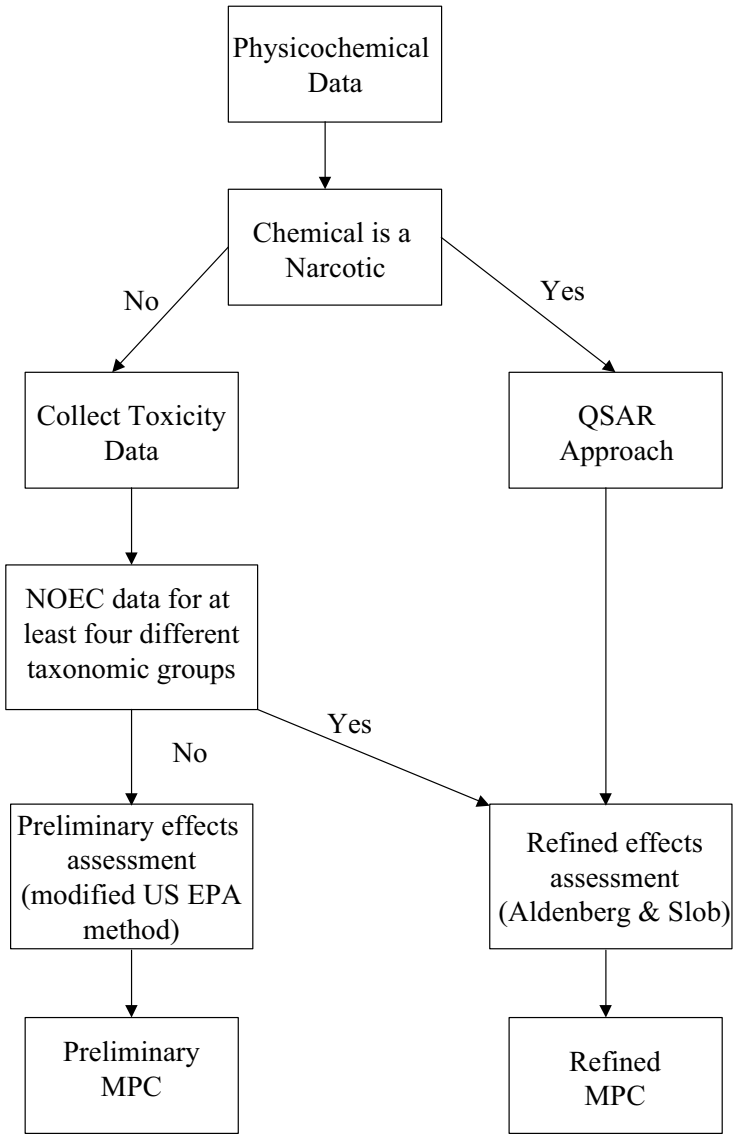


Figure 5 The OECD recommended method of aquatic effects assessment



(Modified from Van de Plassche *et al.* 1993)

Figure 6 The Netherlands scheme for deriving maximum permissible concentrations (MPCs) which are the equivalent of ANZECC water quality guidelines

9.2 The 'proposed assessment factor method'

The CCREM (1991) AF method used by ANZECC (1992) is set out in table 2 (section 4.1). The method uses chronic LOEC toxicity data. It was argued earlier in the review (section 4.1.1.2) that NOEC data is preferable to LOEC data for deriving WQGs. In view of this, it is proposed that the AF method used by ANZECC (1992) be changed to use NOEC data. It was also shown that the use of two assessment factors (ie 2 and 10) for non-persistent and persistent chemicals was scientifically invalid. Therefore, it is recommended that the assessment factor of two be removed. Thus, for acute toxicity data the assessment factor to be used is 100.

If there are less data than the minimum requirements that are recommended by the USEPA and OECD an assessment factor of 1000 should be used. In the proposed framework,

chemicals that fail to meet the minimum requirements can not have WQGs derived. It is however possible for an environmental concern level (ECL) to be derived that will provide an estimate of the aqueous concentration which if exceeded would be the cause for concern.

It is also proposed that the minimum data requirements of the AF method currently used by ANZECC (1992) be modified. The details of these modifications are presented in section 9.5.

9.3 Data required for deriving water quality guidelines

An array of toxicity and physicochemical data are required in order to derive WQGs. Once the data are obtained by searches, their suitability and quality should be determined and compared with the minimum data requirements of each method. The following section outlines these various processes.

9.3.1 Toxicity data searches

It is recommended that for the forthcoming review of the ANZECC WQGs that toxicity data be obtained by conducting searches of the USEPA AQUIRE (1994) database, the Australasian Ecotoxicology database, the in-house literature database of the Ecotoxicology Section of the NSW EPA and various water quality documents of the Dutch, Danish and USA. It is also recommended that only toxicity data published since 1980 be included in the search. Data prior to this is considered to be unreliable due to advances in experimental and analytical capabilities since that time. General physicochemical property data for the chemicals should be obtained from reliable sources such as Verscheuren (1983), Hansch et al (1995) and the ISIS, TOMES and HazChem databases. If the data are not available from these sources the literature should be searched. Finally various software programs could be used to estimate these values eg ASTER (Russom et al 1991).

9.3.2 Assessing the quality of single species toxicity data

The quality of the toxicity data from the AQUIRE database has already been assessed by the USEPA. They have a scheme whereby the information regarding how the toxicity data was generated are evaluated and a score given (see table 13). This scheme classes toxicity data into three categories: complete (C) with a score between 85 and 100; moderate (M) with a score of 51−84 and incomplete (I) with a score of 50 or less. It is recommended that only data with a score greater than 51 (ie C and M data) should be viewed as being of sufficient quality for inclusion in the derivation of Level I and Level II WQGs. Incomplete data were thought to be of such low quality that they should not be used for deriving any WQG.

There may be toxicity data that do not originate from AQUIRE. For example, data are available from the Dutch water quality documents (eg Van de Plassche et al 1993), the Danish water quality documents (Petersen & Pedersen 1995), the English environmental hazard assessment documents (eg Crookes & Howe 1993), the Australasian Ecotoxicology database and journal articles. The Dutch assess the quality of all the data they use and the method is presented in RIVM (1995). The Danish also assess the quality of the toxicity data but no details of the method are provided. It is proposed to accept toxicity data used by the Dutch and Danish as being equivalent to the C and M AQUIRE classes. The British documents do not mention any assessment of the toxicity data and data in the Australasian ecotoxicology database and journal articles are not assessed. It is therefore proposed that an assessment system based on the AQUIRE method should be used to evaluate this unassessed toxicity data (see Appendix 5).

Table 13 A summary of the scheme used in AQUIRE to assess the quality of the toxicity data

Information provided	Possible scores
Exposure duration	20 or 0
Control type	5 or 0
Organism characteristics	5 or 0
Chemical analysis method	5 or 0
Exposure type	5 or 0
Test location	4 or 0
Chemical grade	4 or 0
Test media	4 or 0
Hardness (for freshwater) or salinity (for saltwater)	2 or 0
Alkalinity (for freshwater) or salinity (for saltwater)	2 or 0
Dissolved oxygen	2 or 0
Temperature	2 or 0
PH	2 or 0
Endpoint	20 or 0
Trend of effect	5 or 0
Effect per cent	5 or 0
Statistical significance	4 or 0
Significance level	4 or 0
Total maximum score	100

Modified from AQUIRE (1994)

9.3.3 Assessing the quality of multiple species toxicity tests and data

In section 2.1.1 it was stated that multiple species (MS) toxicity tests yield the most environmentally realistic toxicity data. However, this assumes that the tests were well designed and conducted. If this is not the case the relevance of the data to the environment is reduced. Therefore, the quality of tests should be assessed before using the data to derive WQGs.

The OECD (1992a) provided some guidance on what constitutes a high quality MS test. In summary:

1. They should include fish or shellfish or the endpoints measured should directly relate to these species;
2. The test must represent the basic properties of ecosystems — photosynthesis, nutrient cycling and trophic structure;
3. There should at least be three dose treatments and one suitable control;
4. Chemical and physical properties that can affect exposure to the toxicant or the bioavailability should be measured;
5. Biological endpoints should cover individual, population and community levels;
6. The test should be of sufficient duration to account for the life-history of the organisms and the fate of the toxicant.

It is recommended that these OECD guidelines (OECD 1992a) be used to assess the suitability of using multiple species toxicity data to derive WQGs.

9.4 Australian and New Zealand toxicity data

To increase the ecological relevance of WQGs to Australia and New Zealand it is recommended that all local (Australian and New Zealand) toxicity data of suitable quality should be used to derive WQGs. Similarly, it is recommended that local and overseas data should be given equal weighting. Although some may argue that it would be better to only use local toxicity data or give it a higher weighting than overseas data, this is not recommended for three reasons. Firstly, it is not possible to only use local data because there are only sufficient local data to derive WQGs for a very limited number of chemicals. Secondly, it is not appropriate to increase the weighting of local data because it is so limited that it is unlikely to be representative. Increasing the weighting of local data could result in a skewing of the toxicity data and affect the resulting WQG. Thirdly, as the A&S method is statistically based it is better to have a larger sample size.

As a way of encouraging the generation of more local toxicity data and their use in deriving WQGs it is recommended that the origin of the data should be indicated in the WQG document (table 14). For example, it should be indicated whether there is any Australian and New Zealand toxicity data available for each chemical and whether local data was used in the derivation of the WQG.

Table 14 An example of the type of information that is recommended should be presented in the WQG tables

Chemical	Level I WQG[1] (μg/L)	Level II WQG[2] (μg/L)	Interim WQG[3] (μg/L)	Method used[4]	A&NZ data available[5]	A&NZ data used[6]
Toxicant 1	10	–	–	MS	N	N
Toxicant 2	5	–	–	A&S	Y	N
Toxicant 3	100	–	–	AF	Y	Y
Toxicant 4	–	30	–	A&S		
Toxicant 5	–	90	–	AF		
Toxicant 6	–	–	1000	QA&S		
Toxicant 7	–	–	0.1	QAF		
Toxicant 8	–	–	285	AF		

1 The WQG with the highest confidence that it provides adequate environmental protection.

2 The WQG with a high confidence that it provides adequate environmental protection.

3 The WQG with a low degree of confidence that it provides adequate environmental protection.

4 Method used: MS – the data comes from multiple species toxicity tests and the WQG is calculated using either the Aldenberg and Slob or the proposed assessment factor method, A&S – the data comes from single species toxicity tests and the WQG is calculated using the Aldenberg and Slob method, AF – the data comes from single species toxicity tests and the WQG is calculated using the proposed AF method, QA&S – the data comes from QSARs and the WQG is calculated using the Aldenberg and Slob method, QAF – the data comes from QSARs and the WQG is calculated using the proposed assessment factor method.

5. Y – there is Australian and/or New Zealand toxicity data for the chemical, N there is neither Australian nor New Zealand toxicity data for the chemical.

6. Y – Australian and/or New Zealand toxicity data was used to calculate the WQG, N neither Australian nor New Zealand toxicity data was used to calculate the WQG.

9.5 Minimum data requirements for deriving WQGs

Similar methods to those proposed in the above framework have been used overseas. Therefore, rather than re-inventing the wheel, the data requirements of these methods that are used overseas were assessed for their relevance to Australia and New Zealand.

The A&S method and the AF method used by ANZECC (1992) have different minimum data requirements (table 15). Due to these differences there is the distinct possibility that the

available data for a chemical may only permit one method to be used. Thus, a comparison of the outcomes of the A&S and AF methods in order to determine the WQG would not be possible. This comparison is a major feature of the proposed framework for deriving WQGs. In order to ensure this comparison occurs as often as possible it is recommended that the data requirements be modified so they are more consistent.

Table 15 The minimum data requirements of the assessment factor method used by ANZECC (1992) and Aldenberg and Slob methods

Method	Organism type	Minimum data requirements
AF method used by ANZECC (1992)	Fish	Toxicity data to three species of fish including at least one cold-water species and one warm-water species. Of these two should be chronic toxicity values.
	Invertebrates	Chronic toxicity data to two invertebrate species from different classes, one of which should be a planktonic species.
	Plants	One toxicity value for a vascular plant or algal species. For highly phytotoxic chemicals four acute and/or chronic studies on non-target plant or algal species.
A&S method	Not applicable	Chronic NOEC data from five different species that belong to at least four different taxonomic groups of organisms.

It is not appropriate to modify the data requirements of the A&S method to those of the AF method used by ANZECC (1992) for two reasons. Firstly, it would mean a reduction, from five to four, in the minimum number of chronic NOEC data required and it has been shown (OECD 1992a, Pedersen et al 1994) that when the A&S method is used on less than five data it is not reliable. Secondly, the A&S method assumes that the data are a random sample of species from the environment. As this assumption is already at least partly compromised (see section 4.2.1.1) it is not advisable to further compromise it by requiring a non-random selection of species.

The requirement of the AF method used by ANZECC (1992) that the toxicity data must be for certain types of organisms was examined to see if it could be made more consistent with the requirements of the A&S method. The minimum data requirements of the AF method used by ANZECC (1992) place a disproportionately large emphasis on toxicity data for fish and a disproportionately small emphasis on invertebrate organisms and no consideration is given to any other vertebrate aquatic organisms such as amphibians. The extent of this disproportionate representation of organisms is highlighted by the fact that only 5% of animals (Kingdom Animalia) have backbones (Barnes 1987) of which only a percentage would be fish and that there are more species belonging to the *Arthropoda phylum* (just one of the phyla that would be classified as invertebrates) than all other animal phyla combined (Kirk 1980).

The relative proportion of data required by the AF method used by ANZECC (1992) can be modified to more truly reflect the relative abundance of organisms (ie decrease the number of fish data and increase the number of non-fish data). If this were done then the minimum data requirements of the A&S method and the 'proposed AF method' would be essentially the same. This in turn would increase the probability that both methods could be used thus allowing comparisons of the outcomes. It is therefore recommended that the data requirements of the 'proposed AF method' be changed so they will be more consistent with the A&S method. The proposed minimum data requirements for both the modified AF method used by ANZECC and A&S method at the various levels of WQGs are presented on the next page.

For LEVEL I WQGs

A&S method: Requires chronic NOEC toxicity data for at least five different species that belong to at least four different taxonomic groups (see Appendix 6).

AF method: Requires chronic NOEC toxicity data for at least five different species that have the following composition. There must be one chronic fish NOEC, two chronic NOECs for different taxonomic groups of invertebrates (see Appendix 6), one chronic NOEC for a plant (see Appendix 6) and one chronic NOEC for any of the above groups or a different taxonomic group of organisms (see Appendix 6);

For LEVEL II WQGs

A&S method: Requires acute toxicity data (ie LC50 or EC50) for at least five different species that belong to at least four different taxonomic groups (see Appendix 6);

AF method: Requires acute toxicity data (ie LC50 or EC50) for at least five different species have the following composition. There must be one acute fish datum, two acute data for different taxonomic groups of invertebrates (see Appendix 6), one acute datum for a plant (see Appendix 6) and one acute datum for any of the above groups or a different taxonomic group of organisms (see Appendix 6);

For INTERIM WQGs

A&S method: Requires at least nineteen estimates of chronic toxicity derived by QSARs (Appendix 1). Any experimental data is included with the QSAR estimates.

AF method: Requires at least one toxicity value (either acute or chronic) for each of the following: a fish, an invertebrate and an algae (ie, a total of three toxicity data).

While these minimum data requirements should be applied generally, there may be individual cases where expert judgment may be used to vary the data requirements. Any such modifications to these rules should be noted and the reason for the change explained in the appendixes to the WQGs.

When toxicity data are to be used for the A&S method then the following rules should also apply:

- If several toxicity values are derived for the same effect, for a single species, the geometric mean of the values is calculated and is taken to represent the sensitivity of the species;
- If several toxicity values are derived for different effects, for a single species, the lowest NOEC is taken to represent the sensitivity of the species.

Neither the OECD (1995) nor The Netherlands (Van de Plassche et al 1993) have any minimum data requirements for multiple species toxicity data. The minimum data required, by this proposed framework for all other WQGs is three. Therefore it is proposed that toxicity data be available from at least three different multiple species tests that meet the quality standards (see section 9.3.3).

10 Implications of the proposed framework

Under the proposed framework for WQG derivation different methods may be used and therefore different levels of protection may be provided. The stated aim of the current ANZECC WQGs is

> to protect all forms of aquatic life and all aspects of the aquatic life cycle.... The intention is to protect all life stages during indefinite exposure to the water (ANZECC 1992).

In contrast, the HC5 95% value derived by Aldenberg and Slob aims to protect 95% of species with 95% confidence. For this reason the chances of HC5 95% values under-protecting the environment is low. This is illustrated by all forty of the HC5 values derived in this study theoretically protecting more than 95% of species. In fact, none of the forty HC5 95% values offered less protection than 97% of species.

The difference in the stated levels of protection offered by the two approaches highlights a major implication of the recommended framework. Should the proposed framework be adopted then a change in the stated level of protection may be required. It is argued that the difference is not so much in actual protection but in methodology and approach. Nevertheless, it could be argued that any change in the level of protection to be provided by the ANZECC WQGs was significant.

As it is virtually impossible for either the 'proposed AF method' or the A&S method to protect all organisms in the environment, perhaps the original aim of the 1992 ANZECC WQGs could be retained but become the long-term goal for the National Water Quality Management Strategy.

11 Conclusions

The assessment factor method divides the lowest toxicity value by an assessment factor, the magnitude of which is inversely proportional to the perceived quality of the toxicity data. The assessment factors are empirically chosen and there is debate over the validity of using the acute to chronic ratio. The various assessment factors used by ANZECC in 1992 do not appear to provide the level of protection that was the aim (ie all species). The Aldenberg and Slob method is a statistically based method that introduces risk assessment concepts to the derivation of WQGs and uses all the data rather than just the lowest value as in the assessment factor method. A number of assumptions of the method may be compromised or the subject of ongoing scientific debate. It is acknowledged that multiple species toxicity data obtained from microcosms, mesocosms and field studies are the most environmentally realistic toxicity data. However, the usefulness of such data in deriving WQGs will be limited by the minimal amounts available. Water quality guidelines were derived for forty chemicals using both methods. This revealed that the HC5 95% values derived by the Aldenberg and Slob method were, as a group, significantly lower than the values derived by the assessment factor method and offered a significantly greater degree of protection. However, when the WQGs for individual chemicals were examined the Aldenberg and Slob method did not always produce lower values than the modified USEPA assessment factor method and hence they did not always provide a higher degree of protection. Comparison of the WQGs derived as part of this study with the existing WQGs indicated that irrespective of the method used to derive them, reasonably large decreases might occur. However, it is also possible that the WQGs could be higher than the existing levels for some chemicals.

A new framework for deriving WQGs is proposed. This is a hierarchical scheme in which three different levels of WQGs (Level I, Level II and interim) can be derived. Which level of

WQG can be derived depends on the quantity and type of toxicity data available. The degree of confidence that the WQG will provide adequate environmental protection decreases from Level I, to level II to interim. An important feature of the proposed framework is that, generally, both the 'proposed assessment factor' and the Aldenberg and Slob methods are both used on the available toxicity data with the lowest value becoming the WQG providing the octanol-water partition coefficient and/or bioconcentration factor is less than 1000. If this is not the case then the lower value is adjusted so that it accounts for the potential for bioaccumulation with the adjusted value becoming the WQG. It is also proposed that the effect of water hardness on metal toxicity and the toxicity of mixtures be addressed.

Appendix 1 Quantitative structure-activity relationships (QSARs) used by the Dutch (Van de Plassche et al 1993)

Species	QSAR[1] (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 supcapitata (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/ 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

1 All QSARs were developed by Van Leeuwen et al (1992)

Rules for the use of these QSARs are provided in OECD (1995)

Appendix 2 Variation in Kβ values with sample size

The variation of Kβ with the number of data in the sample is revealed in the following table.

Sample size	Kβ Values	
	HC5 95%	HC5 50%
2	27.70	2.49
5	4.47	1.85
10	3.06	1.73
20	2.49	1.68
50	2.10	1.65
100	1.95	1.64
∝	1.62	1.62

Modified from Aldenberg & Slob 1993

Appendix 3 Comparison of WQGs derived using the ANZECC AF method and the Aldenberg and Slob method

The method used

In order for chemicals to be included in this comparison they needed to have toxicity data for four different taxonomic groups of organisms and have a total of at least five data points (Okkerman et al 1991, OECD 1992). Rules for selecting data prior to input into the Aldenberg and Slob method were:

- If several NOEC values are derived for different effects, for a single species, the lowest NOEC is used to derive the HC5 value;
- If several NOEC values are derived for the same effect, for a single species, the geometric mean of values is calculated and used to derive the HC5 value (Okkerman et al 1991).

The HC5 values were calculated using the ETX computer software system (Aldenberg 1993). The ANZECC WQGs were derived by dividing the lowest NOEC value by 10.

The No observed effect data used in this study are on the following pages.

Chemical	NOEC Values (µg/L)
1,1,1-trichloroethane[ad]	47, 180, 220, 1.3, 7.7
1,1,2-trichloroethane[ad]	47, 180, 220, 10, 13, 18
1,2-dichloroethane[ad]	68, 53, 360, 470, 560, 530, 11, 29
3-chloropropene[ad]	58, 4.1, 3.2, 4.3, 4.2
∝-HCH[ag]	3300, 80, 9, 20, 90, 800
β-HCH[ag]	500, 83, 320, 27, 180
γ-HCH[ag]	150, 250, 950, 500, 440, 330, 11, 4.3, 2.2, 9.1, 2.9, 8.8
Benzo(a)anthracene[ac]	0.093, 0.072, 0.021, 0.020, 0.050, 0.005, 0.004, 0.025, 0.009, 0.024, 0.023, 0.004
Benzopyrene[ac]	0.0063, 0.0053, 0.0165, 0.0154, 0.0034, 0.0031, 0.0194, 0.0181, 0.0173, 0.0027
Benzopyrlene[ac]	0.0012, 0.007, 0.007, 0.0005, 0.002
Berylium[af]	0.08, 0.001, 0.015, 0.002, 0.0085, 0.26, 0.00068
Cadmium[ae]	40, 650, 3100, 1500, 120, 700, 35, 5.5, 17, 2.5, 0.085, 1.0, 4.2, 4.2, 11, 3.0, 31, 4.3, 1.3, 37, 4.4, 1.0, 3.8, 0.2, 9.0, 3.0, 31, 9.0
Cadmium[be]	19, 0.97, 1.6, 0.75, 130, 14, 11, 3.3
Carbofuran[be]	24, 2.5, 0.5, 12, 1.7
Chlorpyrifos[be]	17, 100, 132, 9.2, 50, 16, 1.7, 0.46, 8.5

Appendix 3 cont

Chemical	NOEC Values (µg/L)
Chrysene[ac]	0.0083, 0.0240, 0.0224, 0.0561, 0.0057, 0.0050, 0.0282, 0.0106, 0.0267, 0.0258, 0.0044
Cobalt[af]	1.1, 0.005, 0.058, 0.5, 0.5
Copper[ae]	50, 50, 10, 10, 8.0, 8.0, 12, 4.0, 5.0, 20, 40, 5.0, 15, 34, 40, 8.0, 13, 43, 50, 35, 12, 104, 21, 37, 120, 21, 8.0, 11, 22, 3.0, 22, 13
Copper[be]	370, 1300, 210, 68, 40, 13
Dieldrin[ag]	10, 10, 32, 5.0, 0.12, 0.75
Endrin[ag]	0.067, 10, 0.1, 100, 0.03, 25, 0.12, 0.19
Ethylbenzene[ad]	6, 17, 1, 470, 70
Fluoranthene[ac]	0.20, 0.0123, 2.533, 0.941, 1.893, 1.482, 0.287, 0.027, 0.791, 0.066, 0.061, 0.138, 0.021, 0.017, 0.077, 0.075, 0.075, 0.016
Mercury[ae]	5.0, 35, 2.5, 8.0, 9.0, 0.50, 39, 32, 1.1, 0.02, 0.02, 0.31
Naphthalene[ac]	0.45, 0.26, 2.3, 0.37, 0.021, 129.4, 24.9, 9.2, 22.1, 152.7, 20.5, 1.3, 7.2, 15.4, 1.7, 1.5, 2.5, 1.4, 0.95, 1.9, 2.1, 2.2, 1.1
Pentachlorophenol[ae]	1000, 1000, 100, 1000, 32, 3.2, 50, 100, 23, 23, 3200, 32, 45, 100, 8.9, 32
Phenanthrene[ac]	0.60, 0.032, 0.15, 0.060, 0.042, 6.259, 2.326, 4.990, 4.166, 0.121, 1.012, 2.5288, 0.2319, 0.216, 0.429, 0.107, 0.273, 0.273, 0.280, 0.083
Selenium[af]	0.15, 0.03, 0.05, 0.16, 0.2, 0.005, 0.05, 0.2, 0.26, 1.3, 0.079, 1.0, 4.7, 5.0, 0.0009, 0.031, 0.059, 0.3, 5.6
Sodium arsenate[ah]	4860, 10, 86, 2350, 10 000, 2,400
Thiram[be]	3.8, 2.1, 120, 24, 15, 13, 22, 0.38, 0.6
Trichloroethane[ad]	33, 32, 180, 600, 5.8
Trichloromethane[ad]	63, 93, 550, 110, 9.7

a –toxicity data is for aquatic species

b –toxicity data is for soil species

c – Kalf et al (1995)

d – Van de Plassche et al (1993)

e – Van de Plassche (1994)

f – Van de Plassche & De Bruijn (1992)

g – Van de Plassche et al (1994)

h – Hesse et al (1990)

Appendix 4 Differences between concentration addition and response addition

Two models for toxic additivity exist: concentration additivity and response additivity. The model to be used can be determined based on the theoretical grounds developed by Plackett and Hewlett (1952) providing the mechanism of action of every chemical present is known. However, such information is very seldom available.

Determining the toxicity of a mixture assuming response addition is more complex and requires more information than the concentration addition model. Firstly, the concentration response relationship is needed for every chemical in the mixture. Using this relationship the biological response that corresponds to the concentration of each component in the mixture is determined. The biological responses (BR) are then summed for each component to determine the total effect. As stated in the text this method is not suitable for use with WQGs derived using any assessment factor method. The following method is only applicable to WQGs derived using statistically based extrapolation methods such as the Aldenberg and Slob method. The total effect of a mixture, assuming response addition, is calculated by:

$$\%S = \Sigma S_I \quad (1)$$

where %S is the total percentage of species (%S) likely to be affected and S_i is the percentage of organisms affected by chemical 'i'. If the %S value exceeds the policy level, equals the policy level or is less than the policy level then the WQG is exceeded, is equalled and not exceeded respectively.

The total toxicity of a mixture (TT), assuming concentration additivity, is calculated by:

$$TT = \Sigma Ci/Ti \quad (2)$$

where C_i is the concentration of each component of the mixture and T_i is any measure of the toxicity of each component eg LC50, EC50, LOEC or NOEC. However, the measure of toxicity used must be constant within any one mixture. If the value of TT exceeds 1, equals 1 or is less than 1 then the WQG is exceeded, is equalled and not exceeded respectively.

Appendix 5 The proforma used to assess the quality of toxicity data

This assessment scheme is based on that used by the USEPA AQUIRE database. The final two questions asked differ depending on whether the toxicity data is NOEC/LOEC or LC/EC type data.

Article Number:

Unique Identifier:

Chemical(s):

Question	Value given	
	By Marker	By Checker
Was the duration of the exposure stated?		
Were there appropriate controls (eg a solvent control if solvents are used)?		
Were the characteristics of the test organism stated?		
Were the chemical concentrations measured?		
Was the type of exposure (eg static, flow through) stated?		
Was the test location stated?		
Was the grade or purity of the test chemical stated?		
Was the type of test media used stated?		
Was the hardness (for freshwater) or the salinity (for saltwater) measured and stated?		
Was the alkalinity (for freshwater) or salinity (for saltwater) measured and stated?		
Was the dissolved oxygen content of the test water measured at some stage during or after the test?		
Was the temperature measured during the test?		
Was the pH of the test water measured at some time during the test?		
Was the biological endpoint clearly defined?		
Was there a concentration-response relationship either observable or stated?		
Was the biological effect quantified ie 50% effect, 25% effect?		
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)?		
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)?		
Total score		
Category (C, M, I)*		

Scored by: Date:

Checked by: Date:

Quality Categories: >85 C, 51-84 M, <50 I

Appendix 6 Taxonomically different types of organisms

The types of organisms that are considered taxonomically different when assessing whether the toxicity data meets the minimum data requirements for the Aldenberg and Slob method:

1. fish
2. crustaceans
3. insects
4. molluscs
5. annelids
6. echinoderms
7. rotifers
8. hydra
9. green algae
10. blue algae
11. red algae
12. macrophytes
13. blue-green algae (cyanobacteria)
14. amphibians
15. bacteria (except *Photobacterium phosphoreum/Vibrio fishceri*)
16. protozoans
17. coral
18. fungi

The types of organisms that are considered to be taxonomically different organisms when determining whether the minimum data requirements of the Assessment Factor Method have been met are set out in the following table.

Toxicity data from the Microtox ® system are not recommended for inclusion because the endpoint is a measure of a biochemical effect (see section 8.1 for further details).

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*/ *Photobacterium phosphoreum*), protozoans, coral, fungi and others

References

Abernethy SG, Mackay D & McCarty LS 1988. 'Volume fraction' correlation for narcosis in aquatic organisms: The key role of partitioning. *Environ. Toxicol. Chem*. 7, 469–481.

Aldenberg T 1993. ETX 1.3A: A program to calculate confidence limits for hazardous concentrations based on small samples of toxicity data. The Netherlands National Institute of Public Health and Environmental Protection, Bilthoven, The Netherlands.

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

Allan P 1995. Exerts from a speech to the Australian Centre for Environmental Law outlining the incoming New South Wales Government's Environmental Law Reform Agenda. Saturday 6 May, 1995. *Environmental and Planning Law Journal* 12, 339–344.

Altman PL & Dittmer DS 1962. Cited in ML Dourson & JF Stara 1983. Regulatory history and experimental support of uncertainty (safety) factors. *Regul. Toxicol. Pharmacol.* 3, 224–238.

Ankley GT & Mount DR 1996. Retrospective analysis of the ecological risk of contaminant mixtures in aquatic sediments. *Human Ecol. Risk Assess.* 2, 434–440.

ANZECC (Australian and New Zealand Environment and Conservation Council) 1992. *Australian Water Quality Guidelines For Fresh And Marine Waters.* ANZECC, Australia.

ANZECC & 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 and ARMCANZ, Canberra, Australia.

ANZECC (Australian and New Zealand Environment and Conservation Council) & NHRMC (National Health and Medical Research Council) 1992. *Australian and New Zealand Guidelines for the Assessment and Management of Contaminated Sites.* ANZECC and NHMRC, Australia.

AQUIRE (Aquatic Toxicity Information Retrieval Data Base) 1994. *AQUIRE Standard Operating Procedures*. USEPA, Washington, DC.

Auer CM, Nabholz JV & Baetcke KP 1990. Mode of action and the assessment of chemical hazards in the presence of limited data: Use of structure-activity relationships (SAR) under TSCA, Section 5. *Environ. Health Perspect.* 87, 183–197.

Baird DJ, Barber I & Calow P 1990. Clonal variation and general responses of *Daphnia magna* to toxic stress 1. Chronic life history effects. *Functional Ecology* 4, 399–407.

Baker J 1990. Cited in J Cameron 1993. The Precautionary Principle: Core meaning, constitutional framework and procedures for implementation. *Proceedings of the Precautionary Principle Conference*, Institute of Environmental Studies, University of New South Wales, 20–21 September, 1–21.

Barnes RD 1987. *Invertebrate zoology*. Saunders College Publishing, Philadelphia, USA.

Barnthouse LW, Suter GW II & Rosen AE 1990. Risks of toxic contaminants to exploited fish populations: Influence of life history, data uncertainty and exploitations intensity. *Environ. Toxicol. Chem*. 9, 297–311.

Biodiversity Working Party 1991. *The conservation of biodiversity as it relates to ecologically sustainable development.* ESD Secretariat, DASETT, Canberra, Australia. Cited in ANZECC (Australian and New Zealand Environment and Conservation Council). 1992. *Australian Water Quality Guidelines For Fresh And Marine Waters.* ANZECC, Canberra, Australia.

Bliss CI 1934a. The method of probits. *Science* 79, 38−39.

Bliss CI 1934b. The method of probits: A correction. *Science* 79, 409−410.

Broderius S & Kahl M 1985. Acute toxicity of organic chemical mixtures to the fathead minnow. *Aquatic Toxicology* 6, 307−322.

Bro-Rasmussen F, Calow P, Canton JH, Chambers PL, Silva Fernandes A, Hoffmann L, Jouany JM, Klein W, Persoone G, Scoullos M, Tarazona JV & Vighi M 1994. EEC water quality objectives for chemicals dangerous to aquatic environments (List 1). *Reviews of Environ. Contam. Toxicol.* 137, 83−110.

Cairns J Jr 1995. The case for ecosystem services as toxicological end points. *Human Ecol. Risk Assess*. 1, 171−174.

Calabrese EJ & Baldwin LA 1993. Chemical-specific ecosystem MATC. In *Performing ecological risk assessments*, Lewis Publishers, Boca Raton, USA. 165−183.

Calamari D & Vighi M 1992. A proposal to define quality objectives for aquatic life for mixtures of chemical substances. *Chemosphere* 25, 531−536.

Calow P 1995. Risk assessment: Principles and practice in Europe. *Australasian. J. Ecotoxicol* 1, 11−15.

Cameron J 1993. The Precautionary Principle: Core meaning, constitutional framework and procedures for implementation. *Proceedings of the Precautionary Principle Conference*, Institute of Environmental Studies, University of New South Wales, 20−21 September. 1−21.

CCREM (Canadian Council of Resource and Environment Ministers) 1991. *Canadian Water Quality Guidelines.* Appendix IX. Canadian Council of Resource and Environment Ministers, Inland Water Directorate. Environment Canada, Ottawa, Canada. p. IX-1 to IX-8.

Chapman JC 1995a. The role of ecotoxicity testing in assessing water quality. *Aust. J. Ecol.* 20, 20−27.

Chapman PM 1995b. How should numerical criteria be used? *Human Ecol. Risk Assess.* 1, 1−4.

Chapman PM 1995c. How useful are single-species toxicity tests? *Human Ecol. Risk Assess.* 1, 163−166.

Chapman PM, Cardwell RS & Chapman PF 1996. A warning: NOECs are inappropriate for regulatory use. *Environ. Toxicol. Chem.* 15, 77−79.

Clements RG & Nabholz JV 1994. *ECOSAR. A computer program for estimating the ecotoxicity of industrial chemicals based on structure activity relationships. User's Guide.* Report No. 748-R-93-002. USEPA, Washington DC, USA.

Connell DW 1989. Biomagnification by aquatic organisms: A proposal. *Chemosphere* 19, 1573−1584.

Connell DW 1990. *Bioaccumulation of xenobiotic compounds.* CRC Press, Boca Raton, Florida, USA.

Cotruvo JA 1988. Cited in FR Johannsen 1990. Risk assessment of carcinogenic and noncarcinogenic chemicals. *Crit. Rev. Toxicol.* 20, 341−367.

Crookes MJ & Howe PD 1993. *Environmental hazard assessment: Halogenated naphthalenes*. Department of the Environment, England.

D'Agostino RB & Stephens MA 1986. *Goodness-of-fit techniques.* Marcel Dekker, New York.

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, 1253−1262.

Delos CG 1995. Possible revisions to EPA's procedure for deriving aquatic life criteria. *Proceedings of the Water Environment Federation, Vol 4.* 67th Conference and Exposition, October 15−19, 1994, Chicago, Illinois, USA. 661−667.

Deneer JW, Sinnige TL, Seinen W & Hermens JLM 1988. The joint acute toxicity to *Daphnia magna* of industrial organic chemicals at low concentrations. *Aquatic Toxicology* 12, 33−38.

De Zwart D & Slooff W 1987. Toxicity of mixtures of heavy metals and petrochemicals to *Xenopus laevis. Bull. Environ. Contam. Toxicol* 38, 345−351.

Donkin P 1994. Quantitative structure-activity relationships. In *Handbook of Ecotoxicology*, *Volume II*. ed P Calow, Blackwell Scientific Publications, Oxford, England, 321−347.

Dourson ML & Stara JF 1983. Regulatory history and experimental support of uncertainty (safety) factors. *Regul. Toxicol. Pharmacol.* 3, 224−238.

Emans HJB, Van de Plassche EJ, Canton JH, Okkerman PC & Sparenburgs PM 1993. Validation of some extrapolation methods for effects assessment. *Environ. Toxicol. Chem.* 12, 2139−2154.

Escande J-P 1994. The spirit of René Dubos, or the dangers of a fragmentary approach to ecology. In *Environment and health: A holistic approach*, ed R Kreips, Avebury, Aldershot, United Kingdom, 5−13.

ESD Steering Committee of the Department of Prime Minister and Cabinet 1992. *Draft National Strategy for Ecologically Sustainable Development*. Australian Government Publishing Service, Canberra, Australia.

Evans RD, Harris RS & Bunker JWM 1944. Radium metabolism in rats, and the production of osteogenic sarcoma by experimental radium poisoning. *Amer. J. Roentenol* 52, 353−373.

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

Fredenslund FC, Severinsen M & Andersen MB 1995. *Evaluation of the simple box model for Danish conditions.* Environmental Project No. 307. Danish Ministry of the Environment and Energy, Danish Environment Protection Agency, Copenhagen, Denmark.

Galassi S, Guzella L & Battegazzore M 1994. Biomagnification of PCBs, ppDDE and HCB in the river Po ecosystem (Northern Italy). *Ecotoxicol. Environ. Saf* 29, 174−186.

Gelber RD, Lavin PT, Mehta CR & Schonfield DA 1985. Statistical analysis. In *Fundamentals of aquatic toxicology*, eds GM Rand & SR Petrocelli, Hemisphere Publishing Corp., New York, USA, 110−123.

Gobas FAPC, Zhang X & Wells R 1993. Gastrointestinal magnification: The mechanism of biomagnification and food chain accumulation of organic chemicals. *Environ. Sci. Technol.* 27, 2855−2863.

Goldberg L 1975. Safety evaluation concepts. *J. Assoc. Off. Anal. Chem.* 58, 635−644.

Graney RL, Giesy JP & Clark JR 1995. Field studies. In *Fundamentals of aquatic toxicology. Effects, environmental fate and risk assessment*, ed GM Rand. Taylor and Francis, Washington DC, USA. 257−306.

Grimme LH, Faust M, Bodecker W & Altenburger R 1996. Aquatic toxicology of chemical substances in combinations: Still a matter of controversy. *Human Ecol. Risk Assess* 2, 426−433.

Guckert JB 1993. Artificial streams in ecotoxicology. *J. N. Amer. Bentholog. Soc.* 12, 313−384.

Hamilton MA, Russo RC & Thurston RV 1977. Method for estimating median lethal concentrations in toxicity bioassays. *Environ. Sci. Technol*. 11, 714−719.

Hansch C & Leo A 1995. *Exploring QSAR*. American Chemical Society, Washington DC, USA, 169−217.

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

Hart BT 1974. *A Compilation of Australian Water Quality Criteria.* Australia Water Resources Council, Technical Paper 7, Australian Government Publishing Service, Canberra, Australia.

Hart BT 1982. *Australian water quality criteria for heavy metals.* Australian Water Resources Council, Technical Paper 77, Australian Government Publishing Service, Canberra, Australia.

Hart BT 1996. *A methodology for deriving aquatic guideline values for toxic contaminants*. First Draft. Report to the New Zealand Ministry for the Environment. Water Studies Centre, Monash University, Melbourne, Australia.

Hart BT, Jones MJ & Chapman JC 1995. *A process for the development of guidelines for the protection of aquatic life in New Zealand.* Report to the New Zealand Ministry for the Environment. Water Studies Centre, Monash University, Melbourne, Australia.

Hart WB, Doudoroff P & Greenbank J 1945. Cited in GM Rand, PG Wells & LS McCarty 1995. Introduction to aquatic toxicology. In *Fundamentals of aquatic toxicology. Effects, environmental fate and risk assessment.* 2nd Edition, ed GM Rand. Taylor and Francis, Washington DC, USA, 3−67.

Hayes WJ 1967. Toxicity of pesticides to man: Risks from present levels. *Proc. Royal Soc. London* 167, 101−127.

Hermens J 1989. Quantitative structure-activity relationships of environmental pollutants. In *Handbook of environmental chemistry, Vol. 2E: Reactions and processes*. ed O Hutzinger, Springer-Verlag, Berlin, Germany, 111−162.

Hermens J, Broekhuyzen E, Canton H & Wegman R 1985. Quantitative structure-activity relationships and mixture toxicity studies of alcohols and chlorohydrocarbons: Effects on growth of *Daphnia magna. Aquatic Toxicology* 6, 209−217.

Hermens J, Leeuwangh P & Musch A 1984. Quantitative structure-activity relationships and mixture toxicity studies of chloro- and alkyl-anilines at an acute lethal toxicity level to the guppy *(Poecillia reticulata). Ecotoxicol. Environ. Saf.* 8, 388−394.

Hesse JM, Janus JA, Krajnc EI & Kroese ED 1990. Integrated criteria document arsenicum effects. Appendix to Report No. 758701002. National Institute for Public Health and Environmental Protection, Bilthoven, The Netherlands.

Hoekstra JA & Van Ewijk PH 1993. Alternatives for the no-observed effect level. *Environ. Toxicol. Chem* 12, 187−194.

Holdway DA 1996. The role of biomarkers in risk assessment. *Human Ecol. Risk Assess*. 2, 263−267.

Johannsen FR 1990. Risk assessment of carcinogenic and noncarcinogenic chemicals. *Crit. Rev. Toxicol.* 20, 341−367.

Johnston N, Skidmore J & Thompson G 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 Council.

Jones M, Folt C & Guarada S 1991. Characterising individual, population and community effects on sub-lethal levels of aquatic toxicants: An experimental case study using *Daphnia. Freshwater Biology*. 26, 35−44.

Jongbloed RH, Traas TP & Luttik R 1996. A probabilistic model for deriving soil quality criteria based on secondary poisoning of top predators. II. Calculations for dichlorodiphenyltrichloroethane (DDT) and Cadmium. *Ecotoxicol. Environ. Saf.* 34, 279−306.

Kalf DF, Crommentuijn GH, Posthumus R & Van de Plassche E 1995. *Integrated environmental quality objectives for polycyclic aromatic hydrocarbons (PAHs)*. Report No. 679101018. National Institute of Public Health and the Environment, Bilthoven, The Netherlands.

Kenaga EE 1978. Test organisms and methods useful for the early assessment of acute toxicity of chemicals. *Environ. Sci. Technol*. 12, 3122−1329.

Kenaga EE 1979. Aquatic test organisms and methods useful for assessment of chronic toxicity of chemicals. In *Analyzing the hazard evaluation process*. eds KL Dickson, AW Maki & JJr Cairns, American Fisheries Society, Washington DC, USA. 101−111.

Kenaga EE 1982. Predictability of chronic toxicity from acute toxicity of chemicals in fish and aquatic invertebrates. *Environ. Toxicol. Chem*. 1, 347−358.

Kettle B 1993. Statistics, pragmatism and the Precautionary Principle: Exploring the boundaries of scientific method. In *Proceedings of the Precautionary Principle Conference*. Institute of Environmental Studies, University of New South Wales, 20−21 September, 1−11.

Kimerle RA, Werner AF & Adams WJ 1983. Aquatic hazard evaluation principles applied to the development of water quality criteria. In *Aquatic toxicology and hazard assessment:*

Seventh Symposium. ASTM STP 854. eds RD Cardwell, R Purdy & RC Bahner, American Society for Testing and Materials, Philadelphia, USA, 538−547.

Kirk DL 1980. *Biology today*. Random House, Washington DC, USA.

Könemann H 1980. Structure-activity relationships and additivity in fish toxicities of environmental pollutants. *Ecotoxicol. Environ. Saf* 4, 415−421.

Könemann H 1981. Quantitative structure-activity relationships in fish toxicity studies. 1. Relationship for 50 industrial pollutants. *Toxicol*. 19, 209−221.

Kooijman SALM 1987. A safety factor for LC50 values allowing for differences in sensitivity among species. *Water Research* 21, 269−276.

Kraak MHS, Lavy D, Schoon H, Toussaint M, Peeters HM & Van Straalen NM 1994. Ecotoxicity of mixtures of metals to the zebra mussel *Dreissena polymorpha. Environ. Toxicol. Chem.* 13, 109−114.

Krasovskij GN 1976. Extrapolation of experimental data from animals to man. *Environ. Health Perspect.* 13, 51−63.

Kuhn R, Pattard M, Pernak KD & Winter A 1989a. Results of the harmful effects of selected water pollutants (anilines, phenols, aliphatic compounds) to *Daphnia magna. Water Res*. 23, 495−499.

Kuhn R, Pattard M, Pernak KD & Winter A 1989b. Results of the harmful effects of water pollutants to *Daphnia magna* in the 21 day reproduction test. *Water Res.* 23, 501−510.

La Point TW & Perry JA 1989. Use of experimental ecosystems in regulatory decision making. *Environ. Manag*. 13, 539−544.

Landis WG & Yu M-H 1995. *An introduction to environmental toxicology. Impacts of chemicals upon ecological systems*. Lewis Publishers, Boca Raton, USA, 115−134.

Le Blanc GA 1984. Interspecies relationships in acute toxicity of chemicals to aquatic organisms. *Environ. Toxicol. Chem*. 3, 47−60.

Lee G, Ellersieck MR, Mayer FL & Krause GF 1995. Predicting chronic lethality of chemicals to fishes from acute toxicity test data: Multifactor probit analysis. *Environ. Toxicol. Chem*. 14, 345−349.

Lipman Z 1993. Institutional reform: The New South Wales Environment Protection Authority. *Environmental and Planning Law Journal* 9, 445−457.

Mance G 1987. Factors affecting toxicity. In *Pollution threat of heavy metals in aquatic environments.* Elsevier Applied Science, London, 233−246.

Marcus MD & McDonald LL 1992. Evaluating the statistical basis for relating receiving water impacts to effluent and ambient toxicities. *Environ. Toxicol. Chem.* 11, 1389−1402.

Markich SJ & Camilleri C 1997. *Investigation of metal toxicity to tropical biota: Recommendations for revision of the Australian Water Quality Guidelines.* A Report to the Australian and New Zealand Environment and Conservation Council, Project 95/7.

Mayer FL 1990. *Predicting chronic lethality of chemicals to fishes from acute toxicity test data. Research and development*. Report No. EPA/600/X-90/147. USEPA, Gulf Breeze, Florida, USA.

Mayer FL & Ellersieck MR 1986. *Manual of acute toxicity: Interpretation and data base for 410 chemicals and 66 species of freshwater animals.* US Department of Interior, Fish and Wildlife Service Resource Publication 160, Washington, DC.

Mayer FL, Krause GF, Buckler DR, Ellersieck MR & Lee G 1994. Predicting chronic lethality of chemicals to fishes from acute toxicity test data: Concepts and linear regression analysis. *Environ. Toxicol. Chem.* 13, 671−678.

Mayer FL, Krause GF, Ellersieck MR & Lee G 1992. *Statistical approach to predicting chronic toxicity of chemicals to fishes from acute toxicity test data.* PB92-169655; PB92-503119 (software). National Technical Information Service, Springfield, Vancouver, USA.

Mayer FL, Mayer KS & Ellersieck MR 1986. Relationship of survival to other endpoints and chronic toxicity tests with fish. *Environ. Toxicol. Chem.* 5**,** 737−748.

McCarty LS & Mackay D 1993. Enhancing ecotoxicological modelling and assessment. *Environ. Sci. Technol.* 27, 1719−1728.

McCarty LS & Munkittrick KR 1996. Environmental biomarkers in aquatic toxicology: Fiction, fantasy or functional? *Human Ecol. Risk Assess*. 2, 268−274.

McPherson CA 1995. The role (or lack thereof) of laboratory bioassay tests. *Human Ecol. Risk Assess.* 1, 175−177.

Moore DRJ & Caux PY 1997. Estimating low toxic effects. *Environ. Toxicol. Chem*. 16, 794−801.

Mortimer M & Connell DW 1994. Critical internal and aqueous lethal concentrations of chlorobenzenes with the crab *Portunus pelagicus. Ecotox. Environ. Saf.* 28, 298−312.

Mortimer M & Connell DW 1995. Effects of exposure to chlorobenzenes on growth rates of the crab *Portunus pelagicus. Ecotoxicol. Environ. Saf*. 29, 1881−1885.

Mount DI 1994. Development and current use of single species aquatic toxicity tests. In *Ecological toxicity testing: Scale, complexity and relevance.* eds JJr Cairns & BR Niederlehner, Lewis Publishers, Boca Raton, Florida, USA, 97−104.

Mount DI & Stephan CE 1967. A method for establishing acceptable limits for fish: Malathion and the butoxyethanol ester of 2,4-D. *Trans. Am. Fish. Soc*. 96, 185−193.

Mulhall AM 1997. Models to predict the toxicity of selected phenols and benzenamines to a cladoceran and a marine bacterium*.* Honours Thesis, University of Technology Sydney.

Nabholz JV, Clements RG, Zeeman MG, Osborn KC & Wedge R 1993. Validation of structure activity relationships used by the USEPA's Office of Pollution Prevention and toxics for the environmental hazard assessment of industrial chemicals. In *Environmental toxicology and risk assessment: 2nd Volume. ASTM STP 1216*. eds JW Gorsuch, FJ Dwyer, CG Ingersoll & TW La Point. American Society for Testing and Materials, Philadelphia, USA, 571−590.

Napier G 1992. Application of laboratory derived data to natural aquatic ecosystems*.* PhD thesis, Macquarie University, Sydney, Australia.

New Zealand Ministry of the Environment 1996*. A proposed methodology for deriving aquatic guideline values for toxic contaminants. Document for public comment*. New Zealand Ministry of the Environment, New Zealand.

Nichols JW, Larsen CP, McDonald ME, Neimi GJ & Ankley GT 1995. Bioenergetics-based model for accumulation of polychlorinated biphenyls by nestling tree swallows, *Tachycineta bicolor*. *Environ. Sci. Technol.* 29, 604−612.

Nicholson BC 1984. *Water quality criteria for organic compounds.* Australian Water Resources Council, Technical Paper 82, Australian Government Publishing Service, Canberra, Australia.

Nirmalakhandan N & Speece RE 1988. Structure-activity relationships. Quantitative techniques for predicting the behaviour of chemicals in the ecosystem. *Environ. Sci. Technol*. 22, 606−615.

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.

NSW EPA (New South Wales Environment Protection Authority) 1996. *Discussion Paper NSW State Groundwater Policy*. NSW EPA, Sydney, Australia.

NSW POEA (Protection of the Environment Administration) Act 1991. (NSW), s6(2)(a).

OECD (Organisation for Economic Co-operation and Development) 1992a. *Report of the OECD Workshop on extrapolation of laboratory aquatic toxicity data to the real environment.* OECD Environment Monographs 59, OECD, Paris, France.

OECD (Organisation for Economic Co-operation and Development) 1992b. *Report of the OECD Workshop on quantitative structure-activity relationships (QSARS) in aquatic effects assessment*. OECD Environment Monographs 58, OECD, Paris, France.

OECD (Organisation for Economic Co-operation and Development) 1995. *Guidance document for aquatic effects assessment.* OECD Environment Monographs 92, OECD, Paris, France.

Okkerman PC, Van de Plassche EJ, Emans HJB & Canton JH 1993. Validation of some extrapolation methods with toxicity data derived from multiple species experiments. *Ecotoxicol. Environ. Saf.* 25, 341−359.

Okkerman PC, Van de Plassche EJ, Slooff W, Van Leeuwen CJ & Canton JH 1991. Ecotoxicological effects assessment: A comparison of several extrapolation procedures. *Ecotoxicol. Environ. Saf.* 21, 182−191.

Parkhurst BR 1994. Are single species toxicity test results valid indicators of effects to aquatic communities? In *Ecological toxicity testing: Scale, complexity and relevance,* eds JJr Cairns & BR Niederlehner, Lewis Publishers, Boca Raton, Florida, USA, 105−121.

Patra RWR, Chapman JC, Lim R & Gehrke PC 1995. Effects of temperature on the acute toxicity of endosulfan to silver perch, *Bidyanus bidyanus*. Poster Abstract PW 129. In *Proceedings Second SETAC World Congress (16th Annual meeting)*, Vancouver BC, Canada, 5−9 November 1995. p. PW 252.

Patra RWR, Chapman JC, Lim R & Gehrke PC 1996. Effects of temperature on the toxicity of endosulfan, chlorpyrifos and phenol to Australian *Ceriodaphnia dubia*. Abstract 0112. In *Proceedings of InterSect '96, International Symposium on Environmental Chemistry and Toxicology*, Sydney Australia, 14−18 July 1996. p. O112.

Pawlisz AV & Peters RH 1993. A test of the equipotency of internal burdens of nine narcotic chemicals using *Daphnia magna. Environ. Sci. Technol*. 27, 2801−2806.

Pedersen F, Kristensen P, Damborg A & Christensen HW 1994. *Ecotoxicological evaluation of industrial wastewater.* Miljøprojekt nr. 254. Ministry of Environment, Copenhagen, Denmark.

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

Plackett RL & Hewlett PS 1952. Quantal response of mixtures of poisons. *J. Roy. Stat. Soc. B* 14, 141−163.

Rabbidge J 1995. Evaluation of extrapolation methods for aquatic toxicity data used to derive water quality guidelines: Implications to Australia. MSc Thesis. University of Technology, Sydney, Australia.

Rand GM, Wells PG & McCarty LS 1995. Introduction to aquatic toxicology. In *Fundamentals of aquatic toxicology. Effects, environmental fate and risk assessment.* 2nd Edition. ed GM Rand. Taylor and Francis, Washington DC, USA, 3−67.

Reilly WK 1990. Cited in Z Lipman. 1993. Institutional reform: The New South Wales Environment Protection Authority. *Environmental and Planning Law Journal* 9, 445−457.

Rensvik H 1994. Who sets the environmental standards for tomorrow's industry: Industry, consumers, government, regulators, the green lobby? *Mar. Polltn. Bulltn.* 29, 277−278.

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.

Romijn CAFM, Luttik R & Canton JH 1994. Presentation of a general algorithm to include effect assessment on secondary poisoning in the derivation of environmental quality criteria. Part 2. Terrestrial food chains. *Ecotoxicol. Environ. Saf.* 27, 107−127.

Romijn CAFM, Luttik R, Van de Meent D, Slooff W & Canton JH 1993. Presentation of a general algorithm to include effect assessment on secondary poisoning in the derivation of environmental quality criteria. Part 1. Aquatic food chains. *Ecotoxicol. Environ. Saf.* 26, 61−85.

Rose R, Warne MStJ & Lim R 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.

Ross H 1996. The interaction of chemical mixtures and their implications on water quality guidelines. MSc Thesis, University of Technology Sydney.

Roux DJ, Jooste SHJ & 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.

Russom CL, Anderson EB, Greenwood BE & Pilli A 1991. ASTER: An integration of the AQUIRE database and the QSAR system for use in ecological risk assessments. In *QSAR in environmental toxicology – IV. Proceedings of the Fourth International Workshop*. Veldhoven, The Netherlands, 16−20 September, 1990, Elsevier, Amsterdam, The Netherlands, 667−669.

Schudoma D 1994. Derivation of water quality objectives for hazardous substances to protect aquatic ecosystems: Single-species test approach. *Environ. Toxicol. Water Qual.* 9, 263−272.

Seegert G, Fava A & Cumbie PM 1985. How representative are the data sets used to derive national water quality standards, In *Aquatic toxicology and hazard assessment: Seventh symposium.* eds RD Cardwell, R Purdy & RC Bahner. STP854, American Society for Testing and Materials, Philadelphia, USA, 527−537.

Simonis U 1993. Environmental policy in the Federal Republic of Germany: Curative and precautionary approaches. In *Proceedings of the Precautionary Principle Conference*, Institute of Environmental Studies, University of New South Wales, 20−21 September, 1−14.

Skidmore J & Firth IC 1983. *Acute sensitivity of selected Australian freshwater animals to copper and zinc.* Technical Paper No. 81. Australian Water Resources Council, Australian Government Publishing Service, Canberra, Australia.

Slooff W, Canton JH & Hermens JLM 1983. Comparison of the susceptibility of 22 freshwater species to 15 chemical compounds. 1. (Sub)acute toxicity tests. *Aquat. Toxicol*. 4, 113−128.

Slooff W, Van Oers JA & De Zwart D 1986. Margins of uncertainty in ecotoxicological hazard assessment. *Environ. Toxicol. Chem.* 5, 841−852.

Smith EP & Cairns JJr 1993. Extrapolation methods for setting ecological standards for water quality: Statistical and ecological concerns. *Ecotoxicol.* 2, 203−219.

Sprague JB 1976. Current status of sub-lethal tests of pollutants on aquatic organisms. *J. Fish. Res. Board. Can.* 33, 1988−1992.

Sprague JB 1995. A brief critique of today's use of aquatic toxicity tests. *Human Ecol. Risk Assess* 1, 167−171.

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

Stephan CE & Rogers JW 1985. Advantages of using regression analysis to calculate results of chronic toxicity tests. In *Aquatic toxicology and hazard assessment.* STP 891. eds RC Bahner & DJ Hansen, American Society for Testing and Materials, Philadelphia, PA, USA, 328−338.

Sun K, Krause GF, Mayer FL, Ellersieck MR & Basu AP 1995. Predicting chronic lethality of chemicals to fishes from acute toxicity test data: Theory of accelerated life testing. *Environ. Toxicol. Chem*. 14, 1745−1752.

Sunderam RIM, Cheng DMH & Thompson GB 1992. Toxicity of Endosulfan to native and introduced fish in Australia. *Environ. Toxicol. Chem*. 11, 1469−1476.

Sunderam RIM, Thompson GB, Chapman JC & Cheng DMH 1994. Acute and chronic toxicity of Endosulfan to two Australian cladocerans and their applicability in deriving water quality criteria. *Arch. Environ. Contam. Toxicol.* 27, 541−545.

Suter GW II & Rosen AE 1988. Comparative toxicology for risk assessment of marine fishes and crustaceans. *Environ. Sci. Technol.* 22, 548−556.

Suter GW II, Rosen AE, Lindner E & Parkhurst DF 1987. Endpoints for responses of fish to chronic toxic exposures. *Environ. Toxicol. Chem.* 6, 793−809.

Thomann RV 1981. Equilibrium model of fate of microcontaminants in diverse aquatic food chains. *Canadian J. Aquat. Sci.* 38, 280−296.

Thomann RV & Connolly JP 1984. Model of PCB in the Lake Michigan trout food chain. *Environ. Sci. Technol.* 18, 65−71.

Traas TP, Luttik R & Jongbloed RH 1996. A probabilistic model for deriving soil quality criteria based on secondary poisoning of top predators. 1. Model description and uncertainty analysis. *Ecotoxicol. Environ. Saf.* 34, 264−278.

UNCED 1992. *Global partnership for environment and development.* Guide to Agenda 21. UNCED, Geneva Office.

Underwood AJ 1995. Toxicological testing in laboratories is not ecological testing of toxicology. *Human Ecol. Risk Assess.* 1, 178−182.

USEPA (United States Environmental Protection Agency) 1976. *Quality criteria for water.* Report No. EPA-440/9-76/023. Office of Water Planning and Environmental Protection Agency, Washington DC, USA.

USEPA (United States Environmental Protection Agency) 1984a. *Estimating 'concern levels' for concentrations of chemical substances in the environment.* Environmental Effects Branch Health and Environmental Review, Washington DC, USA.

USEPA (United States Environmental Protection Agency) 1984b. *Development of water quality-based permit limitations for toxic pollutants; National policy. Federal Register* 49, 9016−9019.

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) 1988. *Estimating toxicity of industrial chemicals to aquatic organisms using structure activity relationships. Volume 1.* Report No. 5606/88/00. USEPA, Washington DC, USA.

USEPA (United States Environmental Protection Agency) 1991. *Technical support document for water quality-based toxics control.* Report No. EPA 505/2-90-001. Washington, DC.

USEPA (United States Environmental Protection Agency) 1994. *Water quality standards handbook*. 2nd Edition. EPA Report No. EPA-823-B-94-0059. USEPA, Washington DC, USA.

Van de Meent D & De Bruijn JHM 1995. A modelling procedure to evaluate the coherence of independently derived environmental quality objectives for air, water and soil. *Environ. Toxicol. Chem* 14, 177−186.

Van de Plassche E 1994. *Towards integrated environmental quality objectives for several compounds with a potential for secondary poisoning.* Report No. 679101 012. National Institute of Public Health and Environment Protection, Bilthoven, The Netherlands.

Van de Plassche E, Canton JH, Eijs YA, Everts JW, Jansseen PJCM, Van Koten-Vermeulen JEM, Polder MD, Posthumus R & de Stoppelaar JM 1994. *Towards integrated environmental quality objectives for several compounds with a potential for secondary poisoning: Underlying data.* Annex to Report No. 679101 012. National Institute of Public Health and Environment Protection, Bilthoven, The Netherlands.

Van de Plassche EJ & De Bruijn JHM 1992. *Towards integrated environmental quality objectives for surface water, ground water, sediment and soil for nine trace metals.* RIVM Report No. 679101 005, National Institute of Public Health and Environment Protection, Bilthoven, The Netherlands.

Van de Plassche EJ, Polder MD & 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 der Waterbeemd H & Testa B 1987. The parameterisation of lipophilicity and other structural properties in drug design. *Adv. Drug Res.* 16, 85−125.

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

Van Leeuwen CJ, Verhaar HJM & Hermens JLM 1996. Quality criteria and risk assessment for mixtures of chemicals in the aquatic environment. *Human Ecol. Risk Assess.* 2, 419−425.

Van Straalen NM & Denneman CAJ 1989. Ecotoxicological evaluation of soil quality criteria. *Ecotoxicol. Environ. Saf.* 18, 241−251.

Veith GD, Call DJ & Brooke LT 1983. Structure-toxicity relationships for the fathead minnow, *Pimephales promelas*: Narcotic industrial chemicals. *Can. J. Fish. Aquat. Sci*. 40, 743−748.

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

Vighi M & Calamari D 1996. Quality objectives for aquatic life: The problem of mixtures of chemical substances. *Human Ecol. Risk Assess.* 2**,** 412−418.

VROM (Risk Assessment and Environmental Quality Division; Directorate for Chemicals, External Safety and Radiation Protection; Ministry of Housing, Spatial Planning and the Environment) 1994. *Environmental quality objectives in The Netherlands. A review of environmental quality objectives and their policy framework in The Netherlands*. VROM, Amsterdam, The Netherlands.

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

Wagner PM, Nabholz JV & Kent RJ 1995. The new chemicals process at the Environment Protection Agency (EPA): Structure-activity relationships for hazard identification and risk assessment. *Toxicol. Letters* 79, 67−73.

Ward T & Jacoby C 1995. Deciphering spatiotemporal dynamics: what can mesocosm experiments do to improve predictions of environmental impacts? *Australasian. J. Ecotox*. 1, 51−54.

Warne MStJ 1996. The theory and practise of extrapolation techniques to derive environmental quality criteria: The Dutch approach. In *The health risk assessment and management of contaminated sites. Proceedings of the Third National Workshop on the Health risk assessment and management of contaminated sites.* eds A Langley, B Markey. & H Hill. Contaminated Sites Monograph Series 5, South Australian Health Commission, Adelaide, Australia, 403−416.

Warne MStJ, Connell DW, Hawker DW & Schüürmann G 1989a. Quantitative structure-activity relationships for the toxicity of selected shale oil components to mixed marine bacteria. *Ecotoxicol. Env. Saf.* 17, 133−148.

Warne MStJ, Connell DW, Hawker DW & Schüürmann G 1989b. Prediction of the toxicity of mixtures: Shale oil component mixtures. *Ecotoxicol. Environ. Saf.* 18, 121−128.

Warne MStJ, Connell DW, Hawker DW & Schüürmann G 1990. Development of QSARs based on high performance liquid chromatography capacity factors to describe non-specific toxicity. *Chemosphere* 19, 1113−1128.

Warne MStJ, Connell DW & 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 & Hawker DW 1995. The number of components in a mixture determines whether synergistic and antagonistic or additive toxicity predominate: The funnel hypothesis. *Ecotoxicol. Environ. Saf.* 31, 23−28.

Willis KJ 1994. The effect of temperature on the acute toxicity of Pentachlorophenol to early life stages of *Simocephalus vetulus* (Crustacea: cladocera) and rainbow trout *(Oncorhynchus mykiss)*. MSc Thesis, The University of Waikato, NZ.

Worksafe Australia & NICNAS (National Industrial Chemical Notification and Assessment Scheme) 1991. *Selection principles for priority existing chemicals.* Worksafe Australia, Sydney, Australia.

World Commission on Environment and Development 1987. *Our common future*. Australian Edition. Oxford University Press, Melbourne, Australia.

Glossary

Acute or Acute Toxicity	Having a sudden onset, lasting a short time. Of a stimulus, severe enough to induce a response rapidly. Can be used to define either the exposure or the response to an exposure (effect). Generally the duration of an acute aquatic toxicity test is 4 d or less and mortality is the response measured.
Acute-To-Chronic Ratio (ACR)	A numerical, unitless value, that is the ratio of an acute toxicity test result (ie. LC50) to a chronic toxicity test result (ie. NOEC) where both are expressed in the same units (eg. mg/L). Ideally, the data are for the same specifies and chemical. It is used for estimating the chronic toxicity of a chemical on the basis of its acute toxicity. The ACR should be greater than one.
Additive Toxicity or Additivity	The toxicity of a mixture of chemicals that is approximately equivalent to that expected from the summation of the known toxicities of the individual chemicals present in the mixture.
Aldenberg and Slob Method	A statistical extrapolation procedure that uses all the available NOEC data for a chemical and fits a logistic distribution to the data in order to derive a WQG. The method is able to protect any given percentage of species in an environment between 1% and 99%. The method is used by The Netherlands to derive the equivalent of WQGs and by the OECD to derive environmental effect concentrations.
Antagonism	A phenomenon in which the toxicity of a mixture of chemicals is less than that which would be expected from a simple summation of the toxicities (ie. additive toxicity) of the individual chemicals present in the mixture.
Arbitrary	Selected by individual will, judgement, at random or by convention. It is not selected according to any rule or law.
Assessment Factors (AF)	Arbitrarily chosen values that are designed to account for various uncertainties in extrapolating experimental data to WQGs. The magnitude of the AFs is inversely related to the perceived quality of the toxicity data. The values generally used are 10, 100 and 1000.
Assessment Factor Method	A method for deriving WQGs. This method uses assessment factors of varying magnitude to extrapolate experimental toxicity data to WQGs. The WQG for a chemical is the lowest toxicity value divided by the appropriate AF. Many countries use this method to derive WQGs. However, in all those countries that use this method apart from Canada, Australia and New

	Zealand it is only used when statistical extrapolation methods can not be used.
Arithmetic mean	The arithmetic mean is equal to the sum of the values divided by the number of values.
Bioaccumulation	This is the process by which chemicals are absorbed by aquatic organisms directly from water as well as through exposure through the consumption of food and sediment containing the chemicals.
Bioassay	'A biological assay (or bioassay) is an experiment for estimating the nature, constitution or potency of a material (or of a process), by means of the reaction that follows its application to living matter' (Finney 1978). This general definition includes both the aquatic toxicology and the pharmaceutical usage. The pharmaceutical definition of bioassay is a test used to evaluate the relative potency of a chemical or mixture of chemicals by comparing its effect on a living organism with the effect of a standard preparation on the same type of organism. Bioassays are frequently used in the pharmaceutical industry to evaluate the potency of vitamins and drugs.
Bioconcentration	The process by which there is a net accumulation of a chemical directly from the ambient environment into organisms resulting from simultaneous uptake (eg. by gill or epithelial tissue) and elimination. For aquatic organisms the ambient environment is water.
Bioconcentration Factor (BCF)	The degree to which a chemical can be concentrated in the tissues of an aquatic organism as a result of exposure to water-borne chemical. At steady state during the uptake phase of a bioconcentration test, the BCF is equal to the concentration of a chemical in one or more tissues of the exposed aquatic organisms divided by the average water concentration of the chemical in the test.
Biodegradation	The transformation of a material by the complex enzymatic action of micro-organisms (eg. bacteria, fungi). It usually leads to the disappearance of the parent structure and to the formation of smaller chemical species, some of which are used for cell anabolism. The products of the complete biodegradation of hydrocarbons are CO_2 and H_2O. However, not all compounds completely biodegrade.
Biological End Point	In toxicity testing it is the adverse biological response in question that is measured. End points vary with the level of biological organisation being examined but include changes in biochemical markers or enzyme activities, mortality or survival, growth, reproduction, primary production, and changes in structure (and

	abundance) and function in a community. End points are used in toxicity tests as criteria for effects.
Biomagnification	Result of the processes of bioconcentration and bioaccumulation by which tissue concentrations of bioaccumulated chemicals increase as the chemical passes up through two or more trophic levels. The term implies an efficient transfer of chemical from food to consumer, so that residue concentrations increase systematically from one trophic level to the next.
Chronic or Chronic Toxicity	Involving a stimulus that is lingering or continues for a long time: often signifies periods from several weeks to years, depending on the reproductive life cycle of the aquatic species. Chronic exposure typically induces a biological response of relatively slow progress and long continuance. The chronic aquatic toxicity test is used to study the effects of continuous, long-term exposure to a chemical or other potentially toxic material on aquatic organisms.
Concentration-Response Curve	A curve describing the relationship between different exposure concentrations of a chemical or material and the corresponding percentage response of the exposed test population.
Contaminant	A foreign agent that is present in the environment which may produce a physical or chemical change but may not cause adverse biological effects.
Control	This is a treatment in a toxicity test that duplicates all the conditions of the exposure treatments but contains no test material. The control is used to determine the affect of the basic test conditions (eg. health of test organisms, quality of dilution water) when no toxicant is present.
Default or generic assessment factors	These are the assessment factors most widely used in publicised assessment factor methods. These AFs differ in magnitude by one order of magnitude ie. 10, 100 and 1000.
Degree of Protection	A measure (either quantitative or qualitative) of the extent of protection offered to organisms in the environment. The degree of protection aimed to be provided by the current ANZECC WQGs is to protect all species from life time exposures to chemicals whereas the various statistical extrapolation procedures generally aim to protect 95% of species.
Direct Toxicity	Toxicity that results from and is readily attributable to the toxic agent(s) acting more or less directly at the sites of toxic action in and/or on the exposed organisms that are exhibiting the adverse biological response in question.

Effective Concentration	The concentration of material in water to which test organisms are exposed that is estimated to be effective in producing either a sublethal response in 50% of the test organisms or a 50% reduction of a sublethal characteristic. Both the length of exposure to the toxicant and the sublethal response should be clearly defined eg. 24h EC50 (immobilisation).
Effluent	A complex waste material (eg. liquid industrial discharge or sewage) that is discharged into the environment.
Empirical	A result that is obtained by experiment or observation rather than from a theory.
Fate	Disposition of a material in various environmental compartments (eg. soil or sediment, water, air, biota) as a result of transport, transformation and degradation.
Geometric mean	The geometric mean is equal to the sum of the logarithm of the values divided by the number of values.
Hardness	The concentration of all metallic cations, except those of the alkali metals, present in water. In general, hardness is a measure of the concentration of calcium and magnesium ions in water and is frequently expressed as mg/L calcium carbonate equivalent.
Harmonisation	This is a process that accounts for the inter-compartmental transport of chemicals in deriving environmental quality guidelines for water, soil, sediment and air. This ensures that at least the stated level of protection (ie. protection of 95% of species) is provided to every compartment of the environment. The Netherlands and the Danish EPA are currently evaluating this process.
Hazard	Likelihood that exposure to a chemical will cause an injury or adverse effect under the conditions of its production, use, or disposal.
Indirect toxicity	Adverse effects of toxicity that results from the agent(s) acting on and producing changes in the chemical, physical, and or biological environment external to the organisms under study. For example, a decrease in food for predatory species due to direct toxicity from a chemical to prey may produce adverse effects in the predator species due to starvation rather than inducing any direct chemical toxicity in predator organisms.
Inter-compartmental transport	The process by which pollutants irrespective of where they are released into the environment, are transported to different compartments of the environment (eg. water, sediments, soil, suspended matter, air, biota). Eventually the chemical will be present in all compartments of the

	environment.
Interim WQGs	Water quality guidelines in which there is a low degree of confidence that it will provide the desired degree of environmental protection. Higher quality WQGs include Level I and II.
Inter-species Variation	The variation that occurs between different species.
Intra-species Variation	The differences that occur between individuals or groups of individuals that belong to the same species.
Joint Action	Two or more chemicals exerting their effects simultaneously.
Keystone species	These are species that are integral to the functioning of ecosystems. If these species are eliminated or removed it leads to sever perturbations to the functioning of the ecosystem.
Kolmogorov Smirnov test	This is a statistical test that indicates whether a sample is significantly different to a particular distribution.
Lethal	Causing death by direct action. Death of aquatic organisms is the cessation of all visible signs of biological activity generally in response to a mechanical stimulus.
Lethal concentration (LC50)	The concentration of material in water to which test organisms are exposed that is estimated to be lethal to 50% of the test organisms. The LC50 should be expressed as a time-dependent value (eg. 24-h or 96-h LC50).
Level I WQG	The WQG in which there is the highest degree of confidence that it will provide the desired degree of environmental protection. Lower quality WQGs include Level II and 'interim' WQGs.
Level II WQG	The WQG in which there is a high degree of confidence that it will provide the desired degree of environmental protection. Level I WQGs are of better quality while 'interim' WQGs are of lower quality.
Level of Protection	See degree of protection.
Lowest Observed Effect Concentration (LOEC)	The lowest concentration of a material used in a toxicity test that has a statistically significant adverse effect on the exposed population of test organisms compared with the controls. It is also called lowest observed adverse effect level (LOAEL).
Maximum Acceptable Toxicant Concentration (MATC)	The hypothetical toxic threshold concentration lying in a range bounded at the lower end by the highest tested concentration having no observed effect (NOEC) and at the higher end by the lowest tested concentration having a significant toxic effect (LOEC) in a life cycle (full chronic) or partial life cycle (partial chronic) test. The MATC may be calculated as the geometric mean of

	the LOEC and NOEC.
Mechanism of action	This is the name given to the means by which the toxicant exerts its toxic effect. Generally chemicals can be subdivided into a fairly limited number of mechanisms of action. These include narcosis, polar narcosis, oxidative uncouplers, acetylcholinesterase inhibitors.
Mesocosms	see multiple species tests.
Microcosms	See multiple species tests.
Mixing Zone	An area where an effluent discharge undergoes initial dilution and is extended to cover the secondary mixing in the ambient water body. A mixing zone is an allocated impact zone where water quality criteria can be exceeded as long as acutely toxic conditions are prevented.
Modified USEPA method	This is an assessment factor method that is a modification of the original USEPA assessment factor method. It is used by The Netherlands to derive WQGs and is recommended by the OECD to derive environmental effects concentrations.
Monte Carlo simulations	This is a method of computer simulation that uses probabilistic methods to provide answers to difficult problems. These methods are used extensively in risk assessment.
Multiple species tests	These are toxicity tests in which more than one species is simultaneously exposed to the toxicant. Generally these take the form of simplistic miniature ecosystems. The size and complexity of the miniature ecosystem can vary significantly. Small simple systems are termed microcosms while larger more complex systems are termed mesocosms.
Multiple species toxicity data	Toxicity data obtained from multiple species toxicity tests.
Narcosis	This is a general, non-specific, reversible mode of toxic action that can be produced in most living organisms by the presence of sufficient amounts of many organic chemicals. Effects result from the general disruption of cellular activity. The mechanism producing this effect is unknown, with the main theories being binding to proteins in cell membranes and ‘swelling’ of the lipid portion of cell membranes resulting from the presence of organic chemicals. Hydrophobicity dominates the expression of toxicity in narcotic chemicals.
No Observed Effect Concentration (NOEC)	The highest concentration of a material in a toxicity test that has no statistically significant adverse effect on the

	exposed population of test organisms compared with the controls. This is also called the no observed adverse effect level (NOAEL) or the no observed effect level (NOEL).
Non-persistent	Chemicals that are subject to degradation and therefore do not remain in the environment for prolonged periods of time ie. days to months. While this term was used in ANZECC (1992) it was not defined. An example of a non-persistent chemical is linear alkyl benzene sulfonate (LAS).
Octanol-Water Partition Coefficient (Kow or P)	This is the ratio of a chemical's solubility in *n*-octanol and water at equilibrium. The logarithm of Kow or P (ie. log Kow or log P) is used as an indication of a chemical's propensity for bioconcentration by aquatic organisms. It is also used in quantitative structure-activity relationships (QSARs) to model the toxicity of narcotic chemicals.
Persistent	Chemicals that are resistant to degradation and therefore remain in the environment for prolonged periods of time ie. greater than one year. An example of a persistent chemical is DDT. While this term was used in ANZECC (1992) it was not defined.
Pesticide	A substance used to kill undesirable and unwanted fungi, plants, insects, or other organisms. This generic term is used to describe fungicides, algicides, herbicides, insecticides, rodenticides, nematocides and other substances.
Pollutant	This is a general term for a chemical or non-chemical (eg. increased suspended solids) agent that is present in the environment and causes adverse effects.
Precautionary Principle	This is a general principle developed by the German bureaucracy in the 1960s. It basically states that if there are threats of serious or irreversible environmental damage, lack of full scientific certainty should not be used as a reason for postponing measures to prevent environmental degradation. As such it is not a scientific principle but rather a means by which policy can deal with scientific uncertainty.

Quantitative Structure-Activity Relationships (QSARs)	Relationships between the quantified measures of the structure (and related properties) of chemicals and their activities in biological system. QSARs are used to assisted in explaining (and predicting) the occurrence and mechanism of biological responses to chemicals and to aid in prediction of incidence and magnitude. Although lethal and sublethal toxicity test results are often modelled, other biological activities, such as bioaccumulation, biodegradation, and biotransformation, may also be modelled. QSARs can also be used to explain or predict physicochemical characteristics of chemicals. Also called structure-activity relationships (SARs).
Risk	A statistical concept defined as the expected frequency or probability of undesirable effects resulting from a specified exposure to known or potential environmental concentrations of a material. A material is considered safe if the risks associated with its exposure are judged to be acceptable. Estimates of risk may be expressed in absolute or relative terms. Absolute risk is the excess risk due to actual exposure. Relative risk is the ratio of the risk in the exposed population to the risk in an unexposed population.
Secondary poisoning	See biomagnification.
Simplebox	This is a computer software program that is used to harmonise environmental quality criteria (see harmonisation). The Netherlands and the Danish EPA are evaluating its suitability.
Single species toxicity test	In this type of test only one species is exposed to the toxicant. Generally the test system is very simple having only the organism, water and the toxicant. This is the predominant type of toxicity test.
Single species toxicity data	Toxicity data obtained from single species toxicity tests.
Statistical extrapolation methods	These are methods that use all the available suitable toxicity data for a chemical and fit a statistical distribution to the data in order to derive a water quality guideline. The degree of protection can be modified from 1% to 99% of species however, generally 95% of species are protected.
Statistically Significant Effects	Effects (responses) in the exposed population that are different from those in the controls at a given statistical probability level, typically $P \leq .05$. Biological end points that are important for the survival, growth, behaviour, and perpetuation of a species are selected as criteria for effects. The end points differ depending on the type of toxicity test conducted and the species used. The statistical

	approach also changes with the type of toxicity test conducted.
Stephan et al Method	A statistical extrapolation method that uses all available NOEC data for a chemical and fits a triangular distribution to the data in order to determine WQG. This method is able to protect any given percentage of species in an environment between 1% and 99%. The method is used by the USEPA to develop WQGs.
Sublethal	Below the concentration that directly causes death. Exposure to sub-lethal concentrations of a material may produce less obvious effects on behaviour, biochemical and/or physiological function, and histology of organisms.
Synergism	A phenomenon in which the toxicity of a mixture of chemicals is greater than that which would be expected from a simple summation of the toxicities of the individual chemicals (ie. additive toxicity) present in the mixture.
Toxicant	An agent or material capable of producing an adverse response (effect) in a biological system, seriously injuring structure and/or function or producing death. Natural as well as synthetic chemicals are toxicants. In fact all chemicals, given the concentration is sufficient, will exert toxic effects and are therefore toxicants.
Toxicity Test	The means by which the toxicity of a chemical or other test material is determined. A toxicity test is used to measure the degree of response produced by exposure to a specific level of stimulus (or concentration of chemicals).
Wagner and Løkke Method	A statistical extrapolation method that uses all available NOEC for a chemical and fits a normal distribution to the data in order to derive a WQG. The method is able to protect any given percentage of species in an environment between 1% and 99%. The Danish EPA uses this method.
Water Quality Criterion	An estimate, based on scientific judgments, of the concentration of a chemical or other constituent in water which, if not exceeded, will protect an organism, an organism community, or a prescribed water used or quality with an adequate degree of safety. In some jurisdictions terminology may be different and the term guideline or objective may be substituted for criterion. Criteria may also refer to the scientific principles and data that are used to formulate recommended guidelines or limits.
Water Quality Guideline or Limit	A scientifically based numerical concentration limit or narrative statement recommended to support and maintain a designated water use.

Water Quality Objective	A scientifically based numerical concentration limit or narrative statement (eg. receiving water should sustain a healthy population of salmonids) that has been established to support and protect designated uses of water, often at a specified site or sites. Objectives reflect desired conditions in receiving waters, usually do not carry the force of law (ie. are not water quality standards), and often consider various receiving waters separately.
Whole-Effluent Toxicity (WET)	The total toxic effect of an effluent measured directly with aquatic organisms in a toxicity test.