

An improved method for calculating toxicity-based pollutant loads: Part 2. Application to contaminants discharged to the Great Barrier Reef, Queensland, Australia

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

Author post-print (accepted) deposited by Coventry University's Repository

Original citation & hyperlink:

Smith, RA, Warne, MSJ, Mengersen, K & Turner, RDR 2017, 'An improved method for calculating toxicity-based pollutant loads: Part 2. Application to contaminants discharged to the Great Barrier Reef, Queensland, Australia' *Integrated Environmental Assessment and Management*, vol 13, no. 4, pp. 754-764
<https://dx.doi.org/10.1002/ieam.1860>

DOI 10.1002/ieam.1860

ISSN 1551-3777

ESSN 1551-3793

Publisher: Wiley

This is the peer reviewed version of the following article: Smith, RA, Warne, MSJ, Mengersen, K & Turner, RDR 2017, 'An improved method for calculating toxicity-based pollutant loads: Part 2. Application to contaminants discharged to the Great Barrier Reef, Queensland, Australia' *Integrated Environmental Assessment and Management*, vol 13, no. 4, pp. 754-764, which has been published in final form at <https://dx.doi.org/10.1002/ieam.1860>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

Copyright © and Moral Rights are retained by the author(s) and/ or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This item cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder(s). The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

This document is the author's post-print version, incorporating any revisions agreed during the peer-review process. Some differences between the published version and this version may remain and you are advised to consult the published version if you wish to cite from it.

An Improved Method for Calculating Toxicity-Based Pollutant Loads: Part 2. Application to Contaminants Discharged to the Great Barrier Reef, Queensland, Australia.

Running head: **Toxic Loads in the Great Barrier Reef**

Rachael A. Smith^{*†‡§}, Michael St.J. Warne^{†‡¶#}, 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.

[¶]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 widely used to set pollution reduction targets and assess regulatory compliance for the protection of receiving waterbodies. However, when a pollutant load is comprised of a mixture of chemicals, reducing the overall load (mass) will not necessarily reduce the toxicity by a similar amount. This can be overcome by setting targets based on toxicity-based loads (toxic loads), where the load is modified according to the relative toxicity (expressed as toxic equivalency factors—TEFs) of each toxicant. Here we present the second paper of a two-part series in which a case study is used to demonstrate the application of the toxic load method proposed in Part 1. The toxic load method converts a pollutant load, comprised of multiple chemicals, to a toxicity-based load (toxic load), using a modified TEF approach. The modified approach is based upon the cumulative distribution of relative potency (ReP) estimates of multiple species, and is further improved on previously published TEF methods with the inclusion of two tests to select the percentile of the cumulative ReP distribution which generate TLs: that align with an independent mixture method (test for environmental relevance); and are independent of the reference chemical used (test for robustness). Here, the TL method is applied to mixtures of pesticides that are discharged from agricultural land to the Great Barrier Reef (GBR) in order to test its utility. In this case study, the most environmentally relevant and robust TLs were generated using the 75th percentile of the ReP cumulative distribution. The results demonstrate that it is essential to develop pollution reduction targets based on toxic loads, and making progress to meeting them will lead to a commensurate reduction in toxic effects caused by toxicants in waters of the GBR.

KEY WORDS *POLLUTANT LOADS, GREAT BARRIER REEF, PHOTOSYSTEM II*

HERBICIDES, TOXIC EQUIVALENCY FACTOR, RELATIVE POTENCY

INTRODUCTION

Continued concerns about the long-term health and sustainability of the Great Barrier Reef (GBR) (Haynes and Michalek-Wagner, 2000; Brodie et al 20008; Brodie et al. 2013) led to the Australian and Queensland governments jointly establishing the Reef Water Quality Protection Plan (Reef Plan) in 2003 and revising it in 2009 and 2013 (Australian Government and Queensland Government 2003, 2009, 2013). A key element of Reef Plan was setting aspirational pollutant load reduction targets and agricultural land management practice targets to improve the quality of water entering the GBR from adjacent agricultural land. The pollution load reduction targets included ‘at least a 60 per cent reduction in end-of-catchment pesticide loads in priority areas’ by 2018 (Australian Government and Queensland Government 2013). The long-term goal of these load reduction targets was 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). Fundamental to achieving this long-term goal is the assumption that a reduction in a pollutant load will lead to a commensurate improvement in ecosystem health. However, when the pollutant load consists of a mixture of chemicals with different toxicities, meeting the load reduction target may not achieve the long-term goal of Reef Plan (2013).

Load based targets are often used by regulators and natural resource managers as a means for quantifying and controlling pollutant inputs under compliance, licensing, and water pollution reduction and control programs (US EPA 2000; Raha 2007; Hardy and Koontz 2008; Australian Government and Queensland Government 2013). However, it is well known that the ecological impact of a toxicant is determined by the concentration of the chemical in the environment, the duration and spatial extent of exposure and the sensitivity of the exposed biota (US EPA 1998). A pollutant load on the other hand, is the total mass of a pollutant over a given time period (e.g. tonnes/year).

Two chemicals present together in a catchment with equal concentrations would have equal annual loads; in terms of meeting a load-based pollutant reduction target, they would therefore be interpreted as being equally important. However, if the two chemicals have different toxicities then their ecological impact would be different. Reducing the load of a chemical with a low toxicity will not lead to the same improvement in ecosystem health as an equivalent load reduction of a chemical with a high toxicity (Smith et al. submitted to IEAM). In general, we can assume that ecosystem health would improve with a reduction in the load of chemicals, however, the improvement would not necessarily be of a similar magnitude as the decrease in the load (Smith et al. submitted to IEAM).

In the case of the GBR, a total pesticide load is currently quantified by simply summing the loads of five priority photosystem II (PSII) herbicides which co-occur in Queensland catchments that discharge to the GBR (Waters et al. 2014). Thus, to meet the Reef Plan pollutant reduction target for pesticides, a 60% reduction in total pesticide load is required (Australian Government and Queensland Government 2013), irrespective of the relative amounts of the five herbicides that are reduced (i.e. Reef Plan focusses on reducing the loads of ametryn, atrazine, diuron, hexazinone and tebuthiuron). To achieve this target, it would be logical to focus management efforts on the pesticides with the largest loads. However, if these pesticides had relatively low toxicities compared to the other pesticides present, the toxicity to the GBR would not be reduced by the same amount as the reduction in the total pesticide load. Thus, for managers to ensure that reductions in the pollutant load of a mixture will result in a proportional reduction in toxicity or adverse environmental effects, the relative toxicities of mixture constituents need to be considered.

The Toxic Load (TL) method, presented in Smith et al. (submitted to IEAM), uses a modified toxic equivalency factor (TEF) approach similar to that which has been used for dioxin-like chemicals (van den Berg et al. 2006; Haws et al. 2006; US EPA 2008). The TEF

is a factor which expresses the potency of a test chemical relative to a reference chemical, and the modified approach determines this factor based upon a cumulative distribution of relative potency (ReP) estimates calculated from multiple species. Similar to the TEF, the ReP estimates also express the potency of the test chemical relative to a reference chemical, however the ReP estimate is based on a specific toxic effect of a species (at the population, organism, cellular or biochemical levels) (US EPA 2008). The ReP estimate should only be calculated for instances where the toxic effect on the species from the test and reference chemicals is through the same mode of action (MoA), therefore the concentration-response curves are parallel, and the concentration addition (CA) model of joint action can be applied (Safe 1998). The TEF is then calculated from a selected percentile of the cumulative distribution function (CDF) of the RePs (Smith et al. submitted to IEAM). In order to determine which percentile of the CDF should be used, two novel tests were also introduced to determine the environmental relevance and robustness of the TLs generated from the selected percentile (Smith et al. submitted to IEAM). In summation, the TL method aims to transform the loads of each of the load constituents to a relative toxicity scale such that their weightings are proportional to their potential ecological impact, and thus they can be summed to calculate a total toxicity-based load.

In this paper we aim to show that by converting the pollutant loads to a TL using the method proposed in Smith et al. (submitted to IEAM), load reduction targets can have a greater alignment to improvements in ecosystem health. In addition, we also aim to demonstrate the capacity to optimise the standard TL calculation method to individual scenarios, and determine the TEFs for each of the GBR priority PSII herbicides to generate environmentally relevant and robust TLs.

METHODS

The Smith et al (submitted to IEAM) method for TL calculation was followed and adapted for the Great Barrier Reef case study. The definitions of key terms and formulas referred to in the methods below can be found in Smith et al. (submitted to IEAM). In brief, ReP estimates for a test chemical relative to a reference chemical are first determined for multiple species from ecotoxicity data published in the literature. It is ensured, where possible, that the data used to calculate the RePs are derived from preferably the same study and/or laboratory, and at a minimum, the same endpoint, exposure period and test conditions are used. A CDF is fitted to the RePs of multiple species for the test chemical (relative to the reference chemical). The CDF is considered reliable based on its fit to the ReP values and if it meets the minimum data requirements, i.e. ReP values from five species and four phyla (Warne et al. 2015). TEFs are then calculated from the CDF based on a selected percentile, the pollutant loads are multiplied by the TEF to generate the TLs, and lastly the TLs are tested for their environmental relevance and robustness (see below for a more detailed explanation). This last step is repeated, testing a range of TEFs (calculated from different percentiles of the CDF) for generating the most environmentally relevant and robust TLs. Greater detail of these steps, specific to this case study, are presented below.

Pesticide loads data

Annual loads (Ls) of five priority PSII herbicides (i.e. ametryn, atrazine, diuron, hexazinone and tebuthiuron) from nine catchments (i.e., Johnstone, Tully, Herbert, Haughton, Burdekin, Pioneer, Plane, Fitzroy and Burnett) were obtained from Turner et al. (2013) and Wallace et al. (2014). These five PSII herbicides all have the same MoA; inhibition of electron transport at the plastoquinones Q_A and Q_B in the D1 protein of the electron transport chain which drives photosynthesis in phototrophic species (Jones 2005). These type of

herbicides have previously been shown to follow the CA model (Faust et al. 2001; Backhaus et al. 2004; Magnusson et al. 2010).

For all catchments apart from the Haughton and Plane, the load of the main river (of the same name) was used. In the Haughton and Plane catchments, the loads for Barratta Creek and Sandy Creek respectively, were used.

Loads data for two years (2010 – 11 and 2011 – 12) were chosen to represent inter-annual variation (that may influence calculations) from a longer-term monitoring program (2009 – present). The methods for pesticide monitoring, sample analysis and load calculations are detailed in Turner et al. (2013) and Wallace et al. (2014).

Ecotoxicity data

Ecotoxicity data were sourced, collated and quality checked for calculation of RePs for individual species according to Smith et al. (submitted to IEAM). The US EPA Ecotoxicity Database (US EPA 2015) and the scientific literature were searched for appropriate toxicity data. Only data published since 1980 were accepted due to limitations in the experimental design and methods in earlier work (Warne 2001). In addition, only phototrophic species were considered as PSII herbicides specifically inhibit photosynthesis. Toxicity to non-phototrophic organisms would be exerted through a different mode of action (MoA) and therefore would not be suitable to calculate ReP values (for an explanation see Warne et al. 2015 and Smith et al. submitted to IEAM). Studies using freshwater, estuarine and marine phototrophic species were considered: (i) to ensure there were sufficient data to generate cumulative distribution functions (CDFs) for ReP values; and (ii) so the resulting TEF values would be relevant to all three aquatic ecosystems through which the pesticides are transported. The non-parametric Kruskal Wallis test was used to assess whether there were significant differences in the ReP values of species from freshwater, marine and

estuarine ecosystems. A non-parametric test was used due to the small sample sizes (i.e. sample sizes ranged from 5 – 18) and unequal group sizes.

Species of the GBR from important functional groups, such as coral (and their symbiotic zooxanthellae) and seagrass, were, where available, also included in the data set of each herbicide. In the case where no ecotoxicity data for ecologically relevant endpoints were available for these groups of species, data measuring photosynthetic inhibition were permitted (photosynthesis is not considered as an ecologically relevant endpoint for the derivation of the Australian and New Zealand Guidelines for Fresh and Marine Water Quality, i.e. those that affect the ecological competitiveness of a species e.g. lethality, immobilisation, growth, development, population growth, and reproduction or are relevant to the MoA of the chemical (ANZECC and ARMCANZ 2000; Warne et al. 2015)).

All ecotoxicity data were quality checked using the approach of Hobbs et al. (2005). Matched data sets (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 (Smith et al. submitted to IEAM)) were preferentially selected, however, in cases where the minimum data requirements for generating the CDF (five species from four phyla; Warne et al. 2015) could not be met, data from different studies were considered. In these cases, experimental conditions of the studies were compared and if they were sufficiently similar, the data were used. There was no preference for the duration of exposure in the toxicity tests, except to be ≥ 24 h. If data from multiple exposure periods for a species were available, the geometric mean of the RePs was calculated.

Reference chemicals and ReP distributions

All five priority PSII herbicides were assessed for their suitability to be the reference chemical (for an explanation see Smith et al. submitted to IEAM). Chemicals that met the

minimum data requirements for generating the CDF were considered to be more suitable as the reference chemical than those that did not.

The CDFs of ReP values for each combination of two PSII herbicides (e.g. ametryn and atrazine, ametryn and diuron, ametryn and hexazinone, ametryn and tebuthiuron) were generated using Burrlioz 2.0 (Barry and Henderson 2014). This software fits either a log-logistic or a Burr Type III distribution to the ReP data, depending on the number of data points. The log-logistic distribution is the default when there are 7 or fewer data (in this case ReP values) and the Burr Type III distributions are used when there are 8 or more data.

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

Calculating toxic loads

A TEF is derived from the ReP CFD and is used to calculate the TLs. The TEF is the maximum ReP of a selected percentile of species. Hence, the value of the TEF is dependent on the selected percentile of the CFD, and therefore, so is the value of the TL, as shown from Equations 1 and 2 from Smith et al. (submitted to IEAM):

$$TL_{i,p} = TEF_{i,p} \times L_i \quad \text{Equation 1}$$

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 p^{th} percentile of the ReP CDF of chemical i , and L_i = the load of chemical i (kg or tonnes); and,

$$TL_{mix,p} = \sum_i TL_{i,p} \quad \text{Equation 2}$$

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

Therefore, we examined the degree of variation in the calculated TLs based on a range of selected percentiles: the 50th, 70th, 75th, 80th and 95th percentiles if the test chemical was more toxic than the reference chemical and the 50th, 30th, 25th, 20th and 5th percentiles respectively, when the test chemical was less toxic than the reference chemical (the explanation for this is provided in Smith et al. submitted to IEAM). The above percentiles were used to calculate an annual TL_i for each PSII herbicide and annual TL_{mix} for each catchment.

Test for environmental relevance and robustness

The tests for environmental relevance and robustness (Smith et al. submitted to IEAM) were developed to jointly determine which percentile of the ReP CDF should be used to calculate the TEFs. Firstly, for the TLs to be environmentally relevant, the ratio of TL_i to TL_{mix} must be commensurate with the relative toxicity of chemical i compared to the toxicity of the mixture. This ensures that a reduction in TL_i will generate a relative improvement in ecosystem health. To do this, the $TL_{i,p}:TL_{mix,p}$ (where p is the percentile being tested) is compared to a similar ratio calculated using an independent method (the multi-substance potentially affected fraction, ms-PAF, method (Traas et al. (2002))). Secondly, the test for robustness determines the percentile which generates the most consistent ratio of $TL_i:TL_{mix}$ irrespective of the reference chemical used (Smith et al. submitted to IEAM).

To test which TEF_p generated the most environmentally relevant TLs, we followed the first stage of the ms-PAF procedure for response addition (outlined in Traas et al. 2002) to calculate hazard units (HUs). The methods for calculating the HUs of individual chemicals (HU_i) and mixtures (HU_{mix}) are described in detail in Smith et al. (submitted to IEAM). In

summary, species sensitivity distributions (SSDs) for the five PSII herbicides were generated using the methods of Warne et al. (2015), from which the HU for each PSII herbicide was calculated, according to Traas et al. (2002) (Equation 3):

$$HU_i = \frac{C_i}{\tilde{X}_i^j} \quad \text{Equation 3}$$

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 PSII herbicides in a mixture are then summed resulting in a hazard unit for the mixture (HU_{mix}) (Equation 4):

$$HU_{mix} = \sum_i HU_i \quad \text{Equation 4}$$

The contribution of each constituent to the HU_{mix} ($HU_i:HU_{mix}$) were then calculated and compared to the $TL_{i,p}:TL_{mix,p}$.

From the monitored PSII herbicide concentrations (Turner et al. 2013; Wallace et al. 2014), we calculated Ls, TLs and HUs based on the percentiles previously described, and compared the results. A simple scoring method was applied which gave a score of 1 when the $TL_i:TL_{mix}$ value (for a selected percentile) was within 1 S.D. of the average $HU_i:HU_{mix}$ and a score of 0 when it was outside this range. For each percentile, the percent of cases (catchments \times years) with a score of 1 was determined. The percentile with the highest score was regarded as the one which produced the most environmentally relevant TLs. The $L_i:L_{mix}$ ratios were also compared against the average $HU_i:HU_{mix}$ (± 1 S.D.) ratios and similarly scored in order to assess the environmental relevance of the summed Ls not converted to a TL.

A similar scoring method was applied for the robustness test. For each percentile, the $TL_i:TL_{mix}$ ratio was calculated using multiple reference chemicals. The difference in the $TL_i:TL_{mix}$ between reference chemicals was calculated and a difference $\leq 5\%$ was given a score of 1. For each percentile, the percent of cases (catchments \times years) with a score of 1 was then determined. The percentile with the highest score was regarded as the one which calculated the most robust TLs.

RESULTS & DISCUSSION

Ecotoxicity data and the reference chemicals

Matched ecotoxicity data were sourced from twenty-three studies found in the literature (see Table S1 in the supplementary material for the references). There were a limited number of studies in which ecotoxicity data for species from important functional groups of the GBR and other local species could be used (based on the data selection criteria). The GBR species included for which there were data were; the seagrass species, *Halodule uninervis* and *Zostera muelleri* (Flores et al. 2013), and zooxanthellae of the coral species, *Acropora millepora* (Negri et al. 2011). Other local species that had suitable data were *Nephroselmis pyriformis* and *Navicula sp.* (Magnusson et al. 2008), both estuarine benthic microalgae. Ecotoxicity data for zooxanthellae from the coral species *Seriatopora hystrix* (Jones and Kerswell 2003) were also used when limited data were available. Data for this species were not generally used as the test conditions did not meet all the preferred test requirements; i.e. the endpoint measured was photosynthetic inhibition and the exposure period was < 24 h.

The number of species and phyla for which matched toxicity data were available for each combination of test and reference chemical are presented in Table 1. For ease of

discussion, combinations of test and reference chemicals will be referred to using the following format — test:reference, e.g. atrazine:diuron, where atrazine is the test chemical and diuron is the reference chemical. In all test:reference combinations involving tebuthiuron, there were insufficient matched data to meet the minimum data requirements, although the combinations with diuron, atrazine and hexazinone were close, i.e. five species from three phyla. Combinations between ametryn and hexazinone also had insufficient data to produce reliable ReP CDFs. For all other combinations (excluding tebuthiuron) involving atrazine and diuron as the reference chemicals, reliable ReP CDFs were generated. Thus, atrazine and diuron were considered the best chemicals for use as a reference chemical.

Statistically significant differences in ReP values of species from different ecosystem types (freshwater, marine and estuarine) were tested on all combinations of test chemicals with atrazine and diuron as the reference chemicals (Table S2). In all cases there were no significant differences, which is further demonstrated in the CFDs presented in Figure 1 and Figure S1). Previous studies have compared the sensitivity of marine and freshwater species to different toxicants and the evidence is inconclusive and contradictory (e.g., Hutchinson et al. 1998; Leung et al. 2001; Wheeler et al. 2002), and all have at least one of the following limitations: limited data for each comparison; limited chemicals being compared; the fresh and marine toxicity data having markedly different organism composition, inappropriate species toxicity data being used (e.g. comparing the sensitivity distribution for all species to chemicals with specific modes of action). We note the guidance provided on combining toxicity data for freshwater and marine in the Technical Guidance Document for Deriving Environmental Quality Standards (EC 2011). This states that ‘in principle, ecotoxicity data for freshwater and saltwater organisms should be pooled for organic compounds’ but this ‘must be tested, except where a lack of data makes a statistical analysis unworkable’ (EC 2011).

To determine whether atrazine or diuron was the better reference chemical, the total number of species and phyla that constituted the CDFs for all combinations of test and reference chemicals (for atrazine and diuron) were compared. While atrazine had slightly more data compared to diuron as the reference chemical (i.e., ametryn:atrazine had data for one more species than ametryn:diuron and hexazinone:atrazine had data for one more phyla than hexazinone:diuron, Table 1), these were not sufficiently large differences to prefer atrazine over diuron as the reference chemical. Therefore, all subsequent calculations were done using both chemicals as the reference chemical.

Relative Potency Cumulative distribution functions

The ReP CDFs (Figure 1 and the rest can be found in the Supplementary Material (Figure S1)) indicated that the relative toxicities of the five priority PSII herbicides to the majority of phototrophic species were, in decreasing order, diuron, ametryn, hexazinone, atrazine and tebuthiuron. In most cases, the majority of ReP values (one per species) in each plot, lay either to the left or right of $\text{ReP} = 1$ (Figure 1b - d). However, in one case, ametryn:diuron (Figure 1a), species' ReP values were positioned on both sides of $\text{ReP} = 1$, indicating that diuron was more toxic than ametryn for slightly more than half the species (~65%).

The approximately equal potencies between ametryn and diuron caused some problems when choosing which percentile of the ReP CDF to use for TL calculations (i.e. see **Error! Reference source not found.**, Smith et al. submitted to IEAM). The tests for environmental relevance and robustness of different percentiles of the ametryn:diuron ReP CDF (data not shown) found that the 50th percentile should be used. This phenomenon is likely to hold true for all test chemicals with approximately equal toxicities to the reference chemical.

Toxic Equivalency Factors

The TEF_p values were extracted from all ReP CDFs with atrazine, diuron and hexazinone as the reference chemicals (Table 2). The test for robustness requires all chemicals to be tested as reference chemicals, however, the reliability of the ReP CDFs of hexazinone, ametryn and tebuthiuron (as reference chemicals) was low. It was decided to also include hexazinone as a reference chemical in the robustness test to validate the results of the diuron and atrazine tests; hexazinone was slightly more reliable than ametryn as tebuthiuron:hexazinone had data for five species whereas ametryn:tebuthiuron only had data for four (Table 1)

The values of the TEFs varied over the range of percentiles tested. In some cases the differences were large, for example, ametryn:atrazine TEFs differed by a factor of approximately eight. In other cases, the differences were much smaller; tebuthiuron:atrazine TEFs differed by a factor of 2.4. These variances in the TEFs resulted in differences in the contributions each PSII herbicide had to the TL_{mix} (i.e. the TL_i:TL_{mix} ratio), the degree of which was dependant on the catchment (Figure 2). For example, in the Fitzroy River (Figure 2A), there was up to a 40% difference in the diuron contribution to TL_{mix} depending on which percentile was selected. Similarly, the results from Barratta Creek (Haughton Catchment) showed that the contribution of diuron varied by up to 24% between percentiles (Figure 2B). In contrast, the Pioneer River results showed only small differences between percentiles in the TL_i:TL_{mix} values (Figure 2C). The reason for these differences between catchments is likely a result of the relative composition of the herbicides on a mass basis, which is discussed later. These observed differences in the TL_i:TL_{mix} in some catchments could lead to different management outcomes depending on the percentile used, which is why it is necessary to select a percentile that is environmentally relevant and robust.

There were also notable differences in the calculated TL_{mix} values depending on which percentile was used. The TL_{mix} values (diuron equivalent, i.e. using diuron as the reference chemical) calculated from the 50th percentile (i.e. $TL_{mix,50}$) were the largest and decreased as the percentile used to calculate the TLs increased (Figure 3). The greatest difference in a TL_{mix} calculation from one site was 83% at Burdekin River (2010–11) (Figure 3). Substantial differences in TL_{mix} estimates such as these could influence management actions. For example, a 60% reduction of $TL_{mix,50}$ (diuron equivalent) would require a greater reduction in the mass of pesticides than a 60% reduction of $TL_{mix,95}$. Conversely, the TL_{mix} values calculated using atrazine and hexazinone as the reference chemicals showed the opposite trend; the TL_{mix} values increased from $TL_{mix,50}$ to $TL_{mix,95}$. These results highlight the need for testing which percentile should be used to calculate the TEFs.

Environmental relevance and robustness

For the environmental relevance test, both atrazine and diuron were tested as the reference chemical and the scores were combined to give one score for each percentile tested. Scores from the environmental relevance and the robustness tests showed an inverse relationship (Figure 4). The environmental relevance score was highest at the 50th percentile in which 98% of the tested cases agreed with the HUs calculated by the independent method (ms-PAF) and decreased as the percentile increased. Conversely, robustness was highest at the 95th percentile (94%) and lowest at the 50th percentile (81%) (Figure 4). Accordingly, it was decided that the optimal percentile of the ReP CDF to use in calculating TEFs would have the highest possible scores from both tests. To determine this, the scores were fitted with regression models and the percentile where the two regressions intersected was defined as the optimal for calculating TLs. A 5-parameter sigmoid model and a linear model were fitted to the environmental relevance ($R^2 = 0.99$) and robustness scores ($R^2 = 0.99$),

respectively (SigmaPlot v12.5, Systat Software, San Jose, CA). The models intersected close to the 75th percentile, which scored 92% for environmental relevance and 89% for robustness (Figure 4). Although these scores were not the highest recorded individually for each test, they were the highest values for both scores at the same percentile. Fortuitously, this recommended percentile to calculate TEFs for the GBR case study is the same as that used by the World Health Organisation (van den Berg et al. 2006). But for other case studies this may not be the case. It is too early to say whether the similarities in these results are an indication that the 75th percentile can be used universally for any case study. Once more case studies are tested using the methods outlined in Smith et al. (submitted to IEAM) a more solid conclusion can be made.

The $L_i:L_{mix}$ ratios had a lower environmental relevance score compared to the corresponding ratios for TLs. Only 53% of cases agreed with the HU results (data not shown) which clearly demonstrates the need to weight pollutant loads based on their toxicity to ensure they are environmentally relevant. These results demonstrate that using TLs rather than Ls to quantify load reduction targets would more likely lead to a relative improvement in ecosystem health.

The high scores from the robustness test demonstrated that the method would provide the same result irrespective of which herbicide was used as the reference chemical. Therefore, the selection of reference chemical for calculating the TLs for the GBR, i.e. atrazine or diuron, had to be based on other factors. For example, atrazine is more commonly detected across catchments (Smith et al. 2012) and recent restrictions of diuron application may mean that diuron becomes less relevant to the region's water quality issues in the future. On the other hand, historically, diuron has been used as a reference chemical in expressing the concentration of PSII herbicides in catchments (Lewis et al. 2012) and marine waters of the GBR (Gallen et al. 2013, 2014). Given that stakeholders in the GBR region are familiar

with diuron equivalent concentrations, it makes sense to continue with this approach. Thus, to calculate the TL and TL_{mix} values for this case study, diuron was used as the reference chemical and the 75th percentiles of the ReP CDFs were used to derive the TEFs, except in the case of ametryn:diuron in which the 50th percentile was used.

Differences between loads and toxic loads

The recommended TEFs (Table 2) were used to derive TL and TL_{mix} values for the nine GBR catchments and were compared to their corresponding L_{mix} values. Substantial differences were found in the percent contributions of each PSII herbicide to the mixture depending on whether this was determined in terms of load ($L_i:L_{mix}$) or toxic load ($TL_i:TL_{mix}$) (Figure 5). Overall, the contribution of diuron to the TL_{mix} was larger than its contribution to the L_{mix} . This is logical given diuron is more toxic than atrazine, hexazinone and tebuthiuron. Conversely, due to the low toxicity of atrazine, its contribution to the TL_{mix} was considerably smaller than to the L_{mix} where it was the dominant herbicide. For example, at Barratta Creek (Haughton Catchment), where atrazine had the greatest contribution to L_{mix} (77%), its contribution to the TL_{mix} was reduced to just 11%. For the Fitzroy and Burdekin catchments, the contribution of the tebuthiuron load was also considerably reduced when it was converted to a TL. Tebuthiuron is used in the grazing industry and therefore contributes a large proportion of the load in catchments dominated by grazing, i.e. Burdekin and Fitzroy. However, due to the low toxicity of tebuthiuron compared to the other PSII herbicides, its contribution to the TL_{mix} is minor. Converting Ls to TLs therefore should change the focus of pesticide management. For example, in the Burdekin and Fitzroy catchments management action should shift from reducing tebuthiuron loads to reducing diuron loads, even though the mass is considerably less, in order to achieve the maximum reduction in toxicity and maximum increase in ecosystem health.

Converting Ls to TLs also changes the relative differences in the magnitude of the loads between catchments. The Fitzroy and Burdekin catchments have high L_{mix} values compared to the other catchments, but low TL_{mix} values (Figure 6A and 6B, respectively). This is because they are dominated by the least toxic of the five PSII herbicides; atrazine and tebuthiuron. Catchments with a higher dominance of diuron, e.g. Tully, Herbert and Pioneer rivers, have a higher TL_{mix} relative to the other catchments. The differences observed here between the Ls and TLs have important implications for prioritising catchment areas for management and using the wrong method could lead to less effective allocation of resources and smaller progress to achieving the overall goal of Reef Plan.

CONCLUSIONS

The results from this paper, (i) confirmed the proposition of Smith et al. (submitted to IEAM) that pollutant load mixtures should be converted to a toxic load to ensure ecosystem protection, (ii) determined the optimal percentile of the ReP CFD for calculating PSII herbicide TLs for the GBR, and (iii) demonstrated the case-specific optimisation of the method.

Weighting the mixture load constituents based on their toxicity changed the contribution of each of the PSII herbicides to the mixture load, and the magnitude of the loads for each catchment. In doing so, the toxic loads were shown to be more environmentally relevant than loads following verification by an independent mixture method (ms-PAF). Converting loads to toxic loads will better inform water quality managers on which chemicals and regions they should focus their actions, leading to a more environmentally beneficial allocation of resources.

Determining which percentile to use to calculate the TEF from the ReP CDF was shown to be a critical step in the TL method. The tests for environmental relevance and robustness provided a systematic approach for selecting the optimal percentile to calculate TLs. The 75th percentile was, for Great Barrier Reef catchments, determined to be the optimal percentile for improvements in ecosystem health and ensuring that load reductions were not biased by the toxicity selection of the reference chemical. However, it is not certain that the 75th percentile would be optimal in all case studies, and each should be assessed individually.

As was demonstrated in this paper, the TL method can be optimised for each case study, i.e.: a specific chemical mixture; the choice of reference chemical; inclusion of species from important functional groups of the GBR and other local species; and the TEF percentile - selected for environmental relevance, robustness, or both, as was demonstrated here. However, it may not be applicable to all pollutant load scenarios. For one, there is a requirement for matched ecotoxicity data sets to calculate the RePs. For some chemicals, particularly newer ones, ecotoxicity data may be limited and calculating enough ReP values to meet the minimum data requirements set for fitting the CDF may not be feasible. In addition, the ReP and TEF methods are limited to chemicals with the same MoA, i.e. those with parallel concentration-response curves. Thus for mixtures of chemicals with different MoA, calculating ReP values between two chemicals with non-parallel concentration-response curves would incur a degree of error depending on the variation in the shapes of the concentration-response curves. Thus it is not recommended to use this method for these types of mixtures, unless the error can be quantified.

ACKNOWLEDGEMENTS

The authors are grateful to Andrew Negri (Australian Institute of Marine Science, Townsville) and Stephen Lewis (James Cook University, Townsville) for reviewing the manuscript. In addition, we would like to acknowledge the financial contributions of the Queensland Department of Natural Resources and Mines for the collection of pesticide concentration data used in this study.

REFERENCES

Australian Government and Queensland Government. 2003. Reef Water Quality Protection Plan: For catchments adjacent to Great Barrier Reef World Heritage Area, October 2003. Brisbane (QLD), Australia: Reef Water Quality Protection Plan Secretariat, the State of Queensland.

Australian Government and Queensland Government. 2009. Reef Water Quality Protection Plan 2009 For the Great Barrier Reef World Heritage Area and adjacent catchments. Brisbane (QLD), Australia: Reef Water Quality Protection Plan Secretariat, the State of Queensland.

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.

Backhaus T, Faust M, Scholze M, Gramatica P, Vighi M, Grimme LH. 2004. Joint algal toxicity of phenylurea herbicides is equally predictable by concentration addition and independent action. *Environ Toxicol Chem* 23(2): 258-264.

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.

Brodie, J, Binney, J, Fabricius, K, Gordon, I, Hoegh-Guldberg, O, Hunter, H, O'Reagain, P, Pearson, R, Quirk, M, Thorburn, P, Waterhouse, J. 2008. Synthesis of evidence to support the scientific consensus statement on water quality in the Great Barrier Reef. Brisbane (QLD), Australia: Department of Premier and Cabinet, the State of Queensland.

Brodie, J, Waterhouse, J, Schaffelke, B, Kroon, F, Thorburn, P, Rolfe, P, Johnson, J, Fabricius, K, Lewis, S, Devlin, M, Warne, M, McKenzie, L. 2013. 2013 Scientific Consensus Statement: Land use impacts on Great Barrier Reef water quality and ecosystem condition. Brisbane (QLD), Australia: Reef Water Quality Protection Plan Secretariat, the State of Queensland 2013.

Faust M, Altenburger R, Backhaus T, Blanck H, Boedeker W, Gramatica P, Hamer V, Scholze M, Vighi M, Grimme LH. 2001. Predicting the joint algal toxicity of multi-component s-triazine mixtures at low-effect concentrations of individual toxicants. *Aq Toxicol* 56(1):13-32.

Flores F, Collier CJ, Mercurio P, Negri AP. 2013. Phytotoxicity of four photosystem II herbicides to tropical seagrasses. *PloS One* 8: e75798.

Gallen C, Devlin M, Paxman C, Banks A, Mueller J. 2013. Pesticide monitoring in inshore waters of the Great Barrier Reef using both time-integrated and event monitoring techniques (2012 - 2013). Brisbane (QLD), Australia: National Research Centre for Environmental Toxicology University of Queensland (Entox), University of Queensland.

Gallen C, Devlin M, Thompson K, Paxman C, Mueller J. 2014. Pesticide monitoring in inshore waters of the Great Barrier Reef using both time-integrated and event monitoring techniques (2013 - 2014). Brisbane (QLD), Australia: National Research Centre for Environmental Toxicology University of Queensland (Entox), University of Queensland.

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.

Haynes, D, Michalek-Wagner, K. 2000. Water quality in the Great Barrier Reef World Heritage Area: past perspectives, current issues and new research directions. *Mar Poll Bull* 41: 428-434.

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.

Hutchinson TH, Scholz N, Guhl W. 1998. Analysis of the ECETOC aquatic toxicity (EAT) database IV – Comparative toxicity of chemical substances to freshwater versus saltwater organisms. *Chemosphere* 36(1): 143–153.

Jones RJ, Kerswell AP. 2003. Phytotoxicity of photosystem II (PSII) herbicides to coral. *Mar Ecol Prog Ser* 261: 149-159.

Jones R, 2005. The ecotoxicological effects of Photosystem II herbicides on corals. *Mar Poll Bull* 51(5): 495-506.

Leung KMY, Morritt D, Wheeler JR, Whitehouse P, Sorokin N, Toy R, Holt M, Crane M. 2001. Can saltwater toxicity be predicted from freshwater data? *Mar Poll Bull* 41(11): 1007–1013.

Lewis SE, Schaffelke B, Shaw M, Bainbridge ZT, Rohde KW, Kennedy K, Davis AM, Masters BL, Devlin MJ, Mueller JF, Brodie, JE. 2012. Assessing the additive risks of PSII herbicide exposure to the Great Barrier Reef. *Mar Poll Bull* 65: 280-291.

Magnusson M, Heimann K, Negri AP. 2008. Comparative Effects of Herbicides on Photosynthesis and Growth of Tropical Estuarine Microalgae. *Mar Poll Bull* 56: 1545-1552.

Magnusson M, Heimann K, Quayle P, Negri AP. 2010. Additive toxicity of herbicide mixtures and comparative sensitivity of tropical benthic microalgae. *Mar Poll Bull* 60(11): 1978-1987.

Negri AP, Flores F, Röthig T, Uthicke S. 2011. Herbicides increase the vulnerability of corals to rising sea surface temperature. *Limnol Oceanogr* 56: 471-485.

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

Smith R, Middlebrook R, Turner R, Huggins R, Vardy S, Warne M. 2012. Large-scale pesticide monitoring across Great Barrier Reef catchments—paddock to reef integrated monitoring, modelling and reporting program. *Mar Poll Bull* 65: 117-1.

Smith RA, Warne MStJ, Mengersen K, Turner RDR. Submitted to IEAM. An Improved Method for Calculating Toxicity-Based Pollutant Loads (Toxic Loads): A Multispecies Toxic Equivalency Approach. Submitted to *Integ. Environ. Manag, Assess.*

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.

Turner RDR, Huggins R, Wallace R, Smith RA, Vardy S, Warne MStJ. 2013. Total suspended solids, nutrient and pesticide loads for rivers that discharge to the Great Barrier Reef: Great Barrier Reef Loads Monitoring 2010 – 2011. Brisbane (QLD), Australia: Department of Science, Information Technology, Innovation and the Arts. Water Sciences Technical Report 13, 1.

US EPA [United States Environmental Protection Agency]. 1998. Guidelines for ecological risk assessment (Volume 2). Washington DC, USA: US EPA. EPA/630/R-95.

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]. 2015. ECOTOX User Guide: ECOTOXicology Database System. Version 4.0. [cited 2012/09/09]. Available: <http://www.epa.gov/ecotox/>

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.

Wallace R, Huggins R, Smith RA, Turner RDR, Vardy S, Warne MStJ. 2014. Total suspended solids, nutrient and pesticide loads (2011 – 2012) for rivers that discharge to the

Great Barrier Reef – Great Barrier Reef Catchment Loads Monitoring Program 2011–2012. Brisbane (QLD), Australia: Department of Science, Information Technology, Innovation and the Arts.

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, Innovation and the Arts.

Waters DK, Carroll C, Ellis R, Hateley L, McCloskey GL, Packett R, Dougall C, Fentie, B. 2014. Modelling reductions of pollutant loads due to improved management practices in the Great Barrier Reef catchments – Whole of GBR. Toowoomba (QLD), Australia: Queensland Department of Natural Resources and Mines. Technical Report 1.

Wheeler JR, Leung KMY, Morritt D, Sorokin N, Rogers H, Toy R, Holt M, Whitehouse P, Crane M. 2002. Freshwater to saltwater toxicity extrapolation using species sensitivity distributions. *Env Tox Chem* 21(11):2459–2467.

FIGURES

Figure 1. Cumulative distribution functions (CDF) of the relative potency values for (A) ametryn:diuron, (B) hexazinone:diuron, (C) diuron:atrazine, and (D) tebuthiuron:atrazine. Symbols for ReP values represent freshwater (\circ), marine (\bullet), and estuarine (Δ) species. Blue dotted lines indicate equal toxicity, i.e. $\text{ReP} = 1$. All plots generated by BurrliOZ V2 (Barry and Henderson 2014).

Figure 2. Percent contributions of five PSII herbicides to the toxic load of the mixture ($\text{TL}_i:\text{TL}_{\text{mix}}$) calculated using different toxic equivalence factors (in turn calculated using different percentiles of the relative potency cumulative distribution functions). Contributions are for (A) the Fitzroy River, (B) Barratta Creek (Haughton Catchment) and (C) the Pioneer River, estimated using 2010–11 data.

Figure 3. Difference (expressed as a percentage) in the total toxic loads for each catchment (2010–11) calculated using different percentiles of the relative potency (ReP) cumulative frequency distributions (diuron equivalents) compared to the toxic loads calculated using the 50th percentile ($\text{TL}_{\text{mix},50}$).

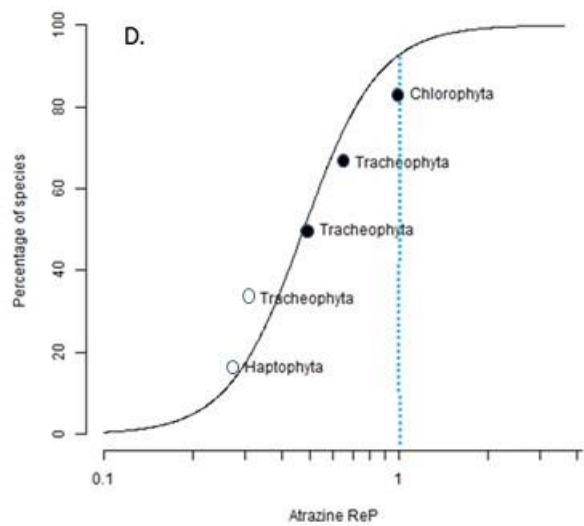
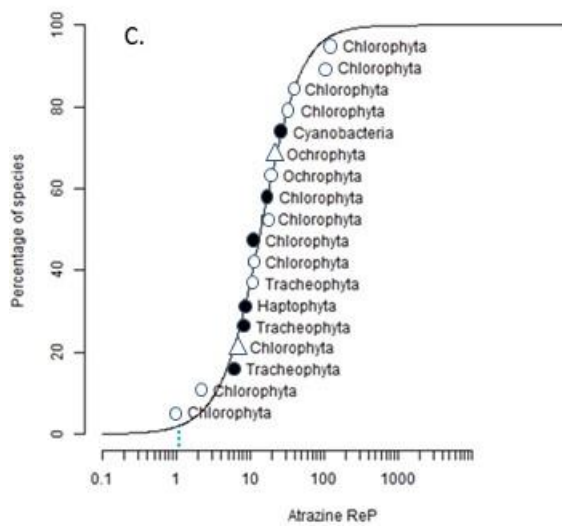
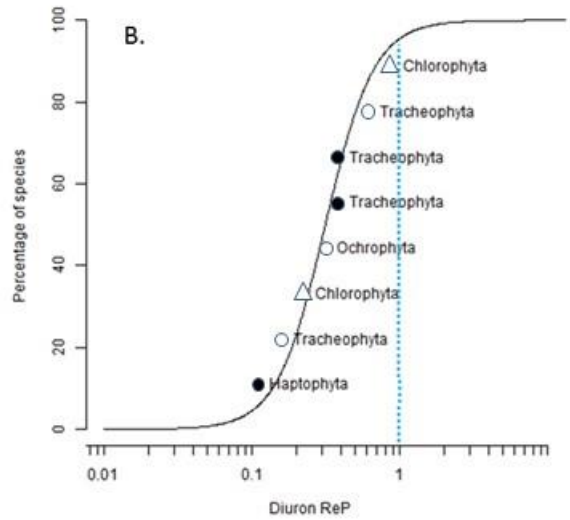
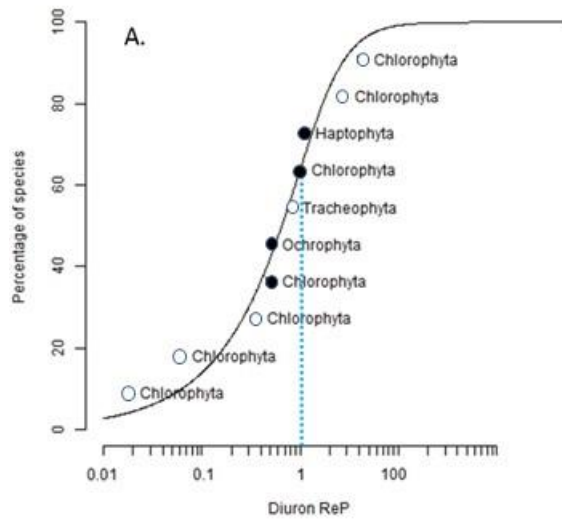
Figure 4. Scores of environmental relevance (\bullet) and robustness (\circ) as a function of the percentile of the relative potency (ReP) cumulative distribution function (CDF). The percentiles tested were the 50th, 70th (or 30th), 75th (or 25th), 80th (or 20th) and 95th (or 5th).

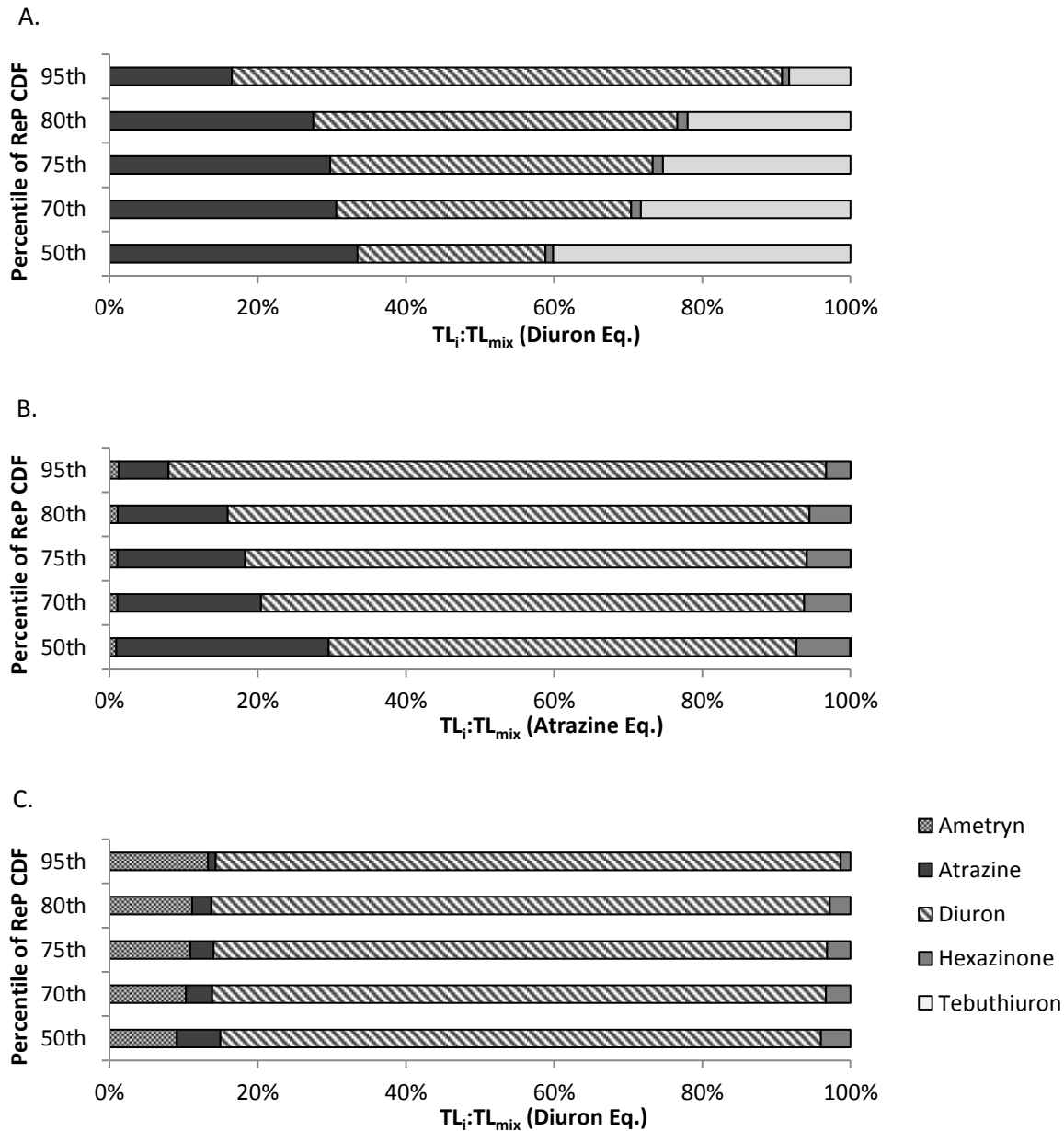
Figure 5. Percent contributions of the constituents of a mixture of five PSII herbicides calculated in terms of (A) the load ($\text{L}_i:\text{L}_{\text{mix}}$) and (B) the toxic load ($\text{TL}_i:\text{TL}_{\text{mix}}$) for nine Great Barrier Reef catchments in 2011–12. Toxic loads were calculated using the diuron equivalent TEFs presented Table 3.

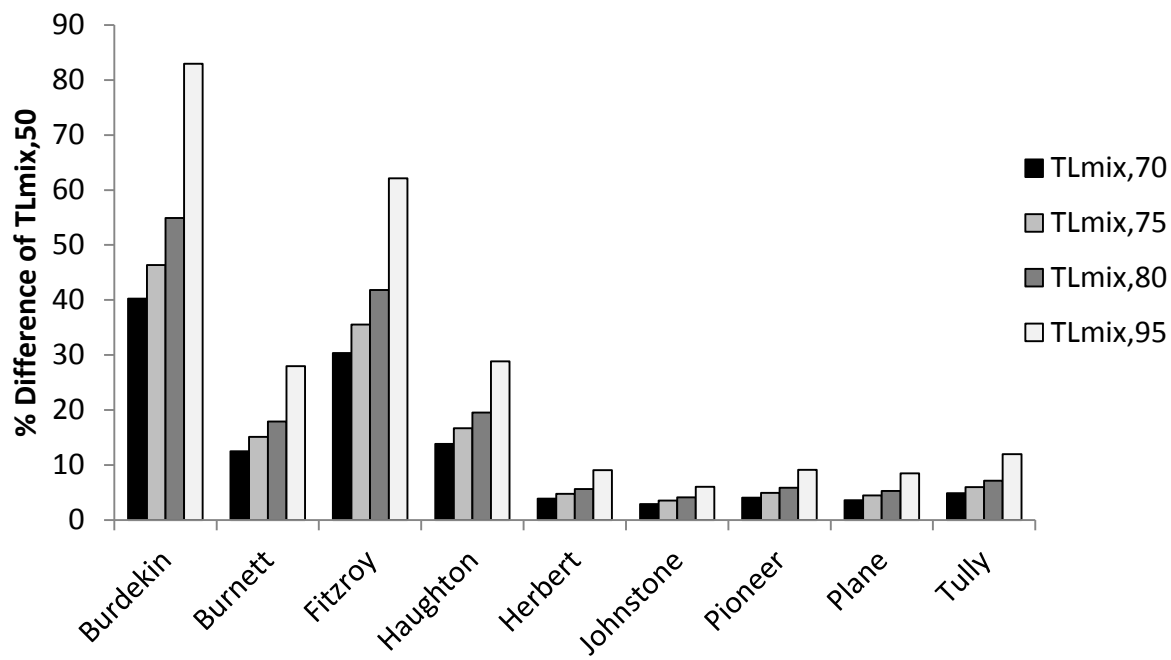
Figure 6. The 2011–2012 annual Total Loads (A) and Total Toxic Load (B) for nine Great Barrier Reef catchments. Total toxic loads were calculated using the bolded TEFs (diuron equivalent) in Table 3.

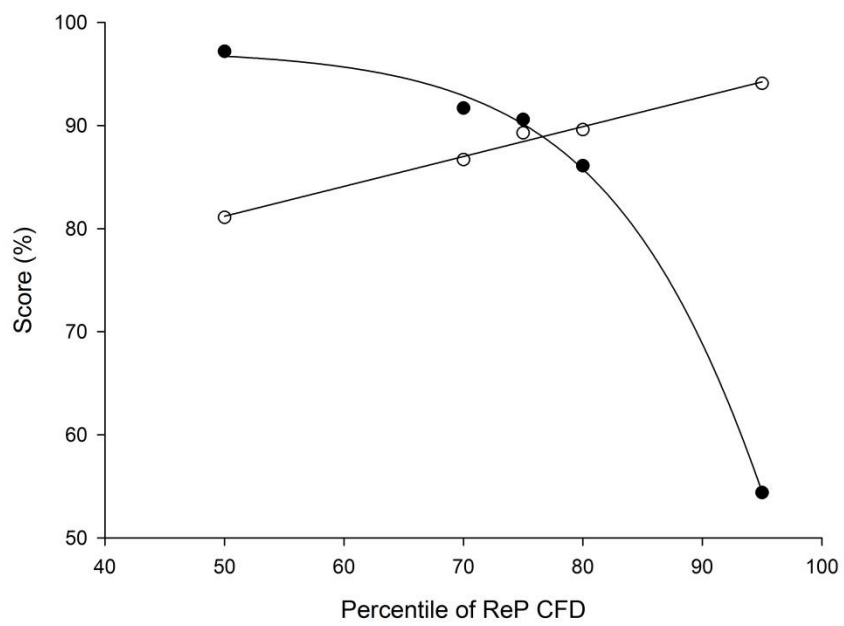
SUPPLEMENTARY FIGURE

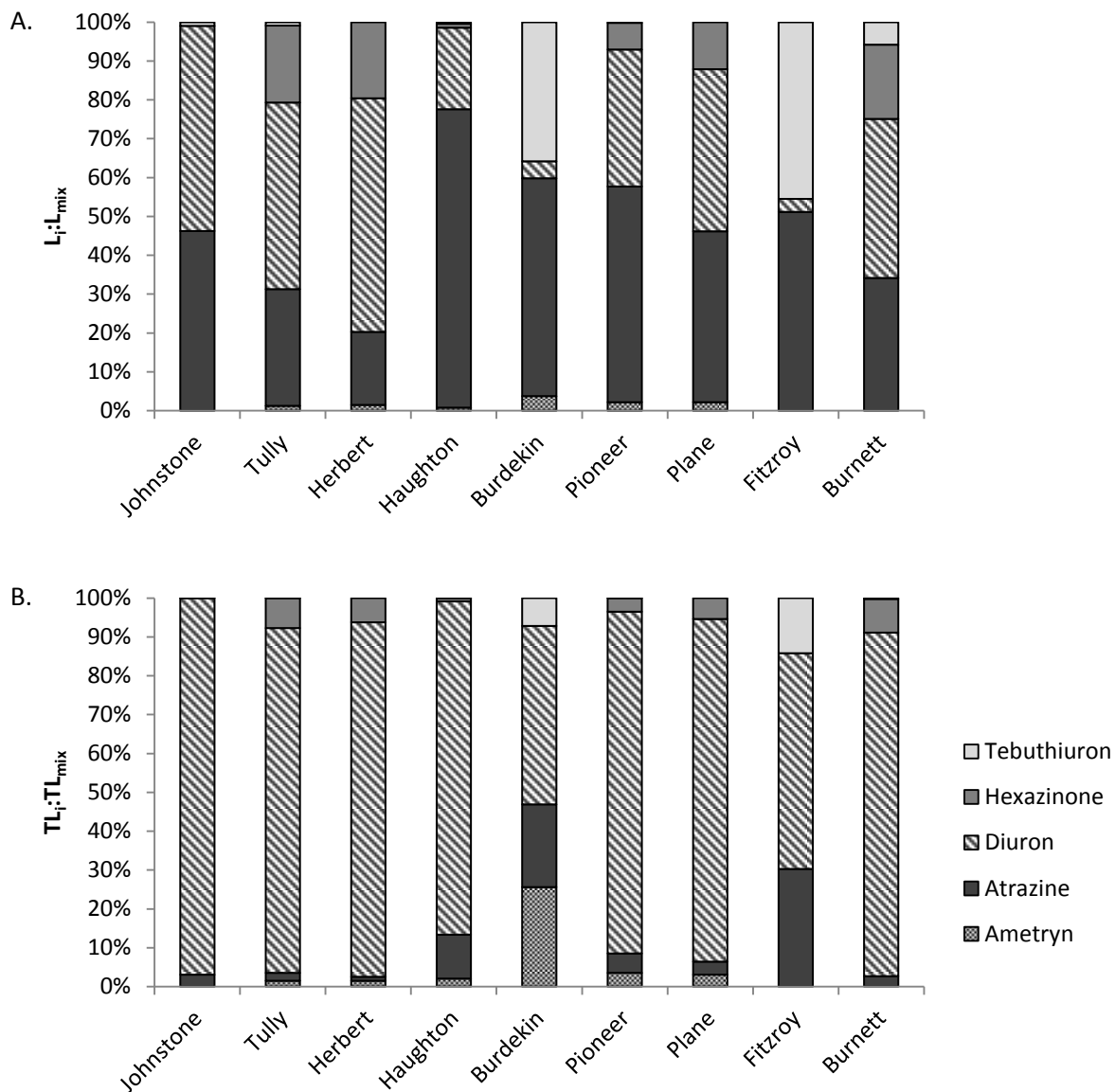
Figure S1. Cumulative distribution functions of the relative potency values for ametryn and atrazine, hexazinone and atrazine, hexazinone and ametryn, tebuthiuron and diuron, tebuthiuron and ametryn, and tebuthiuron and hexazinone.











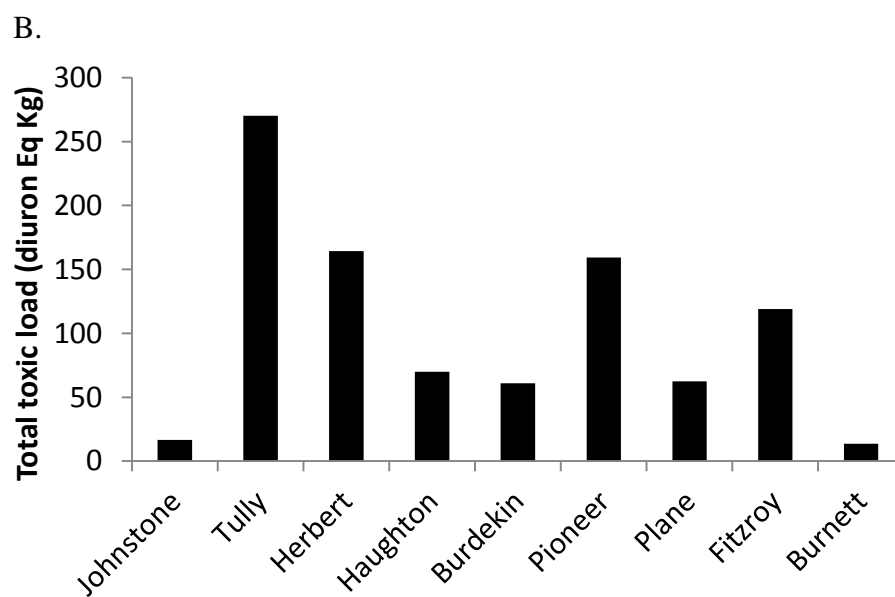
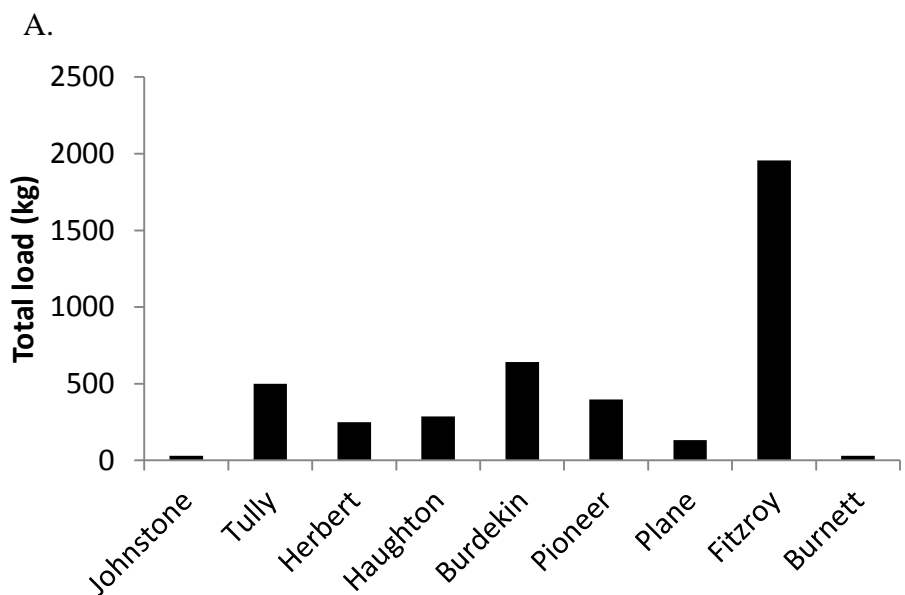


Table 1 Matrix indicating whether (Y = yes) or not (N = no) each of the priority photosystem II herbicides could act as a reference chemical to other PSII herbicides. The values in the parentheses are the number of species and phyla for which relative potency (ReP) values are available. The minimum acceptable number of species and phyla are five and four, respectively. References for the toxicity data underpinning these ReP values are provided in the supplementary material (Table S1).

	Ametryn	Atrazine	Diuron	Hexazinone	Tebuthiuron
Ametryn	-	Y (11/4)	Y (10/4)	N (4/4)	N (4/4)
Atrazine		-	Y (19/6)	Y (9/6)	N (5/3)
Diuron			-	Y(8/4)	N (5/3)
Hexazinone				-	N (5/3)
Tebuthiuron					-

Table 2 Toxic Equivalence Factors (TEFs) used to calculate Toxic Loads for testing ecological relevance and robustness using atrazine, diuron or hexazinone as the reference chemical. Underlined values are the recommended TEFs (refer to the subsequent Environmental Relevance and Robustness section).

Reference Chemical	Percentiles	Ametryn	Atrazine	Diuron	Hexazinone	Tebuthiuron
Atrazine	95th/5th	92	1	78	5.3	0.2
	80th/20th	31	1	32	4.4	0.32
	75th/25th	25	1	27	4.1	0.34
	70th/30th	21	1	24	3.8	0.37
	50th	11	1	14	2.7	0.48
Diuron	95th/5th	0.65*	0.012	1	0.1	0.006
	80th/20th	0.65*	0.031	1	0.19	0.016
	<u>75th/25th</u>	<u>0.65*</u>	<u>0.036</u>	<u>1</u>	<u>0.21</u>	<u>0.019</u>
	70th/30th	0.65*	0.042	1	0.23	0.021
	50th	0.65*	0.07	1	0.31	0.035
Hexazinone	95th/5th	14	0.14	11	1	0.091
	80th/20th	7.9	0.22	5.4	1	0.13
	75th/25th	7.1	0.24	4.8	1	0.14
	70th/30th	6.5	0.26	4.3	1	0.15
	50th	4.8	0.37	3	1	0.19

*The 50th percentile TEF for converting ametryn to diuron equivalents was used in all cases due to their similarity in potencies.