An improved method for calculating toxicity-based pollutant loads: Part 1. Method development

Smith, RA, Warne, MSJ, Mengersen, K & Turner, RDR

An Improved Method for Calculating Toxicity-Based Pollutant Loads: Part 1. Method Development.

Running head: **Toxicity-based pollutant loads (Toxic Loads)**

Rachael A. Smith*[†‡§], Michael St.J. Warne[†‡ǁ#], Kerrie Mengersen[§], Ryan D.R. Turner[†‡§].

[†]Water Quality and Investigations, Science Division, Department of Science, Information Technology and Innovation, Ecosciences Precinct, 41 Boggo Road, Dutton Park, Queensland, 4103, Australia.

[‡]Australian Rivers Institute, Griffith University, 170 Kessels Road, Nathan, Queensland, 4111, Australia.

[§]Science and Engineering Faculty, Queensland University of Technology, 2 George Street, Brisbane, Queensland, 4001, Australia.

[ǁ]National Research Centre for Environmental Toxicology, University of Queensland, 39 Kessels Road, Coopers Plains, Queensland, 4108, Australia.

[#]Centre for Agroecology, Water and Resilience, Coventry University, Priority Street, Coventry CV1 5FB, United Kingdom.

*Address correspondence to: Rachael A. Smith, GPO Box 5078, Brisbane, Qld 4001, Australia, Phone: +61(0)7 3170 5599, Email: rachael.smith@dsiti.qld.gov.au

Other author email addresses:

Michael St.J. Warne: michael.warne@coventry.ac.uk

Kerrie Mengersen: k.mengersen@qut.edu.au

Ryan D.R. Turner: ryan.turner@dsiti.qld.gov.au

ABSTRACT

Pollutant loads are a means for assessing regulatory compliance and setting targets to reduce pollution entering receiving waterbodies. However, a pollutant load is often comprised of multiple chemicals, which may exert joint toxicity on biota. When the ultimate goal for assessing pollutant loads is to protect ecosystems from adverse effects of toxicants, then the total pollutant load needs to be calculated based on the principles of mixture toxicology. In this paper, an improved method is proposed to convert a pollutant load to a toxicity-based load (toxic load) using a modified toxic equivalency factor (TEF) derivation method. The method uses the relative potencies (RePs) of multiple species to represent the response of the ecological community. The TEF is calculated from a percentile of a cumulative distribution function (CDF) fitted to the RePs. The improvements permit the determination of which percentile of the CDF generates the most environmentally relevant and robust toxic loads. That is, environmental relevance ensures that a reduction in the toxic load is likely to result in a corresponding improvement in ecosystem health and robustness ensures that the calculation of the toxic loads is not biased by the reference chemical used. The improved methodology will therefore ensure that correct management decisions will be made and ultimately, a reduction in the toxic load will lead to a commensurate improvement in water quality.

KEY WORDS *Pollutant Loads, Mixtures, Toxic Equivalency Factor, Relative Potency, Multiple species.*

INTRODUCTION

The total mass of pollutants (loads) is often used by regulators and natural resource managers for compliance, licensing, and water pollution reduction and control programs (e.g., US EPA 2000; Raha 2007; Hardy and Koontz 2008; Australian Government and Queensland Government 2013). For example, the United States' Clean Water Act requires the development of total maximum daily load (TMDL) allocations for waterbodies to ensure water quality standards are maintained (US EPA 2000). Additionally, the Reef Water Quality Protection Plan (Reef Plan) (Australian Government and Queensland Government 2013) sets load-based reduction targets for suspended sediment, nutrients and pesticides entering the Great Barrier Reef (GBR), located off the east coast of Queensland, Australia. By meeting the load reduction targets, the 2013 Reef Plan aims to achieve its long-term goal to 'ensure that by 2020 the quality of water entering the Reef from broadscale land use has no detrimental impact on the health and resilience of the GBR' (Australian Government and Queensland Government 2013). In both these examples, there is an inherent assumption that reduction in a pollutant load will lead to improved ecosystem health. Ensuring this is true becomes complicated when the pollutant load consists of multiple chemicals which exert joint toxicity on biota.

Load based targets are an appropriate means of reducing and controlling the amount of pollutants transported to receiving water bodies, however the ecological effects of chemicals are generally controlled by their inherent toxicity, concentration in the ecosystem and the duration of exposure. A load reduction of a single chemical is likely to result in some proportional improvement in ecosystem health and therefore has a degree of environmental relevance. However, when multiple chemicals with different toxicities are present, meeting load reduction targets, that place equal weighting on each chemical, may not have the same level of environmental relevance. For example, a measured percent reduction in the load of a

low toxicity chemical will not lead to the same improvement in ecosystem health as an equivalent reduction in the load of a more toxic chemical. As such load-based pollutant reduction targets could lead to poorly targeted allocation of resources (effort and dollars) and/or perverse environmental outcomes.

Weighting the mixture constituents of a load to be more environmentally relevant can be achieved by incorporating techniques derived from mixture toxicology into the load calculations. The relative potency (ReP) and toxic equivalency factor (TEF) approaches are well-known for their use with mixtures of dioxin-like chemicals (van den Berg et al. 2006; Haws et al. 2006; US EPA 2008) and are well-suited for incorporating into pollutant load calculations (Pedersen et al. 2006). The TEF and ReP methods generate a factor for individual chemicals that can be easily applied to the load calculation, and will weight the constituents according to their relative toxicity. This permits the effects of the mixture constituents to be compared or combined on an equitoxic basis. This approach was first demonstrated by Pedersen et al. (2006) who used the RePs of organophosphorus insecticides based on a single species to weight the total daily maximum loads of mixtures of organophosphate insecticides in order to 'assess the potential ecotoxicological significance of their combined presence'. However, the ReP of one chemical to another varies between species (Putzrath 1997; Compton and Sigal 1999) and, in addition, there is an implicit uncertainty in extrapolating a TEF from one species to the response of a whole community (De Zwart and Posthuma 2005). Therefore, in order for pollutant loads to have a better alignment with potential ecological effects, the RePs of multiple species should be used to calculate the 'toxic load'.

A probabilistic approach using a statistical distribution of RePs from a representative group of species has been suggested (Finley et al. 2003; Haws et al. 2006) to calculate RePs and TEFs. While this approach will account for multiple species, in many cases a single value

from a distribution is preferred or required (van den Berg et al. 2006). In the case of toxic load calculations, a single value is preferable, particularly for ease of calculation and communication for compliance and licensing. Deriving a single value from a representative distribution of species' ReP values is comparable to deriving environmental quality guidelines from species sensitivity distributions (SSDs). A SSD is a cumulative distribution function (CDF) which describes the variation in the sensitivities of a sample of species that occur in an ecosystem to a toxicant or mixture of toxicants. One of the assumptions of SSD methods is that the sensitivity of the sample of species is representative of the assessed ecological community (Posthuma et al. 2002). A SSD determines the concentration of a chemical that should theoretically protect p % of species (termed either the hazardous concentration to the selected percent of species e.g. HC5 or the protective concentration e.g. PC95). Defining the context of a selected percentile of a ReP CDF is more complex and does not necessarily relate to environmental protection. Specifically, a selected percentile (p) of the ReP CDF represents p% of species for which the relative potency of the test chemical is up to x times more toxic than the reference chemical (assuming the reference chemical is less toxic than the test chemical). Thus, there is still a question remaining as to which percentile of the ReP CDF should be used to derive a TEF that is environmentally relevant.

In the case of dioxin-like chemicals the World Health Organisation used a combination of unweighted ReP distributions, expert judgment, and point estimates to derive the TEFs (van den Berg et al. 2006). As a result, the TEFs derived from a range of percentiles that principally fell within the 50th and 75th percentiles of the ReP distribution, with the majority being closer to the 75th percentile in order to be "health protective" (van den Berg et al. 2006). Environment Canada and Health Canada (2001) calculated TEFs for nonylphenol and its ethoxylates by taking the mean of the ReP values. Similarly, Kennedy et al. (2010) calculated diuron TEFs for other herbicides by averaging the RePs of coral and microalgae.

In all of these cases however, there was no indication as to whether the percentile used to calculate the TEFs provided an appropriate degree of protection (i.e. environmental relevance). Furthermore, there was no test to demonstrate that the results weren't biased by the chosen reference chemical and the same environmental outcome would be achieved if the reference chemical was changed, i.e. environmental robustness.

This paper proposes a method that converts annual pollutant loads to toxic loads using a modified TEF approach that includes tests to maximise the environmental relevance and robustness of the TEF values. Another paper (Smith et al. submitted to IEAM) tests the applicability of the new method to a case study – pollutants discharged from agricultural land to the Great Barrier Reef.

Definitions of key terms

The definitions of key terms used in this paper are provided below.

Cumulative Distribution Function (CDF) - describes the probability that a variable will be equal to or less than a specified value. In the case of SSDs, the CDF curve describes the distribution of ecotoxicity data to a chemical in which species are ranked from the most to the least sensitive (Posthuma et al. 2002). From a CDF, we can therefore determine the percentage of species that would theoretically experience adverse effects by any specified concentration of a chemical.

Load (L) - the estimated mass (e.g. t, kg) of a chemical that passes a specified point in a waterway. Loads are usually calculated on a daily, event or annual basis. The load of an individual chemical (i) is referred to in this paper as L_i. The load of a mixture of chemicals (L_{mix}) is equal to the sum of the loads of each chemical in the mixture.

Matched toxicity data - toxicity data from studies conducted within the same laboratory where multiple chemicals are tested under the same test conditions to a consistent set of organisms.

Relative Potency (ReP) - is the estimate of the potency of the chemical being considered (henceforth referred to as the 'test' chemical) relative to a reference chemical, that both cause a specified toxic effect in a population, organism, cell or biochemical reaction (US EPA 2008). The test and reference chemicals must have the same mode of action (MoA), have parallel concentration-response curves, and conform with the concentration addition (CA) model of joint action (Safe 1998). The toxicity data of the test and reference chemicals used to calculate a ReP must be derived from the same species and study design, preferably using 'matched toxicity data'. For one test chemical, there may be a range of ReP values calculated from multiple species, and multiple ReP values for one species based on different ecotoxicity endpoints.

Toxic equivalence factor (TEF) – the estimate (based on one or more studies) of the potency of a test chemical relative to a reference chemical that causes a toxic effect. In this paper TEFs are determined using a cumulative distribution function of the ReP values of multiple species.

Toxic Load (TL) - the product of the TEF for each chemical in a mixture multiplied by the load of each individual chemical. The toxic load of chemicals with the same MoA can then be summed to generate a toxic load for the mixture (TL_{mix}). The TL is expressed as an equivalent mass of the reference chemical, e.g. diuron equivalents (Eq).

General method for calculating toxic loads

The steps of the proposed general method for calculating TLs are presented in Figure 1. The name 'general method' was used because it outlines the 'general' process, however

modifications may be required to optimise the method for particular jurisdictions or for specific sites (see Smith et al. submitted to IEAM, for an example of this). The general method is largely adapted from published procedures used for calculating RePs and TEFs (Safe 1998; US EPA 2008), and generating SSDs for water quality guideline (WQG) derivation (e.g. Warne 2001; Warne et al. 2015).

The principle variation to these earlier methods for determining the TEF from a cumulative distribution function (CDF) is the inclusion of two novel procedures to ensure the most environmentally relevant and robust TLs are generated. The two new procedures are a test for environmental relevance and a test for robustness (Step 6, Figure 1).

Nominating the reference chemical

Calculating RePs requires a reference chemical to which the potency of the test chemicals being considered can be compared. Published methods for calculating RePs of particular classes of chemicals have a suggested standard reference chemical. For example, 2,3,7,8-tetrachlorodibenzo-p-dioxin (2,3,7,8 TCDD) or the polychlorinated biphenyl congener 126 (PCB 126) are the standard reference chemicals for dioxin-like chemicals (van den Berg 2006; US EPA 2008), and estradiol is the traditional reference for endocrine disruptor chemicals with estrogen mediated receptor responses (e.g. Gutendorf and Westendorf 2001). The Ontario Ministry of the Environment (OME) was the first to propose the TEF methodology (OME 1984; Haws et al. 2006) and suggested the most toxic and well-studied member (2,3,7,8 TCDD) of the dioxin-like class of chemicals should be used as the reference chemical. However, the 'most toxic' chemical may not be easy to define for many chemical classes, due to inter-species variations in toxicity. Even in the case of the dioxin-like chemicals, more recent research has demonstrated that 2,3,7,8-TCDD is not the most

toxic chemical to all species, with other dioxin-like chemicals having potencies up to four times that of 2,3,7,8-TCDD (Finley et al. 2003). Furthermore, a fixed reference chemical may not be suitable in all circumstances. For example, it may be important in communicating results to use a reference chemical that is one of the mixture constituents present at a site. Ultimately, we consider that the most important requirement in nominating a reference chemical is that there are sufficient matched data sets between the reference chemical and each of the mixture constituents to generate reliable ReP CDFs. A reliable CDF is one that is based on toxicity data that meets the minimum data requirements to generate a CDF (for details see Step 5). The best reference chemical is therefore the chemical that can make reliable ReP CDFs with the most chemicals in the suite of chemicals being considered.

Collating and screening toxicity data

Toxicity data required for calculating the RePs of multiple species can be sourced from the scientific literature. The suggested methods for collating and screening suitable toxicity data for calculating RePs are as follows:

- Toxicity data should be collected from published peer-reviewed studies, published laboratory reports and/or ecotoxicity databases – un-published data should only be used as a last resort and if a copy of a document stating the method used is made publicly available (Warne et al. 2015);

- Preference should be given to using matched toxicity data sets (see earlier definition). This overcomes differences in the data due to variable test conditions and intra-laboratory variation. Where there are insufficient matched data available, data from separate studies can be used providing key test conditions were identical or similar (e.g. identical species, test conditions, measures of toxicity, endpoints).

- If the MoA of a chemical is specific to a group of organisms, only species belonging to the target group should be used (Warne et al. 2015). For example, photosystem II (PSII) herbicides are far more toxic to photosynthetic organisms and therefore only phototrophic species should be used to calculate RePs.

- Preference should be given to using median lethal (LC50) or effect (EC50) concentration data as these are the points in the concentration-response curve with the least error. This is more important than using a more sensitive measure of toxicity such as the EC10 or NOEC type data that are mainly preferred in deriving WQGs (ANZECC and ARMCANZ 2000; EC 2011; Warne et al. 2015).

- Preference should be given to toxicity data with exposure durations that are relevant to the MoA of the chemical and test organism.

- Preference should be given to using toxicity data that measure ecologically relevant endpoints – those that affect the ecological competitiveness of a species (e.g. lethality, immobilisation, growth, development, population growth, and reproduction – ANZECC and ARMCANZ 2000; Warne et al. 2015) or are relevant to the MoA of the chemical.

Determining the quality of toxicity data

All toxicity data that pass the screening process should be assessed by a data quality checking scheme similar to what is used for generating water quality guidelines (e.g. Klimisch et al. 1997; Durda and Preziosi 2000; Hobbs et al. 2005; Schneider et al. 2009; Agerstrand et al. 2014 ; Isigonis et al. 2015; Warne et al. 2015). Only data of sufficient quality should be used to derive RePs and TEFs.

Calculating relative potencies (RePs)

The ReP is calculated according to Equation 1.

$$REP_i = \frac{ECx_r}{ECx_i} \qquad \qquad \text{Equation 1}$$

where, ECx_i is the concentration of chemical 'i' that effects x% of a population of the test

organism, and ECx_r is the concentration of the reference chemical that effects x% of a

population of the test organism for the same endpoint. We recommend using the midpoint of

the concentration-response curve (i.e. the EC/LC50) as the least amount of error generally

occurs at this point in the curve.

A ReP of 1 indicates that the test and reference chemicals are equally toxic, a ReP

value < 1 indicates that the test chemical is less toxic than the reference chemical, and a ReP

value > 1 indicates that the test chemical is more toxic than the reference chemical.

Fitting a cumulative distribution function to ReP values

This step adopts the approach of Finley et al. (2003) and Haws et al. (2006) by fitting

the ReP data to a CDF, analogous to the SSD concept. If multiple toxicity tests have been

conducted or multiple experimental conditions were employed (e.g. multiple endpoints

measured, different test durations) then species may have multiple ReP values. A single ReP

value is required to represent each species in the CDF. Data reduction methods such as those

used to derive one toxicity value per species in SSDs (e.g. van de Plassche et al. 1993; Warne

et al. 2015) should be used in calculating ReP values. It is recommended that at least the

minimum data requirements for generating a SSD are used when fitting ReP values to a CDF.

Note however, that these minimum data requirements vary with the jurisdiction. For example,

the minimum data requirement for deriving water quality guidelines in Australia and New

Zealand is toxicity data for at least five species from at least four phyla (ANZECC and

ARMCANZ 2000), in the USA 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 are required (Stephan et al. 1985) and in Europe data for at least ten species that belong to at least eight taxa are required (EC 2011).

The single ReP values for each species should be collated and analysed using a CDF method (e.g. BurrliOZ V2 (Barry and Henderson 2014) and SSD Generator (US EPA 2012)). This approach, will permit the calculation of TEF values from the CDF for different percentiles of the ReP CDF.

Selecting the percentile of the ReP cumulative distribution function to calculate the toxic equivalency factors

It is important to realise that ReP CDFs can be located to the right of REP equals one (i.e. the test chemical is more toxic than the reference chemical), to the left of one (i.e. the test chemical is less toxic than the reference chemical), or either side of one (i.e. for some species the test chemical is more toxic than the reference chemical and for other species it is equal to or less toxic than the reference chemical). Thus, to determine the TEF that is 'representative' of a defined percentile of species, we need to adjust our calculations to account for the position of the species' ReP values relative to the reference chemical (i.e. REP = 1). To illustrate the implications of the above, the CDFs of the ReP values of a chemical that is less toxic (Chemical A) and more toxic (Chemical B) than a reference chemical are presented in Figure 2. The distribution of ReP values for chemical A sit principally to the left of the reference chemical (the logarithm of the ReP value for the reference toxicant is 0 i.e. $\log_{10} 1 = 0$) while the distribution of ReP values for chemical B sit principally to the right of the reference chemical. Thus, the TEF that represents the ReP values of chemical B (relative to the reference chemical) for $\leq 95\%$ of species, can be calculated using the 95th percentile of the distribution (shaded area B). However, the TEF that represents the ReP values of chemical A (relative to the reference chemical) for $\leq 95\%$ of

species would be calculated using the 5th percentile of the distribution (shaded area A) (as this chemical is generally less toxic than the reference chemical). Table 1 provides examples of percentiles that could be used to calculate TEF values depending on the toxicity of the test chemical relative to the reference chemical.

The next step examines which percentile of the ReP CDF should be selected to calculate the TEFs and subsequently the TLs. The percentile that is selected determines the value of the TEF which in turn determines the contribution of each constituent to the TLmix, and therefore, will determine which chemicals become the focus of management action. For example, in Smith et al. (submitted to IEAM) it was demonstrated that, depending on which percentile was selected, the relative contributions of the mixture constituents to the TLmix varied by up to 40%. Such a large variation in results could lead to different management actions and outcomes depending on which percentile was used. Therefore, a percentile needs to be selected that will generate environmentally relevant and robust TLs.

We recommend an iterative approach (Figure 3) in which a percentile is selected to calculate a TEF, the TLs are then calculated from the TEFs, and lastly the TLs are tested for environmental relevance and robustness. This process is repeated for different percentiles until the percentile that generates the optimal (i.e. the most environmentally relevant and robust) set of TEFs and TLs is determined.

Testing toxic equivalency factors and toxic loads for environmental relevance and robustness

Calculating toxic loads

The TL$_{mix}$ is calculated by first generating the TLs of each of the mixture constituents using the following equation:

$$TL_{i,p} = TEF_{i,p} \times L_i \qquad\qquad \text{Equation 2}$$

where $TL_{i,p}$ is the toxic load of chemical i for percentile *p* of the ReP CDF of chemical i, $TEF_{i,p}$ is the TEF corresponding to the pth percentile of the ReP CDF of chemical i, and L_i = the load of chemical i (kg or tonnes).

The TLs for each constituent are then summed using the CA model of joint action to calculate TL_{mix}, i.e.:

$$TL_{mix,p} \; = \; \sum_i TL_{i,p} \hspace{3cm} \text{Equation 3}$$

where, $TL_{mix,p}$ is the toxic load of the mixture for the pth percentile of the ReP CDF, and $TL_{i,p}$ is the toxic load of chemical i for the pth percentile of the ReP CDF.

The ratio of the TL of each mixture constituent to the TL_{mix} can then be calculated, i.e.:

$$TL_{i,p} : TL_{mix,p}. \hspace{3cm} \text{Equation 4}$$

Test for environmental relevance

The test for environmental relevance compares $TL_{i,p} : TL_{mix,p}$ values generated from a selected percentile, against a similar ratio calculated from an independent method for estimating mixture toxicity. The multisubstance-potentially affected fraction (ms-PAF) method (see Traas et al. 2002 for a detailed description) is the recommended independent method.

The TEF method outlined in this paper and the ms-PAF method are similar in that both are probabilistic techniques that use a sample of the population to generate a CDF (SSD in the case of the ms-PAF). The ms-PAF method differs from the TEF method by estimating the percent of species that would be affected by the mixture in question. An advantage of using ms-PAF as an independent method is that the toxicity data required for producing SSDs do not have to be matched, as with RePs. This means that larger toxicity datasets with a better

representation of species and phyla are often available to use with the ms-PAF method, and therefore, a more reliable estimate of an ecosystem response is produced. Unlike the TEF method in which there is a preference for EC/LC50 toxicity values, the No Observed Effect Concentration (NOEC), No Effect Concentration (NEC), and/or EC/LC10 values are often preferred for generating SSDs particularly for environmental conservation (e.g. Warne et al. 2015). For this reason, we recommend the use of NOEC/NEC/EC/LC10 toxicity values for ms-PAF calculations.

The ms-PAF method (for chemicals with the same MoA) is normally a two-stage process, however in this study we only need to conduct the first stage; calculating the hazard units (HUs). The HUs of each mixture constituent are calculated from their individual SSDs, according to Traas et al. (2002) (Equation 5);

$$HU_i = \frac{C_i}{\tilde{X}_i^j} \qquad \text{Equation 5}$$

where HU_i is the hazard unit for chemical i, C_i is the concentration of chemical i in a sample, and $\tilde{X}_i^j$ is the median EC/LCx (e.g. EC10) of species j to m exposed to chemical i.

The HU values of the mixture constituents are then summed resulting in a hazard unit for the mixture (HU_{mix}) (Equation 6).

$$HU_{mix} = \sum_i HU_i \qquad \text{Equation 6}$$

Using the above equations the contribution of each constituent to the HU_{mix} can be determined ($HU_i:HU_{mix}$) and compared to the corresponding contribution calculated using the TL method (i.e. $TL_{i,p}:TL_{mix,p}$). This should be done for each of the selected percentiles. The percentile which generates $TL_i:TL_{mix}$ ratios most similar to the $HU_i:HU_{mix}$ ratios for all constituents would be considered the most environmentally relevant.

Test for robustness

It is important to select the percentile of the ReP CDF which, for any mixture constituent, generates equal contributions to the TL_{mix} irrespective of the reference chemical used. By doing this, it means that changing the reference chemical will not change the contribution of each constituent and the overall assessment of the risk the mixture poses. This would be advantageous if, for example, more toxicity data became available for one of the other mixture constituents, the constituents of the mixture changes, or the usual reference chemical is in the process of being phased out or its use restricted. In any case, the contributions of the mixture constituents to the TL_{mix} should remain the same and not depend on the reference chemical chosen. Therefore, the percentile which generates a TL_i:TL_{mix} ratio most similar amongst multiple reference chemicals would be considered the most robust.

Adopting the TEFs that generate the most relevant and robust toxic loads

Depending on the case being examined the percentile could be selected to optimise either the environmental relevance or the robustness or both. We recommend the latter. In the case study presented in Smith et al. (submitted to IEAM), environmental relevance decreased with increasing percentile of the ReP CDF while robustness increased with increasing percentile of the ReP CDF. This meant that no percentile existed that scored the highest in both tests. Therefore, the percentile was selected which had the highest possible scores from both tests by fitting regression models to the scores of each test and identifying where the two regression lines met. The two models intersected close to the 75th percentile, therefore this value was chosen as the percentile to generate the TEFs and hence calculate the TLs. Coincidently, the 75th percentile was the preferred percentile by WHO when generating TEFs for dioxin (van den Berg et al. 2006). Until, more case studies similar to Smith et al.

(submitted to IEAM) have been conducted it will not be possible to determine if the 75th percentile is the best to use universally.

CONCLUSIONS

Calculating toxic loads makes ecotoxicological sense when dealing with mixtures of toxicants and will inform water quality managers on which chemicals and regions they should focus their actions, leading to better allocation of resources. In this study we developed an environmentally relevant and robust method for calculating toxic loads using a modified TEF approach. This paper proposes that converting loads of a chemical mixture to toxic loads was more environmentally relevant than placing equal weighting on each of the mixture constituents. However, care must be taken in selecting which percentile of the ReP CDF is used to calculate the toxic loads, as the percentile chosen can have a marked effect on the environmental relevance and robustness of the toxic load, the relative magnitude of the toxic load, and the relative contribution of each constituent to the toxic load. Hence, using the wrong percentile could lead to misguided management decisions, e.g. reducing a toxic load by more than what is needed, which could have unnecessary economic costs, or not reducing a toxic load enough which could lead to environmental degradation. A systematic approach that determines the best percentile, such as that developed in this study, is necessary.

The toxic load method is not without limitations. The requirement for matched ecotoxicity data sets means that this method may not be suitable for some chemicals, particularly newer ones where ecotoxicity data are limited. In addition, the requirement for chemicals to have the same MoA and parallel concentration-response curves means that this method would not be applicable for some pollutant mixtures.

ACKNOWLEDGEMENTS

REFERENCES

Ågerstrand, M., Edvardsson, L., & Rudén, C. (2014). Bad reporting or bad science? Systematic data evaluation as a means to improve the use of peer-reviewed studies in risk assessments of chemicals. *Human and Ecological Risk Assessment: An International Journal*, 20(6), 1427-1445.

Australian Government and Queensland Government. 2013. Reef Water Quality Protection Plan 2013. Securing the health and resilience of Great Barrier Reef World Heritage Area and adjacent catchments. Brisbane (QLD), Australia: Reef Water Quality Protection Plan Secretariat, the State of Queensland.

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

Barry, S and Henderson B. (2014). Burrlioz 2.0. Canberra, Australia: Commonwealth Science and Industrial Research Organisation. Available from: https://research.csiro.au/software/burrlioz/ . Accessed December 24, 2014.

Compton R, Sigal EA. 1999. The use of toxic equivalency factors (TEFs) in ecological risk assessment: Strengths and limitations. *Hum Ecol Risk Assess,* 5: 33-42.

De Zwart D, Posthuma L. 2005. Complex mixture toxicity for single and multiple species: Proposed methodologies. *Environ Toxicol Chem* 24: 2665-2676.

Durda JL, Preziosi DV. (2000). Data quality evaluation of toxicological studies used to derive ecotoxicological benchmarks. *Hum Ecol Risk Assess* 6: 747-765.

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

Environment Canada and Health Canada. 2001. Priority substances list assessment report: Nonylphenol and its ethoxylates. Canada: Environment Canada and Health Canada, Minister of Public Works and Government Services.

Finley B, Connor K, Scott P. (2003). The use of toxic equivalency factor distributions in probabilistic risk assessments for dioxins, furans, and PCBs. *J Toxicol Environ Health Part A* 66, 533-550.

Gutendorf B, Westendorf J. 2001. Comparison of an array of in vitro assays for the assessment of the estrogenic potential of natural and synthetic estrogens, phytoestrogens and xenoestrogens. *Toxicol* 166, 79-89.

Hardy SD, Koontz TM. 2008. Reducing nonpoint source pollution through collaboration: policies and programs across the US states. *Environ Manage* 4: 301-310.

Haws LC, Su SH, Harris M, DeVito MJ, Walker NJ, Farland WH, Finley B, Birnbaum LS. (2006). Development of a refined database of mammalian relative potency estimates for dioxin-like compounds. *Toxicol Sci* 89: 4-30.

Hobbs DA, Warne MSJ, Markich SJ. (2005). Evaluation of criteria used to assess the quality of aquatic toxicity data. *Integr Environ Assess Manage* 1: 174-180.

Klimisch HJ, Andreae M, Tillmann U. (1997). A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data. *Regul Toxicol Pharmacol* 25: 1-5.

Kennedy K, Paxman C, Dunn A, O'Brien J, Mueller J F. 2010. Monitoring of organic chemicals in the Great Barrier Reef Marine Park and selected tributaries using time integrated monitoring tools (2008-2009). Brisbane (QLD), Australia: National Research Centre for Environmental Toxicology University of Queensland (Entox), University of Queensland.

OME (Ontario Ministry of the Environment) (1984). Scientific criteria document for standard development. No. 4–84. Polychlorinated Dibenzop- dioxins (PCDDs) and polychlorinated dibenzofurans (PCDFs). Ontario, Canada: Intergovernmental Relations and Hazardous Contaminants Coordination Branch, Ontario Ministry of the Environment.

Pedersen JA, Yeager MA, Suffet IH (Mel). 2006. Organophosphorus insecticides in agricultural and residential runoff: Field observations and implications for total maximum daily load development. *Environ Sci and Technol* 40:2120-2127.

Posthuma, L., Traas, T. P., & Suter, I. I. (2002). General introduction to species sensitivity distributions. In: Posthuma L, Suter II GW, Traas TP, editors. Species Sensitivity Distributions in Ecotoxicology. Boca Raton (FL), USA: Lewis Publishers. p 3-11.

Putzrath RM. 1997. Estimating relative potency for receptor-mediated toxicity: reevaluating the toxicity equivalence factor (TEF) model. *Regul Toxicol Pharmacol* 25: 68–78.

Raha D. 2007. Paradigm Shift in Water Environment Protection in New South Wales. *Aust J Water Resour* 11: 67-78.

Safe SH. 1998. Hazard and risk assessment of chemical mixtures using the toxic equivalency factor approach. *Environ Health Perspect*, 106(Suppl 4): 1051-1058.

Schneider K, Schwarz M, Burkholder I, Kopp-Schneider A, Edler L, Kinsner-Ovaskainen A, Hartung T, Hoffmann S. 2009. "ToxRTool", a new tool to assess the reliability of toxicological data. *Toxicol Lett* 189: 138-144.

Smith RA, Warne MStJ, Mengersen K and Turner RDR. Submitted to IEAM. Application of Toxicity-Based Pollutant Loads (Toxic Loads) to Contaminants Discharged to the Great Barrier Reef, Queensland, Australia. Submitted to *Integ. Environ. Manag, Assess*.

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. Washington DC, USA:US EPA. US EPA Report No. PB-85-227049.

Traas TP, Van de Meent D, Posthuma L, Hamers T, Kater BJ, De Zwart, D, Aldenberg, T. (2002). The potentially affected fraction as a measure of ecological risk. In: Posthuma L, Suter II GW, Traas TP, editors. Species Sensitivity Distributions in Ecotoxicology. Boca Raton (FL), USA: Lewis Publishers. p 315-344.

US EPA [United States Environmental Protection Agency]. 2000. Revisions to the Water Quality Planning and Management Regulation and Revisions to the National Pollutant Discharge Elimination System Program in Support of Revisions to the Water Quality Planning and Management Regulation. Washington DC, USA: US EPA. 65 Federal Register, 43, 586.

US EPA [United States Environmental Protection Agency]. 2008. Framework for Application of the Toxicity Equivalence Methodology for Polychlorinated Dioxins, Furans,

and Biphenyls in Ecological Risk Assessment. Washington DC, USA: Office of the Science Advisor, US EPA. US EPA 20460.

US EPA [United States Environmental Protection Agency]. 2012. *Species Sensitivity Distribution Generator, Ver 1. Cincinnati (OH), USA:* US EPA, Office of Research and Development, National Center for Environmental Assessment. *Available from:* http://www.epa.gov/caddis/da_software_ssdmacro.html

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

van den Berg M, Birnbaum LS, Denison M, De Vito M, Farland W, Feeley M, Fiedler H, Hakansson H, Hanberg A, Haws L, Rose M, Safe S, Schrenk D, Tohyama C, Tritscher A, Tuomisto J, Tysklind M, Walker N, Peterson RE. (2006). The 2005 World Health Organization reevaluation of human and mammalian toxic equivalency factors for dioxins and dioxin-like compounds. *Toxicol Sci* 93, 223-241.

Warne MStJ. 2001. Derivation of the ANZECC and ARMCANZ Water Quality Guidelines for Toxicants. *Australas J Ecotoxicol* 7, 123 – 136.

Warne MStJ, Batley GE, van Dam RA, Chapman JC, Fox DR, Hickey CW, Stauber JL. 2015. Deriving Australian and New Zealand Water Quality Guideline Values for Toxicants. Brisbane (QLD), Australia: Department of Science, Information Technology and Innovation.

FIGURES

Figure 1. Flow diagram of steps for calculating toxic loads using a modified TEF approach.

Figure 2. Probability density function of log ReP values. The distribution of ReP values for chemical A sit principally to the left of the reference chemical (i.e. log ReP = 0) while the distribution of ReP values for chemical B sit principally to the right of the reference chemical. The shaded area represents the ReP values for 95% of species for chemical A and B and are calculated using the 5th and 95th percentiles, respectively.

Figure 3. Key steps within steps six and seven of the Toxic Loads general method.

Table 1 Example of corresponding percentiles of the relative potency cumulative distribution function (ReP CDF) for chemicals less toxic and more toxic than the reference chemical.

Percentiles of the ReP CDF	
Less toxic than the reference chemical	More toxic than the reference chemical
5	95
10	90
20	80
50	50
80	20
90	10
95	5

1

```
┌─────────────────────────────────────────────┐
│   1. Nominating the reference chemical        │
└─────────────────────────────────────────────┘
                      ↓
┌─────────────────────────────────────────────┐
│   2. Collating and screening toxicity data    │
└─────────────────────────────────────────────┘
                      ↓
┌─────────────────────────────────────────────┐
│   3. Determining the quality of toxicity data │
└─────────────────────────────────────────────┘
                      ↓
┌─────────────────────────────────────────────┐
│   4. Calculating relative potencies (ReP)     │
└─────────────────────────────────────────────┘
                      ↓
┌─────────────────────────────────────────────┐
│   5. Fitting a cumulative distribution        │
│   function to ReP values                       │
└─────────────────────────────────────────────┘
                      ↓
┌─────────────────────────────────────────────┐
│   6. Selecting the percentile of the ReP       │
│   cumulative distribution function to          │
│   calculate toxic equivalency factors          │
└─────────────────────────────────────────────┘
                      ↓
┌─────────────────────────────────────────────┐
│   7. Testing toxic equivalency factors and     │
│   toxic loads for environmental relevance &    │
│   robustness                                    │
└─────────────────────────────────────────────┘
                      ↓
┌─────────────────────────────────────────────┐
│   8. Adopting the toxic equivalency factors    │
│   that generate the most relvant and robust    │
│   toxic loads                                   │
└─────────────────────────────────────────────┘
```

2

3

Select percentile of the CDF to be tested

Calculate TEFs for selected percentile

Calculate TLs for representative subset of results

Test environmental relevance and robustness

If NOT sufficiently relevant and robust

If sufficiently relevant and robust

Use selected TEF to calculate TLs for all results