

Does toxicity test variability support bioavailability model predictions being within a factor of 2?

Gwilym A. V. Price^{A,B,*} , Jenny L. Stauber^{B,C}, Sarah Stone^{A,B}, Darren J. Koppel^D, Aleicia Holland^{B,C} and Dianne Jolley^B

Environmental context. Having appropriate and robust models used for developing water quality guidelines is critical for sound environmental management. Methods used to validate models have only been demonstrated appropriate for a small portion of data types used in these models. This study has found that models using certain data types would be more appropriately validated using alternative evaluation criteria. This study serves as an important reference for developing and evaluating robust models.

For full list of author affiliations and declarations see end of paper

***Correspondence to:**

Gwilym A. V. Price
Faculty of Science, University of
Technology Sydney, Ultimo, NSW 2007,
Australia
Email: gwilym.price@csiro.au

Handling Editor:

Kevin Wilkinson

Received: 13 May 2022
Accepted: 11 July 2022
Published: 8 September 2022

Cite this:

Price GAV et al. (2022)
Environmental Chemistry
19(3 & 4), 177–182. doi:10.1071/EN22050

© 2022 The Author(s) (or their
employer(s)). Published by
CSIRO Publishing.
This is an open access article distributed
under the Creative Commons Attribution-
NonCommercial-NoDerivatives 4.0
International License (CC BY-NC-ND)

OPEN ACCESS

ABSTRACT

Rationale. Bioavailability-based toxicity models for metals often have performance assessed by whether it can predict toxicity data within a factor of 2 of their paired observed toxicity data. This method has only been verified for median effect values (EC50) for acute fish and daphnia data, however toxicity models have been developed for a much broader range of effect levels (i.e. EC10/EC20) and species (e.g. microalga). This study tested whether the factor-of-2 rule is appropriate for a wider range of organisms and effect concentrations than previously studied.

Methodology. Toxicity estimate data from repeated tests conducted under the same conditions were collated to assess variation in results and compare this variation to a range of 4 (a factor of 2 above and below the mean) and a range of 9 (a factor of 3 above and below the mean) to assess if a factor-of-3 rule may be more appropriate for some species and effect levels.

Results and discussion. Overall, the factor-of-2 rule is broadly applicable for metal toxicity to a range of species for EC50 data. The EC10 datasets highlighted that larger variability exists in low effect levels and supported the use of a factor-of-3 rule, while the either the factor-of-2 or factor-of-3 rule could be applied to microalgae. The level of performance evaluation chosen may depend on the application of the bioavailability model. This study also found that while repeated toxicity test data is routinely generated, it is rarely published. Publication of such data would enable expansion of the present study to include inter-laboratory comparisons, an important consideration as most bioavailability models are based on data pooled from multiple sources.

Keywords: bioavailability, biotic ligand model, ecotoxicology, metal toxicity, model predictions, model validation, reference toxicants, water quality.

Introduction

In the last several decades there has been increased development, use, and interest in incorporating metal bioavailability models into regulatory water quality guidelines/criteria (Brix *et al.* 2020). Simple univariate regression models, such as the hardness-adjustment algorithm (USEPA 1985) have been used in water quality guidelines since the late 1980s. However, over the following decades more complex models, such as the biotic ligand model (BLM) (Di Toro *et al.* 2001) and multiple linear regression (MLR) models (Brix *et al.* 2017), which incorporate multiple water chemistry parameters have been developed to better predict metal toxicity to aquatic organisms.

Recently, bioavailability models have been developed using low effect levels (e.g. data based on effect concentrations that cause a 10% (EC10) and/or 20% (EC20) effect) based on chronic toxicity data. These inherently have higher uncertainty than models based on

higher effect levels, such as EC50 values, as there is typically greater uncertainty at the EC10 and EC20 values in a concentration-response model.

The increase in complexity, type and use of these models has resulted in a need for validation methods to test a model's predictive capacity. Garman et al. (2020) outlined several methods of model performance evaluation, including regression slope bias analysis and whether model predictions are within a factor of 2 (i.e. range of 4) of the observed toxicity estimate (e.g. EC50). Use of the factor-of-2 rule was first proposed by Di Toro et al. (2001) and Santore et al. (2001), and was based on a single dataset using a 96-h acute lethality test on larval fathead minnows (*Pimephales promelas*) exposed to copper (Erickson et al. 1996).

Several recent papers have called for the need for further assessment of the among-test variability for tests conducted under the same conditions in order to determine if the factor-of-2 rule is widely applicable (Garman et al. 2020; Peters et al. 2021). This rule has been examined by two other studies. Santore and Ryan (2015) assessed variation in *Daphnia magna* zinc acute lethality tests, while Meyer et al. (2018) examined a larger toxicity dataset where *D. magna* neonates were exposed separately to cadmium, copper, nickel or zinc. Additionally, Meyer et al. (2018) reanalysed *P. promelas* data from Erickson et al. (1996). Both studies found that the factor-of-2 rule was generally applicable across the two species.

No study has yet investigated the suitability of the factor-of-2 rule for microalga, despite several bioavailability models being recently developed (DeForest et al. 2018; Croteau et al. 2021; Peters et al. 2021) nor for low effect levels or chronic toxicity data for any organism. Peters et al. (2018) suggested that a factor-of-3 rule (i.e. a range of 9) may be more appropriate for low effect level and chronic data given the inherent increased uncertainty associated with these. Assessing the suitability of validation techniques like the factor-of-2 rule for these types of toxicity data is important as they are often preferred over acute EC50 data for water quality guideline development (Batley et al. 2018).

In this study, we report an analysis of the appropriateness of the factor-of-2 rule and the proposed factor-of-3 rule using an extensive collection of repeated toxicity datasets including freshwater and marine invertebrates and microalgae. Acute and chronic data across a range of endpoints at low and high effect levels were assessed. The results of these analyses serve as an important reference point for developing and evaluating bioavailability model performance.

Methods

Data sources

Toxicity estimate data were taken from a previously unpublished internal quality control reference toxicant database comprised of standardised tests used to assess test

repeatability and organism culture performance over time (CSIRO, unpubl. data). Test species consist of both freshwater and marine organisms. Additional data was sourced from published reference toxicant data in Stone et al. (2022) and Meyer et al. (2018).

Datasets were defined as having the same endpoint, test duration, test vessel, initial organism density (e.g. cell density for microalgae) and test water (laboratory prepared waters with the same chemical characteristics). A minimum of 5 datapoints (i.e. ECx values) were required per dataset and only tests with measured concentrations were included. In total 29 datasets representing 12 species (including microalgae, invertebrates, and fish), 547 toxicity tests, 3 contaminants (copper ($n = 21$), nickel ($n = 6$) and zinc ($n = 2$)), acute ($n = 11$ datasets) and chronic ($n = 18$) endpoints, and EC10 ($n = 7$ datasets) and EC50 ($n = 22$ datasets) data were collated (Supplementary Table S1).

Calculations and statistics

All statistical analyses were performed using the statistical environment RStudio (version 1.1.423; RStudio Team 2020) with figures produced using the extension packages *ggplot2* (Wickham 2016) and *ggpubr* (Kassambara 2020).

All data was tabulated into datasets to calculate means, standard deviations, and percentiles. Upper-lower prediction ratios (ULPRs) were used to assess toxicity dataset variability and were calculated as per Meyer et al. (2018) (Eqns 1, 2). Both 90- and 95%-ULPRs were calculated for each dataset using untransformed and log₁₀-transformed data. For the log-transformations all toxicity estimates within a dataset were transformed prior to calculations as per Meyer et al. (2018). Percentiles were calculated as shown in Eqn 3, using the mean toxicity estimate within a dataset, the z-score of normal distribution (Z) and the standard deviation (σ) of the dataset.

$$95\% - \text{ULPR} = \frac{97.5\text{th percentile}}{2.5\text{th percentile}} \quad (1)$$

$$90\% - \text{ULPR} = \frac{95\text{th percentile}}{5\text{th percentile}} \quad (2)$$

$$\text{Percentile} = \text{mean} \pm (Z \times \sigma) \quad (3)$$

Untransformed and transformed data were tested for normality using the Shapiro–Wilk test. ULPRs were compared to a range of 4 and 9, which are the ranges of deviation from observed toxicity values that the factor-of-2 and factor-of-3 rules suggest is a satisfactory fit for bioavailability models.

Several of the lower percentile calculations (i.e. 2.5th and 5th) for the untransformed data were less than 0 and therefore ULPRs could not be calculated. Additionally, several untransformed datasets were not normally distributed ($P < 0.05$). Therefore, for consistency log-transformed results were discussed in the present study as all log-

transformed results were normally distributed (except *Ceriodaphnia dubia* exposed to copper) and the transformation negates the issues of negative values at lower percentile calculations.

Applying a 95% prediction limit to assess toxicity variability may be unrealistic when considering bioavailability models developed with data pooled from numerous studies, laboratories and timepoints. This is often the case with most models in the literature predicting far less than 95% of data within a factor of 2 (Besser *et al.* 2021; Brix *et al.* 2021; Santore *et al.* 2021), with Peters *et al.* (2021) suggesting a model be deemed acceptable if 50% of data lies within a factor of 2 and 90% within a factor of 3 for lower effect levels. Based on this, the present study will discuss the results in terms of a 90%-ULPR and results for 95%-ULPRs are provided in Supplementary Table S2 to serve as a direct comparison to other studies.

Additionally, median, geomeans, coefficient of variation (CV) and maximum/minimum ratios (MMR) were calculated with all results provide in Supplementary Table S2. Significant differences between ULPRs was tested using the non-parametric unpaired Wilcoxon test following normality testing with the Shapiro–Wilk test. Standard deviation (s.d.) was used to specify variability (i.e. ± 1 s.d.) and ULPRs are expressed as median (interquartile range) throughout.

Results and discussion

Across all datasets the range of 90- and 95%-ULPRs were 1.6–12.7 (median: 3.2 (2.8–4.4)) and 1.7–20.6 (median: 3.9 (3.1–5.8)), respectively. All calculations and results for both untransformed and log-transformed data are provided in Supplementary Table S2.

EC10 versus EC50

In general, the ULPRs for the EC10 data in the present study do not support the factor-of-2 rule but do support a factor-of-3 rule. The ULPRs for the EC50 data support the factor-of-2 and agree with the findings of Meyer *et al.* (2018).

There were seven datasets based on EC10 values, which is much less than the 22 datasets based on EC50 values reflecting that EC50 values are more common acceptability criteria in reference toxicant testing (see examples in Stone *et al.* (2022) and Price *et al.* (2022)). The seven available EC10 datasets were comprised of 92 toxicity tests across four species (two invertebrates and two microalgae) and two contaminants (copper and nickel) with both acute and chronic data (Supplementary Table S1). The median 90%-ULPR for all EC10 datasets was 5.6 (3.5–7.1) (Fig. 1). Of the seven datasets, only two had 90%-ULPRs < 4 (complying with the factor-of-2 rule); however, six of the seven datasets had 90%-ULPRs < 9 complying with the factor-of-3 rule as suggested by Peters *et al.* (2018). The one dataset that fell

outside the factor-of-3 rule had a 90%-ULPRs of 12.7 for *Chlorella* sp. copper EC10 data. Meyer *et al.* (2018) reported similar increases in variability for *Daphnia magna* cadmium EC50 data, which was explained by age-related differences in cadmium sensitivity to *D. magna* neonates. However, this is unlikely the case for *Chlorella* sp. with this greater variability likely to reflect *Chlorella* sp. being highly sensitive to copper with a median EC10 value of $0.5 \mu\text{g Cu L}^{-1}$, which is close to the instrument detection limits (inductively coupled plasma – atomic emission spectroscopy (ICP-AES)).

For the EC50 data, 22 datasets comprising 455 toxicity tests across 12 species (invertebrates, microalgae and fish) and three contaminants (copper, nickel and zinc) were available (Supplementary Table S1). In comparison to the EC10 ULPRs, the EC50 ULPRs were lower and less variable (Fig. 1), with the median 90%-ULPR for all EC50 datasets being 3.2 (2.8–4.1) (Table 1). Of the 22 datasets, 18 datasets had 90%-ULPRs < 4 and all 90%-ULPRs were < 9. The ULPRs for the EC50 data in this study support the factor-of-2 rule and agree with the findings of Meyer *et al.* (2018).

Several contaminant- and species-matched datasets were available, which allowed for the comparison of EC10 and EC50 variability from the same tests (i.e. EC10 and EC50 values derived from the same concentration-response curve), rather than across all datasets. One matched dataset was available for a freshwater microalga, *Chlorella* sp., exposed to copper and three matched datasets were available for the marine copepod *Acartia sinjiensis*, exposed to copper (both acute and chronic) and nickel.

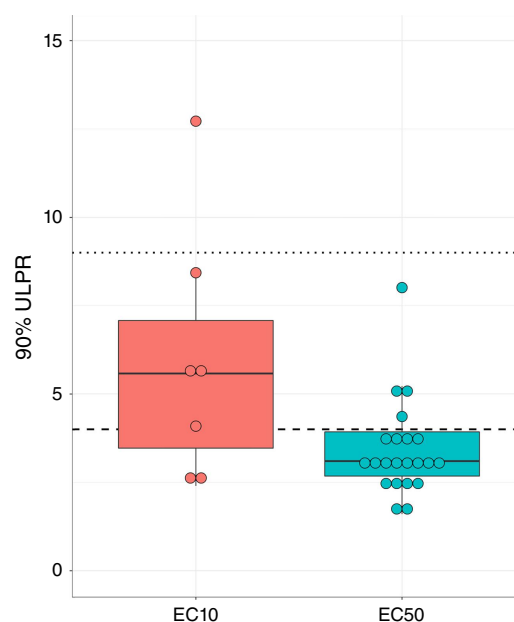


Fig. 1. Boxplots showing the 90% ULPRs for all EC10 and EC50 datasets. Boxplots span the interquartile range, with median shown, whiskers are $1.5 \times \text{IQR}$. Dots show ULPRs for individual datasets. Dashed line indicates the factor-of-2 rule threshold of 4 and the dotted line indicates the factor-of-3 threshold of 9.

Table 1. The 90% upper-lower prediction ratios for all EC10 and EC50 datasets and for the contaminant- and species-matched datasets. All individual dataset upper-lower prediction ratios are provided in Supplementary Table S1. n = total datapoints in dataset.

Dataset	n	90%-ULPR	
		Based on EC10	Based on EC50
All data	93/455 ^A	5.6 (3.5–7.1) ^B	3.2 (2.8–4.1) ^B
<i>Chlorella</i> sp. (copper)	11	13	4.9
<i>Acartia sinjiensis</i> (copper – acute)	37	5.7	2.7
<i>Acartia sinjiensis</i> (copper – chronic)	5	8.4	3.4
<i>Acartia sinjiensis</i> (nickel – chronic)	9	2.8	1.4

^AEC10 datapoints/EC50 datapoints.

^BMedian (IQR).

When comparing the *Chlorella* sp. copper matched EC10 and EC50 ULPRs, the EC10 90%-ULPR was much larger, at 12.7, while the EC50 90%-ULPR was 4.9. Larger EC10 ULPRs compared to the matched EC50 ULPRs were also found in the *A. sinjiensis* data (Table 1). These matched comparisons provide further evidence to support comments of Peters et al. (2018) about larger EC10 variability, especially given these matched EC10 and EC50 values are estimates from the same concentration-response models.

Chronic versus acute

Comparing chronic datasets to acute datasets at both the EC10 and EC50 level did not result in any differences between ULPRs (Fig. 2). When considering all datasets, the 90%-ULPR for EC50 values was not significantly different between chronic and acute datasets ($P = 0.08$, $z = 1.77$). The overall median 90%-ULPR for the acute and chronic EC50 values was 2.8 (2.4–3.1), and 3.3 (3.0–4.0), respectively. For the EC10 values, there were limited datasets available for acute toxicity ($n = 2$); however, ULPRs were similar for both the acute and chronic datasets, as shown in Fig. 2.

Only three datasets were available to compare contaminant- and species-matched acute and chronic ULPRs, with data available for the marine copepod *A. sinjiensis* and the marine urchin *Heliocidaris tuberculata*. *A. sinjiensis* acute and chronic copper EC50 ULPRs were similar, with acute and chronic EC50 90%-ULPRs of 2.9 and 3.3, respectively. Matched acute and chronic EC50 ULPRs for *H. tuberculata* had similarly small differences between the two datasets, with acute and chronic EC50 90%-ULPRs of 2.8 and 4.0, respectively. The acute and chronic copper EC10 ULPRs for *A. sinjiensis* had a larger difference compared to the EC50 ranges above, with acute and chronic EC10 90%-ULPRs of 5.7 and 8.4, respectively.

The comparisons with all datasets and the EC50 contaminant- and species-matched datasets suggest that

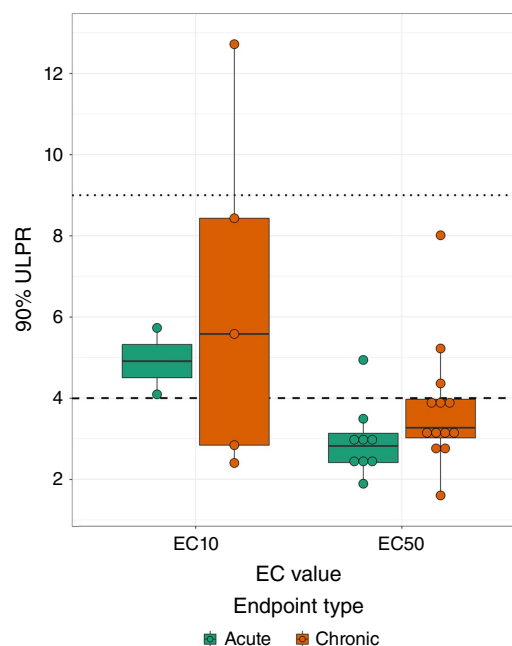


Fig. 2. Boxplots comparing the acute and chronic 90% ULPRs for all EC10 and EC50 datasets. Boxplots span the interquartile range, with median shown, whiskers indicate the $1.5 \times \text{IQR}$. Dots show ULPRs for individual datasets. Dashed line indicates the factor-of-2 rule threshold of 4 and the dotted line indicates the factor-of-3 threshold of 9.

acute and chronic test variability are similar. In addition, the median values for the acute and chronic 90%-ULPRs were both < 4 , broadly suggesting that the data, regardless of whether it is chronic or acute, supports the factor-of-2 rule. The EC10 contaminant- and species-matched datasets suggest differences may be present between acute and chronic EC10 data, as the chronic ULPR range is much larger than the acute range. However, this is likely related to the greater variability in EC10 data (discussed earlier), rather than specific differences between acute and chronic data. Furthermore, this variability is based on a single species and contaminant. More data would be useful to assess the differences in variability at the EC10 level for matched acute and chronic data.

Microalgae

Of the 10 EC50 microalgal datasets, 5 were freshwater species and 5 were marine species (Supplementary Table S1). The median EC50 90%-ULPRs for the freshwater and marine species were 4.0 (3.8–4.4) and 2.9 (2.7–3.1), respectively. These were not significantly different ($P = 0.095$, $z = 1.77$), but the marine species did generally appear to have lower ULPRs (Supplementary Fig. S1). As freshwater microalgal species are the current focus of algae bioavailability modeling, the discussion will focus on these results (DeForest et al. 2020; Croteau et al. 2021). All data analysis and results for

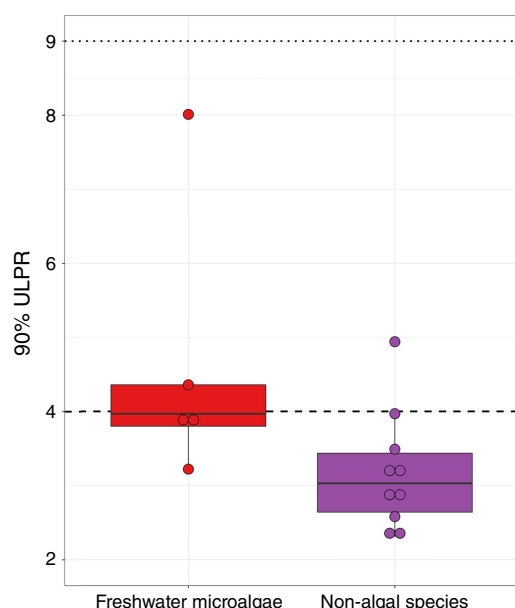


Fig. 3. Boxplots comparing the 90% ULPRs for all freshwater microalgae and non-algal EC50 datasets. Boxplots span the inter-quartile range, with median shown, whiskers indicate the $1.5 \times \text{IQR}$. Dots show ULPRs for individual datasets. Dashed line indicates the factor-of-2 rule threshold of 4 and the dotted line indicates the factor-of-3 threshold of 9.

the marine species are provided in Supplementary Table S2, and the results were similar to the freshwater species.

The 90%-ULPR for freshwater microalgal EC50 data ranged from 3.2 to 8.0, with a median value of 4.0 (3.8–4.4). Based on the median, the factor-of-2 rule may be suitable, with 3 of the 5 datasets with ULPRs < 4 , however a factor-of-3 rule appears more applicable with all ULPRs < 9 .

Comparing the freshwater microalgae and non-algal species ULPRs shows a small, yet significant difference ($P = 0.027$, $z = 2.27$), with non-algal species having a slightly lower median 90%-ULPR of 2.9 (2.4–3.3) compared to the 4.0 (3.8–4.4) for microalgae (Fig. 3). When comparing the microalgae ULPRs to other commonly used taxa for bioavailability modelling, such as daphnids and fish, the ULPRs for microalgae were similar. The only daphnid dataset in the present study was for *C. dubia* exposed to copper, which had a 90%-ULPR of 3.1. The *D. magna* data used in Meyer *et al.* (2018) had a median 90%-ULPR of 2.4 and *P. promelas* had a median 90%-ULPR of 3.5. In general, the microalgal test variability does not appear to be much larger than non-algal species.

Recommendations

Reference toxicant data is routinely generated during contaminant toxicity studies, however the data is rarely published. Publication of such data would enable expansion of the present study to other organisms, contaminants and

effect levels allowing for further assessment of validation techniques for bioavailability modelling. This would also allow for inter-laboratory comparisons which is important given most bioavailability models are developed using data from numerous sources. However, care is needed when making such comparisons as different laboratories do not necessarily use the same culture and/or testing media.

Conclusions

The data in the present study indicated that the factor-of-2 rule is broadly applicable for metal toxicity to a range of species for EC50 data, generally agreeing with the previous analysis by Meyer *et al.* (2018). The EC10 data highlighted that larger variability exists in low effect levels and supported the use of the factor-of-3 rule as recommended by Peters *et al.* (2018). Overall, either the factor-of-2 or factor-of-3 rule could be applied to microalgal data and the rule chosen for model performance evaluation may depend on the application of the bioavailability model. Given that most bioavailability models are developed using data from numerous sources, future assessments of inter-laboratory variability for matched tests (i.e. the same species and conditions) would be valuable. However, this may be difficult as differences in sensitivities can arise from small changes in water chemistry between laboratories and strains of the same species.

Supplementary material

Supplementary material is available [online](#).

References

- Batley GE, Van Dam RA, Warne M, Chapman J, Fox DR, Hickey CW, Stauber JL (2018) Technical Rationale for Changes to the Method for Deriving Australian and New Zealand Water Quality Guideline Values for Toxicants - Update of 2014 version. Prepared for the revision of the Australian and New Zealand Guidelines for Fresh and Marine Water. (CSIRO: Canberra, ACT, Australia)
- Besser JM, Ivey CD, Steevens JA, Cleveland D, Soucek D, Dickinson A, Genderen EJ, Van Ryan AC, Schlekot CE, Garman ER, Middleton ET, Santore RC (2021) Modeling the Bioavailability of Nickel and Zinc to *Ceriodaphnia dubia* and *Neocloeon triangulifer* in Toxicity Tests with Natural Waters. *Environmental Toxicology and Chemistry* **40**, 3049–3062. doi:10.1002/etc.5178
- Brix KV, Deforest DK, Tear L, Grosell M, Adams WJ (2017) Use of Multiple Linear Regression Models for Setting Water Quality Criteria for Copper: A Complementary Approach to the Biotic Ligand Model. *Environmental Science and Technology* **51**, 5182–5192. doi:10.1021/acs.est.6b05533
- Brix KV, DeForest DK, Tear L, Peijnenburg W, Peters A, Middleton ET, Erickson R (2020) Development of Empirical Bioavailability Models for Metals. *Environmental Toxicology and Chemistry* **39**, 85–100. doi:10.1002/etc.4570
- Brix KV, Tear L, Santore RC, Croteau K, DeForest DK (2021) Comparative Performance of Multiple Linear Regression and Biotic Ligand Models for Estimating the Bioavailability of Copper in Freshwater. *Environmental Toxicology and Chemistry* **40**, 1649–1661. doi:10.1002/etc.5012
- Croteau K, Ryan AC, Santore R, DeForest D, Schlekot C, Middleton E, Garman E (2021) Comparison of Multiple Linear Regression and

- Biotic Ligand Models to Predict the Toxicity of Nickel to Aquatic Freshwater Organisms. *Environmental Toxicology and Chemistry* **40**, 2189–2205. doi:10.1002/etc.5063
- DeForest DK, Brix KV, Tear LM, Adams WJ (2018) Multiple linear regression models for predicting chronic aluminum toxicity to freshwater aquatic organisms and developing water quality guidelines. *Environmental Toxicology and Chemistry* **37**, 80–90. doi:10.1002/etc.3922
- DeForest DK, Brix KV, Tear LM, Cardwell AS, Stubblefield WA, Nordheim E, Adams WJ (2020) Updated Multiple Linear Regression Models for Predicting Chronic Aluminum Toxicity to Freshwater Aquatic Organisms and Developing Water Quality Guidelines. *Environmental Toxicology and Chemistry* **39**, 1724–1736. doi:10.1002/etc.4796
- Di Toro DM, Allen HE, Bergman HL, Meyer JS, Paquin PR, Santore RC (2001) Biotic Ligand Model of the Acute Toxicity of Metals. 1. Technical Basis. *Environmental Toxicology and Chemistry* **20**, 2383–2396. doi:10.1002/etc.5620201034
- Erickson RJ, Benoit DA, Mattson VR, Leonard EN, Nelson Jr HP (1996) The effects of water chemistry on the toxicity of copper to fathead minnows. *Environmental Toxicology and Chemistry* **15**, 181–193. doi:10.1002/etc.5620150217
- Garman ER, Meyer JS, Bergeron CM, Blewett TA, Clements WH, Elias MC, Farley KJ, Gissi F, Ryan AC (2020) Validation of Bioavailability-Based Toxicity Models for Metals. *Environmental Toxic and Chemistry* **39**, 101–117. doi:10.1002/etc.4563
- Kassambara A (2020) Package ggpubr; 'ggplot2' Based Publication Ready Plots. R package version 0.4.0. Available at <https://CRAN.R-project.org/package=ggpubr>
- Meyer JS, Traudt EM, Ranville JF (2018) Is the Factor-of-2 Rule Broadly Applicable for Evaluating the Prediction Accuracy of Metal-Toxicity Models? *Bulletin of Environmental Contamination and Toxicology* **100**, 64–68. doi:10.1007/s00128-017-2258-4
- Peters A, Merrington G, Schlekot C, De Schampelaere K, Stauber J, Batley G, Harford A, van Dam R, Pease C, Mooney T, Warne M, Hickey C, Glazebrook P, Chapman J, Smith R, Krassoi R (2018) Validation of the nickel biotic ligand model for locally relevant species in Australian freshwaters. *Environmental Toxicology and Chemistry* **37**, 2566–2574. doi:10.1002/etc.4213
- Peters A, Merrington G, Stauber J, Golding L, Batley G, Gissi F, Adams M, Binet M, McKnight K, Schlekot CE, Garman E, Middleton E (2021) Empirical Bioavailability Corrections for Nickel in Freshwaters for Australia and New Zealand Water Quality Guideline Development. *Environmental Toxicology and Chemistry* **40**, 113–126. doi:10.1002/etc.4900
- Price GAV, Stauber JL, Holland A, Koppel DJ, Van Genderen EJ, Ryan AC, Jolley DF (2022) The influence of hardness at varying pH on zinc toxicity and lability to a freshwater microalga, *Chlorella* sp. *Environmental Science: Processes & Impacts* **24**, 783–793. doi:10.1039/d2em00063f
- RStudio Team (2020) RStudio: Integrated Development for R. Available at <http://www.rstudio.com/>
- Santore RC, Ryan AC (2015) Development and application of a multi-metal multibiotic ligand model for assessing aquatic toxicity of metal mixtures. *Environmental Toxicology and Chemistry* **34**, 777–787. doi:10.1002/etc.2869
- Santore RC, Di Toro DM, Paquin PR, Allen HE, Meyer JS (2001) Biotic ligand model of the acute toxicity of metals. 2. Application to acute copper toxicity in freshwater fish and *Daphnia*. *Environmental Toxicology and Chemistry* **20**, 2397–2402. doi:10.1002/etc.5620201035
- Santore RC, Croteau K, Ryan AC, Schlekot C, Middleton E, Garman E, Hoang T (2021) A Review of Water Quality Factors that Affect Nickel Bioavailability to Aquatic Organisms: Refinement of the Biotic Ligand Model for Nickel in Acute and Chronic Exposures. *Environmental Toxicology and Chemistry* **40**, 2121–2134. doi:10.1002/etc.5109
- Stone S, Koppel D, Binet MT, Simpson SL, Jolley DF (2022) Pulse-Exposure Toxicity of Ammonia and Propoxur to the Tropical Copepod *Acartia sinjiensis*. *Environmental Toxic and Chemistry* **41**, 208–218. doi:10.1002/etc.5235
- Stephan CE, Mount DI, Hansen DJ, Gentile JH, Chapman GA, Brungs WA (1985) 'Guidelines for deriving numerical national water quality criteria for the protection of aquatic organisms and their uses.' (US Environmental Protection Agency: Washington, DC, USA)
- Wickham H (2016) 'Elegant Graphics for Data Analysis.' (Springer-Verlag: New York, USA)

Data availability. The data that support this study are available in the article and accompanying online supplementary material. Code (R) used to make calculations reported in the article can be made available upon request to the corresponding author.

Conflicts of interest. There are no conflicts to declare.

Declaration of funding. Gwilym Price was supported by an Australian Government Research Training Program Scholarship.

Author affiliations

^AFaculty of Science, University of Technology Sydney, Ultimo, NSW 2007, Australia.

^BCSIRO Land and Water, Lucas Heights, NSW, Australia.

^CDepartment of Environment and Genetics, School of Life Science, Centre for Freshwater Ecosystems, La Trobe University, Albury/Wodonga Campus, Vic., Australia.

^DAustralian Institute of Marine Science, Crawley, WA, Australia.