Assessment Factors in Species Sensitivity Distributions for the Derivation of Guideline Values for Aquatic Contaminants

David R. Fox[1,2]* and Graeme E. Batley[3]

[1]Environmetrics Australia, Beaumaris, Victoria 3193, Australia
[2]University of Melbourne, Parkville, Victoria 3010, Australia
[3]CSIRO Land and Water, Locked Bag 2007, Kirrawee NSW 2232, Australia

*Email: david.fox@environmetrics.net.au

Abstract

The development of the Species Sensitivity Distribution (SSD) more than 30 years ago was in direct response to the many criticisms concerning the use of subjective Assessment (or Application) Factors (AFs) in widespread use at the time. While not perfect, SSD modelling is statistically defensible whereas AFs are not. While intuitively appealing, we believe recent guidance recommending the use of AFs in conjunction with SSD modelling is concerning and has the potential to impose unnecessary, time-consuming, and expensive follow-up investigations on both regulators and the regulated. This paper outlines our concerns and presents results of more contemporary analyses to quantify the impact of arbitrary scaling of SSD model outputs.

Key words: assessment factor, species sensitivity distribution, guideline value, model averaging

Introduction

Species sensitivity distributions (SSDs) are widely used for the derivation of water quality guideline values (GVs), water quality criteria, and environmental quality standards (EQSs) for chemical contaminants in aquatic environments (ANZG 2018; Stefan 2002; CCME 2007: EC 2011, 2018). The development of the SSD methodology has been slow and incremental and despite several well-documented shortcomings (Fox 2016 and references therein), it remains the most scientifically rigorous, statistically defensible, and ecotoxicologically relevant approach.

The purpose of this paper is to highlight what we perceive to be a disturbing trend whereby the statistical rigour embedded in the SSD approach is being compromised by a desire to apply an 'assessment (or application) factor' (AF) (BC 2019; EC 2011; Belanger and Carr 2019) to the output of the SSD - typically the HC_5 (concentration hazardous to 5% of species) also known as the PC_{95} (concentration protective of 95% of species).

The European Commission (EC 2011) recommended applying a default AF of 5 but noted that this may be reduced 'where evidence removes residual uncertainty'. Their input dataset should have at least 10 datapoints for at least 8 taxonomic groups. As a minimum, the following points are considered:

"• the overall quality of the database and the endpoints covered, e.g., if all the data are generated from "true" chronic studies (e.g., covering all sensitive life stages);

• the diversity and representativity of the taxonomic groups covered by the database, and the extent to which differences in the life forms, feeding strategies and trophic levels of the organisms are represented;

• knowledge on presumed mode of action of the chemical (covering also long-term exposure). Details on justification could be referenced from structurally similar substances with established mode of action;

• statistical uncertainties around the HC_5 estimate, e.g., reflected in the goodness of fit or the size of confidence interval around the 5th percentile, and consideration of different levels of confidence (e.g. by a comparison between the median estimate of the HC_5 with the lower estimate (90% confidence interval) of the HC_5);

• comparisons between field and mesocosm studies, where available, and the HC_5 and mesocosm/field studies to evaluate the level of agreement between laboratory and field evidence."

A more recent European review of current thinking on the use of SSDs (ECETOC 2014) recommended a default AF of 2 for the HC_5 of the SSD, but indicated that this value can 'be further refined based on characteristics of the toxicity data, i.e., representativeness, mode of action, interspecies variability and uncertainty'. In British Columbia (BC 2019), the minimum AF is 2, applied to a dataset of 15, with appropriate taxonomic coverage and no additional residual uncertainties, otherwise an AF of 5 is applied.

In Australia and New Zealand, the SSDs are applied to a dataset of at least 8 (and desirably 15) and the resultant guideline values (GVs) are classified as having a reliability that is very high, high, moderate or low, based on the adequacy of the sample size, and the goodness of fit to the SSD model and whether the dataset contains all chronic data or a mixture of chronic and converted acute data (Warne et al. 2018). No AFs are applied for reasons that we discuss below. Implicit in the GV applications is that the derived values are not to be used in a punitive manner, but, if exceeded are used as triggers for further investigations involving other lines of evidence, typically evaluated in

a weight of evidence framework (ANZG 2018). It is worth noting that, in Europe, the EQS derived using an SSD is a regulatory, legally binding limit, rather than a guideline value.

Dealing with uncertainty

At the heart of the assessment factor approach is an understandable desire to adjust a GV to account for the many and varied sources of uncertainty that may affect its quantification. While large assessment factors are commonly used as the default approach to GV derivation where there are few toxicity data, our criticisms are mainly reserved for their use in conjunction with species sensitivity distribution (SSD) modelling. We set aside the companion process of using ACRs (acute-to-chronic ratios) to ostensibly convert acute data into chronic equivalents. While the choice of the ACR is somewhat arbitrary, it doesn't have to be and as was shown by Fox (2006), this practice can be put on a more substantive statistical footing.

SSD modelling is an inherently statistical approach whereas the determination and application of an assessment factor is not. We have no issue with *scaling* (we do it all the time when we change units), but what assessment factors are doing is using an HCp (obtained as a point estimate from the fitted SSD) whose statistical properties (e.g. bias and variance/uncertainty) are well understood and then adjusting it in a way that lacks transparency and reproducibility while simultaneously altering the claimed level of protection. We see no point in going through the mechanics of a rigorous statistical SSD modelling exercise only to subjectively modify the results.

While we readily acknowledge the rightful place of subjectivity in the form of *expertise*, *knowledge*, and *scientific understanding,* what we are objecting to here is fundamentally different. To be clear, the assessment factor approach is different to a Bayesian analysis whereby the subjective component is incorporated and handled in a logical, transparent, and statistically credible manner. The scaling of an SSD-derived GV by an arbitrary constant (typically a number between 2 and 10) undermines the statistical underpinnings of the SSD methodology. The quantification of the AF to be applied in any given situation is based on a *subjective* evaluation of either the uncertainties around the derivation of the HC_5 (EC 2011) or "the residual uncertainty of the WQG" (BC 2019). However determined, the rationale behind the AF is that it results in a more stringent GV that imparts a greater (although unquantified) level of protection to the ecosystem. This is deemed necessary to account for the many sources of uncertainty in the SSD modelling process. But this implicitly assumes that the net effect of these uncertainties is to inflate the HC_5 which we believe is unverifiable in any given instance. We have seen SSD analyses based on either poor data, poor modelling, or a combination of both, that resulted in ultra-conservative HCp values. To further reduce these by a factor of 2-10 would be both unrealistic and unjustifiable.

The issue of whether to accommodate the uncertainty in an SSD-derived HCp used to establish default guideline values (DGVs) in Australia and New Zealand was contemplated more than 20 years ago during the preparation of the Australian and New Zealand Water Quality Guidelines (ANZECC/ARMCANZ 2000). At the time it was suggested that the lower limit of a suitably chosen confidence interval on the HCp estimate be used for this purpose. It was quickly recognised that, at the time, this proposition was unworkable in practice for the following reasons: (i) there was no clear basis for choosing the level of confidence; (ii) even for relatively low levels of confidence, the lower concentration bound was often close to or below naturally occurring concentrations (and sometimes even *negative*); and (iii) the resulting metrics (e.g. a 95:90 'trigger-value' where the first number is the level of protection and the second number is the level of confidence) were confusing for many people. The Guidelines document (ANZECC/ARMCANZ 2000) also contemplated an alternative approach using beta-content tolerance intervals (Fox 2000), although this was not vigorously pursued. However, given the persistent use of assessment factors, further investigation into the use of tolerance intervals as an alternative to subjective assessment factors is possibly warranted. Tolerance intervals are subtly different to confidence intervals. The focus of the latter is an unknown population parameter (such as the true mean or true variance), whereas a tolerance interval is a probability statement about the fraction of a population contained in some (random) interval.

The approach reflected in the most recent Guidelines document (ANZG 2018) is to use an SSD model to obtain a point estimate of the HCp together with its standard error. This latter quantity is a direct measure of the precision of the estimate which is influenced by stochastic variation in toxicity data which in turn affects the quality of the fitted SSD model and its predictions. Further advances in SSD modelling have indicated how uncertainty in the toxicity data can be explicitly incorporated into the SSD modelling framework (Fox 2010) as well as accommodating model uncertainty (Fox et al. 2021, Thorley and Schwarz 2018).

It has been argued that in data-poor environments where the sample size is too small to meaningfully fit an SSD, a predicted no effect concentration (PNEC) be obtained "by dividing the lowest toxicity value in the substance's dataset by a certain assessment factor" (Okonski et al. 2021). We believe this is an inappropriate approach to environmental protection for the following reasons:

- It is based on a *single* observation (the smallest value) and discards the rest of the data.
- This single data value is entirely dependent on the (subjective) choice of the most sensitive species in the concentration-response experiments.

- Extreme values (such as the minimum and maximum) are notoriously unreliable due to their large sampling variability.
- This highly variable, single data point is then scaled by an arbitrary amount.
- As an estimate of a NEC, the inferred level of protection is 100%, however there is no way of knowing what level of protection is afforded by scaling the minimum concentration "by a certain assessment factor".

With a sample size of effectively *n = 1*, the only plausible option is to obtain more data. The resulting often ultra-conservative GVs are driven by the needs of regulatory agencies for a number that they can apply, but because the GVs in such cases will frequently be exceeded, they will almost always be a trigger for further studies to confirm the absence of effects at the measured environmental concentrations. In Australia and New Zealand, such low reliability GVs were termed Environmental Concern Levels (ECLs) (ANZECC/ARMCANZ 2000) and suggested as only interim working levels. An example of their conservatism is the GV for aluminium in marine waters where the ECL was 0.5 µg/L compared to a more recently derived high reliability value of 24 µg/L based on 11 data points in an SSD representing 6 taxonomic groups (Golding et al. 2015). The problems such disparate values can cause industry are obvious. Equally there may be a cost if, as in some jurisdictions, the regulator needs to demonstrate that there is an impact. Published PNECs derived in the same manner can be equally over-protective depending on the magnitude of the applied AFs.

Clearly, what is required are methods and tools that reduce the subjectivity associated with SSD modelling rather than increasing it. Recent advances in the use of statistical model-averaging may provide one such approach.

Advances in the SSD methodology

In 2011, the European advice was to apply log-logistic fits for the SSD, although the use of other statistical approaches was acknowledged (EC 2011). The latest advice is that 'the choice of a distribution function other than log-normal or log-logistic should be clearly explained' (EC, 2018). In 2000, Australia and New Zealand pioneered the application of a Burr type III distribution using their Burrlioz software (ANZECC/ARMCANZ, 2000; Campbell et al. 2000), updated in 2014 (Barry and Henderson 2014).

A recently completed 2.5-year study by Australian and Canadian researchers undertook a comprehensive assessment of the statistical underpinnings of SSD modelling with particular emphasis on the use of Burr III distributions in the Burrlioz software (Fox et al. 2022) and the model-averaging approach used in the R package `ssdtools` (Thorley and Schwarz 2018). A major

recommendation arising from this work was that both jurisdictions use model-averaged SSDs and the `ssdtools` software for all future water quality GV derivations (Fox et al. 2022).

Model-averaged SSDs

The strength of model-averaged SSDs is that they (i) remove the need to pick a single 'best' probability model and (ii) the importance of any individual distribution is based entirely on information-theoretic metrics and not on subjective assessment. This latter feature combined with the quantification of uncertainties and the provision of confidence intervals negates the need to further 'adjust' estimated HCp values using arbitrary AFs. By way of example, we have reanalysed the linear alkylbenzene sulfonate (LAS) data supplied as supplemental information by Belanger and Carr (2019) and shown in our Supplementary Information Table S1.

The `ssdtools` package (Thorley and Schwarz 2018) uses eight, 1-component distributions as well as two, 2-component *mixture distributions* (Fox et al. 2021). A plot of the empirical cumulative distribution function (*cdf*) and fitted distributions from `ssdtools` for the LAS data is shown in Figure 1. It is evident that the single-component distributions display a wide variety of left-tail behaviours while the two 2-component mixture distributions have abrupt left-tails that are no doubt driven by the unique ability of these distributions to model the atypical toxicity values of the smallest three data points. Whether a long tail or a short tail is appropriate for these data represents a major source of uncertainty which the proponents of AFs handle by fitting a single (1-component) distribution and scaling the resultant HCp by an arbitrary amount. We are perplexed as to how one meaningfully determines a value to adjust for this model uncertainty. In contrast, the `ssdtools` model output provides the quantitative information required (Table 1) (discussed in detail by Fox et al. 2021). By examining the weight column in Table 1, we immediately see that the inverse Pareto distribution is not supported at all while the two, 2-component distributions afford the best representation of these data (in reality, only one of these distributions is required since the fits are almost identical).

For the LAS data, the model-averaged HC_5 is estimated to be 0.223 mg/L with a standard error of 0.091 mg/L and a 95% confidence interval ranging from 0.148 mg/L to 0.497 mg/L. Individual distributions provided HC_5 estimates ranging from 0.0096 mg/L for the Weibull to 0.255 mg/L for the log-Gumbel. The log-logistic estimated HC_5 is 0.208 mg/L and applying an AF of 5 gives a GV of 0.042 mg/L. We estimate that at this low concentration, the level of species protection is 99.7% and not the assumed 95%.

The reliability of any SSD-derived GV will be critically dependent on the quality of the input data. In Australia and New Zealand, this requires the toxicity testing to pass established QA/QC criteria, having sufficient data to meet the criterion for a 'high reliability" GV, and considering the possibility of bimodality in the dataset (Warne et al. 2018).

In the next section, we more fully investigate the impact of arbitrarily scaling an HC_5 obtained from SSDs fitted to benchmark toxicity data sets in the recently curated R package `ssddata` (Fisher and Thorley 2021).

The `ssddata` toxicity datasets

A key recommendation in Fox et al. (2021) was the establishment of a readily accessible collection of toxicity data sets that displayed a variety of distributional shapes and tail behaviours that could be used for the testing and evaluation of statistical methodologies. The R package `ssddata` (Fisher and Thorley 2021) was created for that purpose and is available on github at https://github.com/open-aims/ssddata and CRAN https://cran.r-project.org/src/contrib/ssddata_1.0.0.tar.gz . The package includes a range of datasets sourced from the Canadian Council of Ministers of the Environment (CCME), the Australian Institute of Marine Science (AIMS), the Commonwealth Scientific and Industrial Research Organisation (CSIRO), and the Australian and New Zealand water quality guidelines website (ANZG 2018), as well as anonymous datasets supplied by the Department of Agriculture Water and the Environment (DAWE) and other parties. Also available in the `ssddata` package is a dataset containing various software fits to the `ssddata` toxicity data that can be used for comparison purposes (see https://open-aims.github.io/ssddata/reference/ssd_fits.html). Table 2 provides a complete list of the contents of the package.

Figures 2 and 3 show respectively, the histograms and empirical *cdfs* for each of the 25 data sets in the ssddata package. HCp estimates for p = {1,5,10 and 20} were obtained for each of the 25 datasets in `ssddata` using the `ssd_fits_bcanz` function in the R package `ssdtools` (Thorley and Schwarz 2018). `ssd_fits_bcanz` returns model-averaged results based on the default set of distributions adopted by both the Australian/New Zealand and Canadian jurisdictions. After dividing these HCp estimates by a range of AFs {1, 2, 4, 8, and 16}, the fitted SSD was used to obtain estimates of the fraction of species protected at each of the common protection levels {80%, 90%, 95%, and 99%}. The results are summarised in Figure 4 based on the averages for each of the 25 distributions for each of the applied AFs. As expected, this shows that the effect of applying an AF is to increase the level of protection – which is the whole point of an AF. Although the profile lines of Figure 4 indicate a predictable trend in the relationship between *actual* and *assumed* levels of protection, we caution against an attempt to quantify these based on this limited study. Clearly the

greatest impact of AFs is at the lower levels of assumed protection. This simply reflects the increased 'headroom' at these levels, i.e., the AF can increase the assumed level of protection by 20 percentage points at the 80% assumed protection level, whereas at the 99% level of protection, this increase is limited to 1 percentage point. Interestingly, the *slope* of the profile lines in Figure 4 *decreases* as the AF increases to the point where for an AF=16, the fraction protected is almost a constant ~98%. Higher AFs are intended to compensate for higher levels of uncertainty and/or data quality/paucity issues. Our results suggest that in such cases, the level of protection afforded by an AF-adjusted HCp is independent of p which effectively negates the use of HCps.

Conclusions

To derive default water quality GVs for toxicants, we have shown that, provided that the number of datasets meets the minimum requirements (ANZG 2018), the application of a state-of-the-art model-averaging software such as `ssdtools` will yield the most statistically defensible HC_5 values. The use of assessment factors designed to account for 'uncertainties' offers no demonstrable advantages and may indeed lower values below natural environmental concentrations or analytical detection limits. This places requirements for unnecessary additional expensive experimentation on the 'polluter' to demonstrate a lack of impact to the regulator.

ACKNOWLEDGEMENTS

We are grateful to our colleagues Dr. Jenny Stauber (CSIRO) and Dr. Lisa Golding (CSIRO) for their valuable comments on an earlier draft of this paper. We would also like to acknowledge the anonymous reviewers whose feedback helped clarify and improve the final version of this paper.

REFERENCES

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

ANZG (2018). Australian and New Zealand Guidelines for Fresh and Marine Water Quality. Australian and New Zealand Governments and Australian state and territory governments, Canberra ACT, Australia. Available at: www.waterquality.gov.au/anz-guidelines

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

BC (2019). Derivation of water quality guidelines for the protection of aquatic life in British Columbia. Water Quality Guideline Series, WQG-06. British Columbia Ministry of Environment and Climate Change Strategy, Victoria B.C., Canada.

Belanger SE, Carr GJ. (2019). SSDs revisited: Part II—Practical considerations in the development and use of application factors applied to species sensitivity distributions. *Environmental Toxicology and Chemistry,* **38,** 1526–1541.

Campbell E, Palmer M, Shao Q, Warne M, Wilson D. 2000. Burrlioz: A flexible approach to species protection. Proceedings, TIES/SPRUCE 2000 Conference, 4–8 September 2000, University of Sheffield, Sheffield, UK.

CCME (2007). Protocol for the Derivation of Water Quality Guidelines for the Protection of Aquatic Life. Winnipeg, Manitoba: Canadian Council of Ministers of the Environment. Available online at: http://ceqg-rcqe.ccme.ca/download/en/220

EC (2011). Common Implementation Strategy for the Water Framework Directive (2000/60/EC). Guidance Document No. 27. Technical Guidance for Deriving Environmental Quality Standards. Technical Report-2011-055, European Commission.

EC (2018). Technical guidance for deriving environmental quality standards. Guidance Document No. 27. Updated version. European Commission. Available from: https://rvs.rivm.nl/sites/default/files/2019-04/Guidance%20No%2027%20-%20Deriving%20Environmental%20Quality%20Standards%20-%20version%202018.pdf.

ECETOC (2014). Workshop Report 28 – Estimating toxicity thresholds for aquatic ecological communities from sensitivity distributions. European Centre for Ecotoxicology and Toxicology of Chemicals, Brussels, Belgium. Available at: https://www.ecetoc.org/wp-content/uploads/2021/10/ECETOC_WR_28.pdf

Fisher, R Thorley, J. (2021) ssddata: Species Sensitivity Distribution Data. https://cran.r-project.org/package=ssddata

Fox DR (2016). Contemporary methods for statistical design and analysis. In Marine Ecotoxicology: Current Knowledge and Future Issues' (Eds J Blasco, P Chapman, O Campana, M Hampel) pp 35–70. (Elsevier, San Diego, CA, USA).

Fox DR, van Dam RA, Fisher R, Batley GE, Tillmanns AR, Thorley J, Schwarz CJ, Spry DJ, McTavish K (2021). Recent developments in ssd modelling. *Environmental Toxicology and Chemistry*, **40,** 293-308.

Fox DR, Fisher R, Thorley, JL, Schwarz, C (2022). Joint investigation into statistical methods underpinning the derivation of toxicant guideline values in Australia and New Zealand. Report prepared for the Department of Agriculture, Water and the Environment. (Environmetrics Australia, Beaumaris, Vic. and the Australian Institute of Marine Science, Perth, WA). Available from https://bit.ly/3l4H0Av

Fox DR. (2006). Statistical Issues in Ecological Risk Assessment. *Human and Ecological Risk Assessment*, 12, 120-129.

Golding LA, Angel BM, Batley GE, Apte SC, Krassoi R, Doyle CJ (2015). Derivation of a water quality guideline for aluminium in marine waters. *Environmental Toxicology and Chemistry,* **34,** 141–151.

Stephan CE. (2002). Use of species sensitivity distributions in the derivation of water quality criteria by the U.S. Environmental Protection Agency. In Posthuma L, Suter GW, Traas TP, eds, Species Sensitivity Distributions in Ecotoxicology. CRC, Boca Raton, FL, USA, pp 211–220.

Thorley J, Schwarz C (2018). ssdtools: An R package to fit Species Sensitivity Distributions. *Journal of Open Source Software*, **3,** 1082. doi: 10.21105/joss.01082.

Warne M, Batley GE, van Dam RA, Chapman JC, Fox DR, Hickey CW, Stauber JL (2018). Revised Method for Deriving Australian and New Zealand Water Quality Guideline Values for Toxicants – update of 2015 version. Prepared for the revision of the Australian and New Zealand Guidelines for Fresh and Marine Water Quality. 48 pp. (Australian and New Zealand Governments and Australian state and territory governments, Canberra, ACT, Australia).

Conflict of Interest

Graeme Batley is an Editor for Environmental Chemistry and was blinded from the peer review process for this manuscript

Declaration of Funding

This research did not receive any specific funding

Data availability statement

This paper contained no new or unpublished data

Table 1. Goodness-of-fit summary statistics from ssdtools for the fitted distributions for LAS shown in Figure 1.

Distribution	ad	ks	cvm	aic	aicc	bic	delta	Weight
burrIII3	0.321	0.122	0.041	76	77.6	78.9	5.2	0.023
gamma	0.454	0.139	0.049	76.1	76.8	78.0	4.41	0.035
gompertz	0.464	0.139	0.051	76.1	76.8	78.0	4.41	0.034
invpareto	2.77	0.383	0.569	89.3	90.0	91.2	17.6	0
lgumbel	0.591	0.14	0.089	74.9	75.6	76.8	3.21	0.063
llogis	0.341	0.126	0.044	74.1	74.8	75.9	2.38	0.095
llogis_llogis	0.195	0.112	0.028	68.3	72.9	73	0.45	0.25
lnorm	0.342	0.123	0.043	73.2	73.9	75.1	1.49	0.148
lnorm_lnorm	0.2	0.109	0.025	67.8	72.4	72.5	0	0.313
weibull	0.41	0.138	0.042	75.9	76.6	77.8	4.2	0.038

ad= Anderson-Darling statistic; ks= Kolmogorov-Smirnov test statistic; cvm= Cramer-von Mises statistic; aic= Akaike's information criterion; aicc= Akaike's information criterion corrected for sample size; bic= Bayesian information criterion; delta=aicc-min{aicc}; weight is a measure of the support for the distribution and is a function of delta (see Fox et al. 2021).

322

323 Table 2. Toxicity data sets available in the R package ssddata

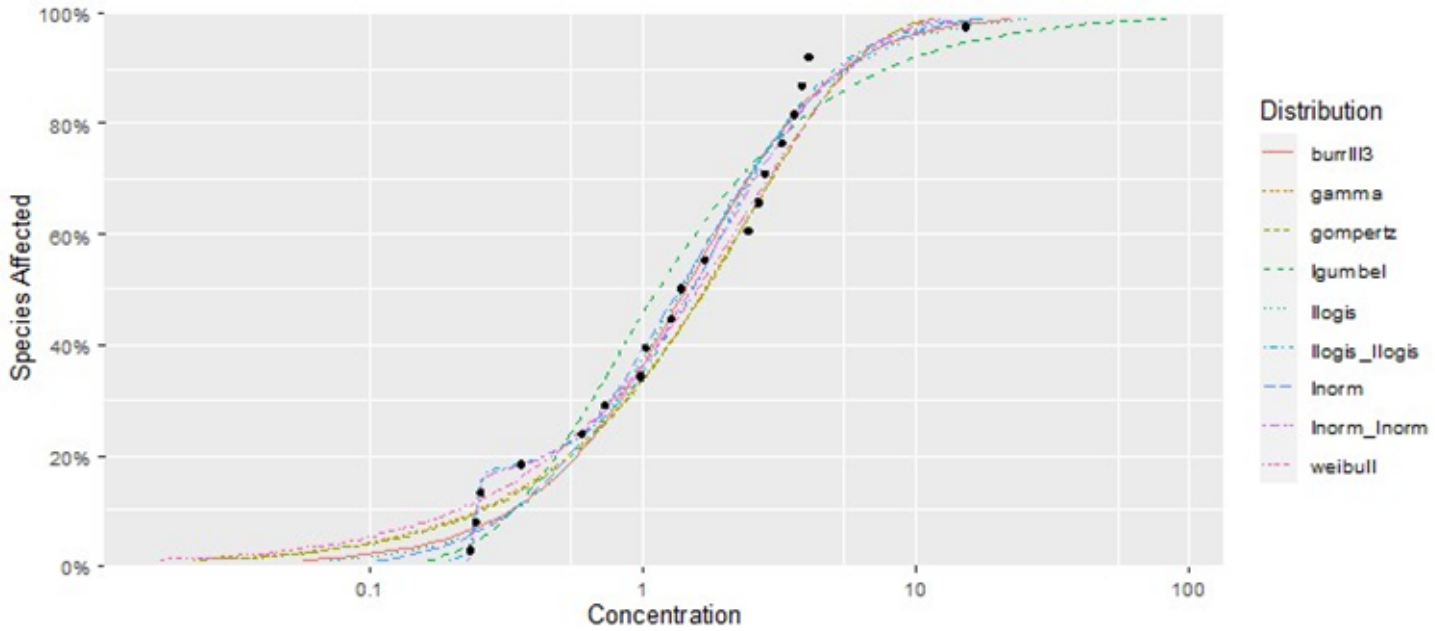
Dataframe name	Description
aims_aluminium_marine	Species Sensitivity Data for aluminium_marine
aims_data	Species Sensitivity Data provided by AIMS
aims_gallium_marine	Species Sensitivity Data for gallium_marine
aims_molybdenum_marine	Species Sensitivity Data for molybdenum_marine
anon_a	Anonymous Species Sensitivity Data anon_a
anon_b	Anonymous Species Sensitivity Data anon_b
anon_c	Anonymous Species Sensitivity Data anon_c
anon_d	Anonymous Species Sensitivity Data anon_d
anon_data	Anonymous Species Sensitivity Data
anon_e	Anonymous Species Sensitivity Data anon_e
anzg_data	ANZG Species Sensitivity Data
anzg_metolachlor_fresh	Species Sensitivity Data for metolachlor_fresh
ccme_boron	CCME Species Sensitivity Data for ccme_boron
ccme_cadmium	CCME Species Sensitivity Data for ccme_cadmium
ccme_chloride	CCME Species Sensitivity Data for ccme_chloride
ccme_data	CCME Species Sensitivity Data
ccme_endosulfan	CCME Species Sensitivity Data for ccme_endosulfan
ccme_glyphosate	CCME Species Sensitivity Data for ccme_glyphosate
ccme_silver	CCME Species Sensitivity Data for ccme_silver
ccme_uranium	CCME Species Sensitivity Data for ccme_uranium
csiro_chlorine_marine	Species Sensitivity Data for chlorine_marine
csiro_cobalt_marine	Species Sensitivity Data for cobalt_marine
csiro_data	Species Sensitivity Data provided by CSIRO
csiro_lead_marine	Species Sensitivity Data for lead_marine
csiro_nickel_fresh	Species Sensitivity Data for nickel_fresh
ssd_fits	Species Sensitivity Distribution Fit Data

324

325

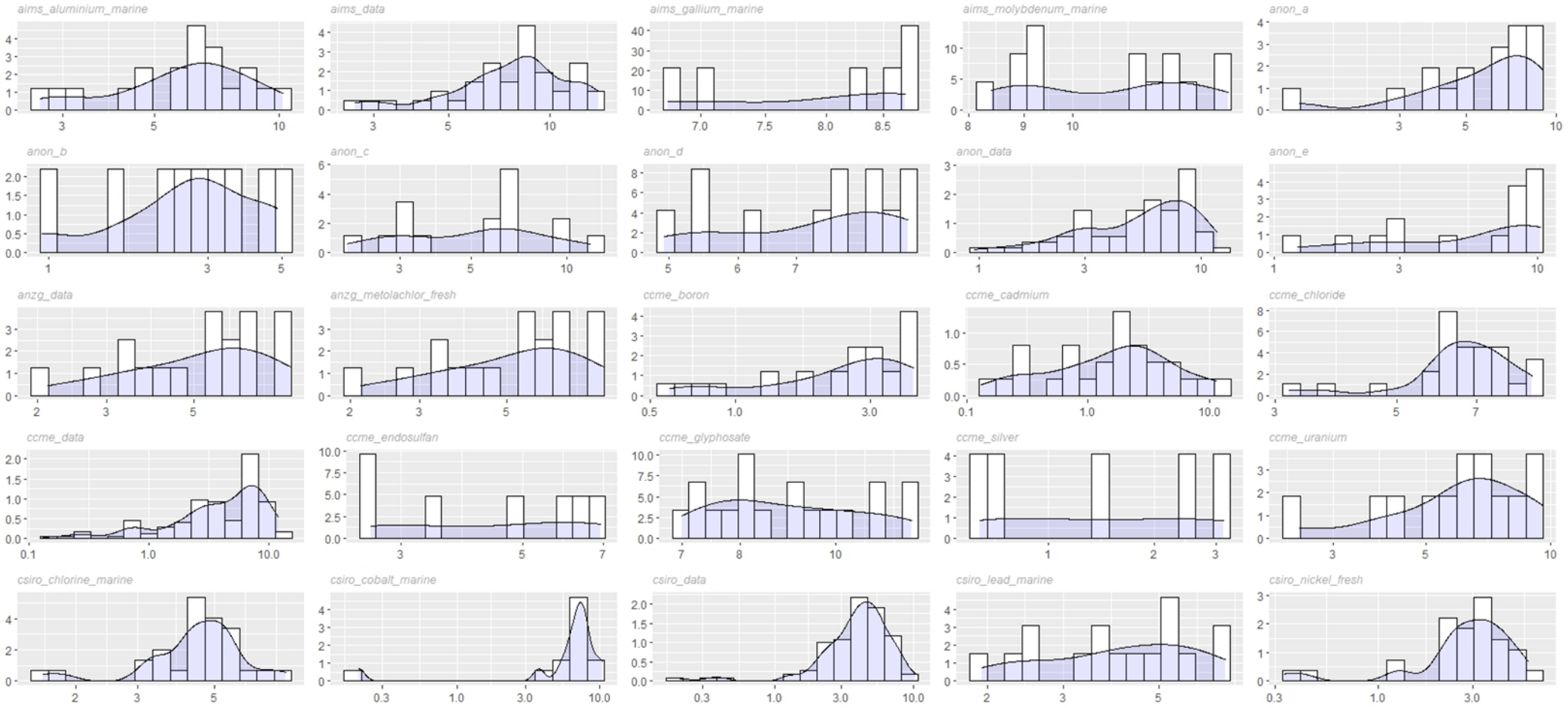
326

327



328

Figure 1. Values of the empirical cumulative distribution function (solid points) and fitted distributions from the ssdtools package for LAS toxicity data in Table S1

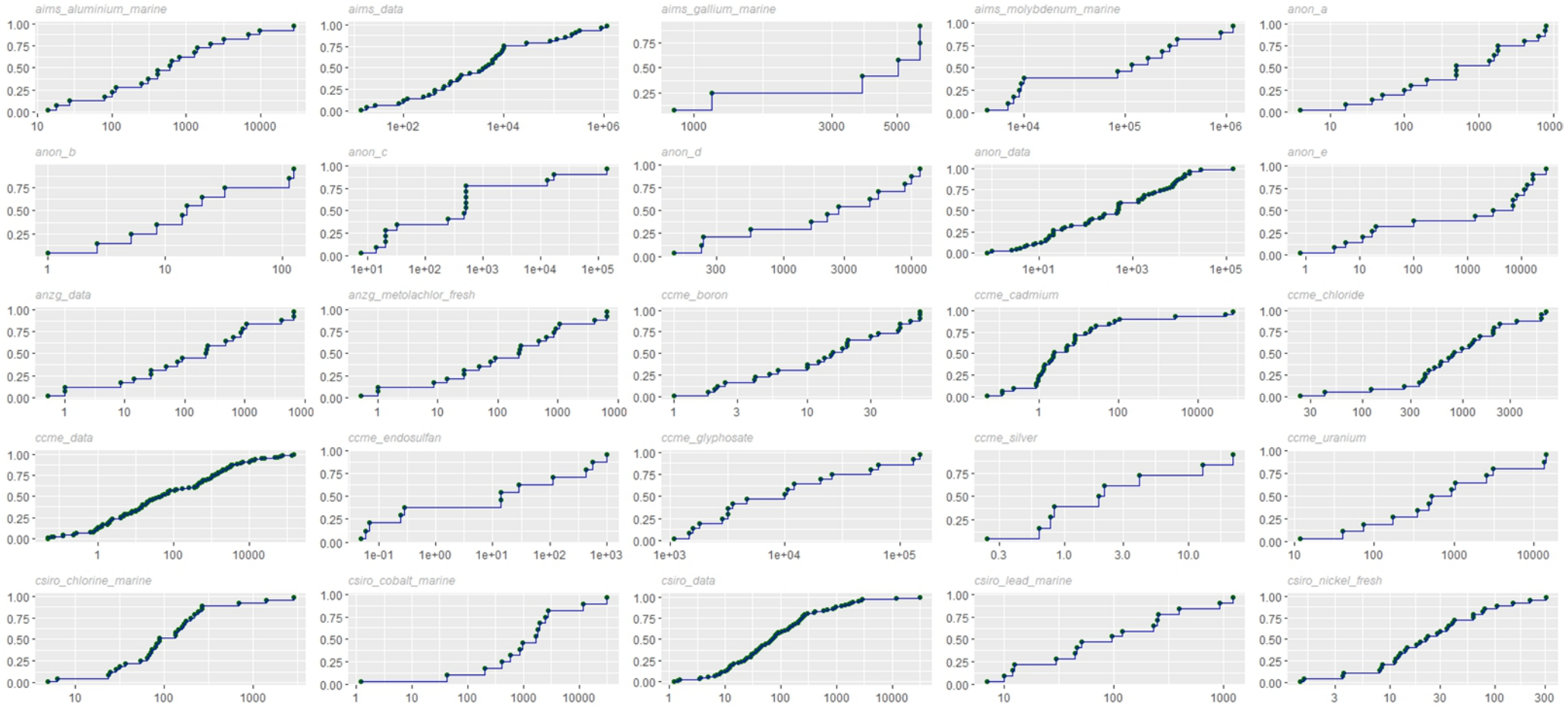


329

330 Figure 2. Histograms with density smooth overlay for 25 toxicity data sets in ssddata package. Vertical scale is probability density; horizontal scale is
331 log(concentration).

332

333

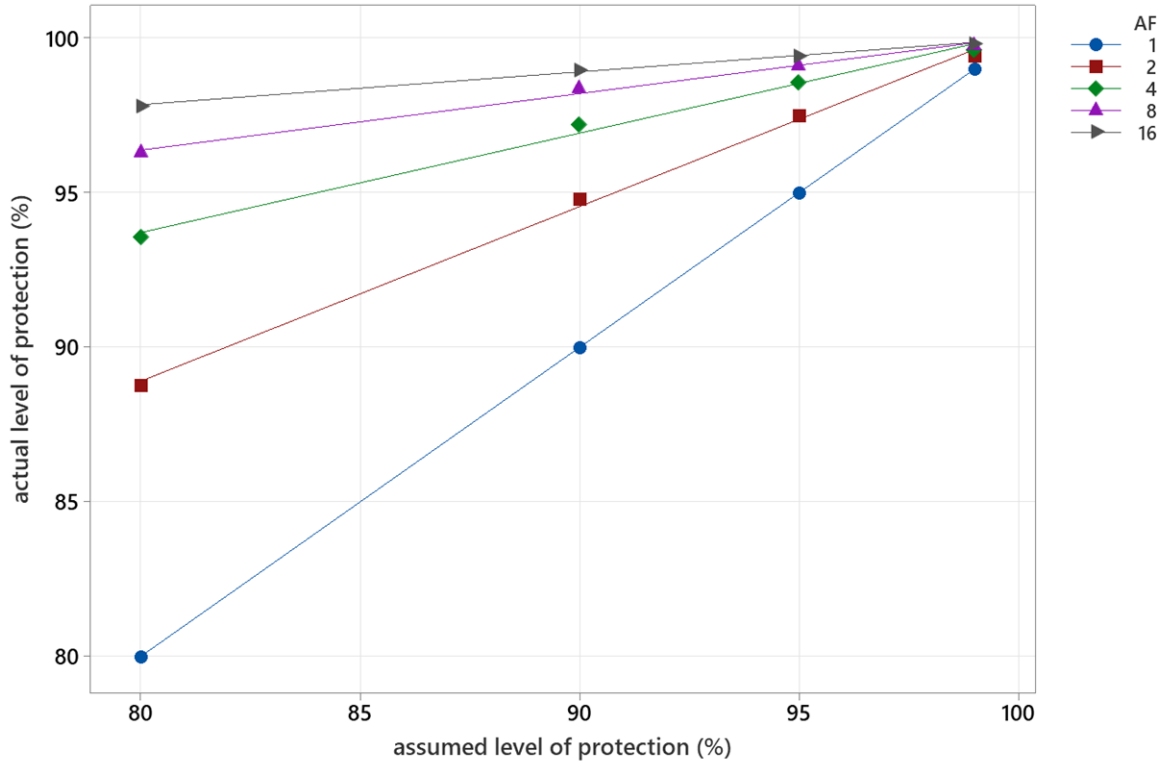


334

335 Figure 3. Empirical cdfs for 25 toxicity data sets in ssddata package. Vertical scale is fraction affected; horizontal scale is log(concentration).

336

337



338

339 Figure 4. Relationship between assumed level of protection and estimated level of protection after
340 the ssd-derived HC is divided by various assessment factors (plotted points are averages taken over
341 all 25 data sets in the ssddata package)

342

343

Supporting Information

Assessment Factors in Species Sensitivity Distributions for the Derivation of Guideline Values for Aquatic Contaminants

David R. Fox[1,2]* and Graeme E. Batley[3]

[1]Environmetrics Australia, Beaumaris, Victoria 3193, Australia
[2]University of Melbourne, Parkville, Victoria 3010, Australia
[3]CSIRO Land and Water, Locked Bag 2007, Kirrawee NSW 2232, Australia

Table S1. Chronic toxicity data for linear alkylbenzene sulfonate (LAS) normalised to a chain length of 11.58 alkyl carbons for use in an SSD[a]

Rank	Taxon	Mean normalized value (mg/L)
1	*Onchorhychus mykiss*	0.2349
2	*Lemna minor*	0.2448
3	*Tilapia mossambica*	0.2554
4	*Corbicula fluminea*	0.3591
5	*Ceriodaphnia dubia*	0.5979
6	*Microcystis aeruginosa*	0.7250
7	*Pimephales promelas*	0.9849
8	*Lepomis macrochirus*	1.0215
9	*Hyallela Azteca*	1.2636
10	*Daphnia magna*	1.3702
11	*Brachionus calyciflorus*	1.6799
12	*Desmodesmus subspicatus*	2.4337
13	*Paratanytarsus parthenogenica*	2.6603
14	*Chironomus riparius*	2.7868
15	*Poecelia reticulata*	3.2407
16	*Chlorella kessleri*	3.5491
17	*Elimia sp*	3.8574
18	*Elodea canadensis*	4.0562
19	*Pseudokirchneriella subcapitata*	15.2717

[a] Data are taken from the Supplementary Information in Belanger and Carr (2019) originally derived from Belanger et al. (2016)

References

Belanger SE, Carr GJ. (2019). SSDs revisited: Part II—Practical considerations in the development and use of application factors applied to species sensitivity distributions. *Environmental Toxicology and Chemistry* **38,** 1526–1541.

361 Belanger SE, Brill JL, Rawlings JM, Price BB (2016). Development of acute toxicity Quantitative
362 Structure Activity Relationships (QSAR) and their use in linear alkylbenzene sulfonate species
363 sensitivity distributions. *Chemosphere* **155**, 18-27.

364

365