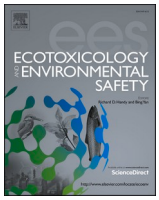
Ecotoxicology and Environmental Safety 278 (2024) 116379

Contents lists available at ScienceDirect

Ecotoxicology and Environmental Safety

journal homepage: www.elsevier.com/locate/ecoenv



Which distribution to choose for deriving a species sensitivity distribution? Implications from analysis of acute and chronic ecotoxicity data

Miina Yanagihara [a,b,*], Kyoshiro Hiki [c], Yuichi Iwasaki [d,**]

[a] KWR Water Research Institute, Groningenhaven 7, Nieuwegein 3433 PE, the Netherlands
[b] Center for Marine Environmental Studies, Ehime University Bunkyo-cho 3, Matsuyama, Ehime 790-8577, Japan
[c] Health and Environmental Risk Division, National Institute for Environmental Studies, 16-2 Onogawa, Tsukuba, Ibaraki 305-8506, Japan
[d] Research Institute of Science for Safety and Sustainability, National Institute of Advanced Industrial Science and Technology (AIST), 16-1 Onogawa, Tsukuba, Ibaraki 305-8569, Japan

ARTICLE INFO

Edited by Mohamed Abdel-Daim

Keywords:
Ecological risk assessment
Hazard assessment
Model selection
Statistical distribution

ABSTRACT

Species sensitivity distributions (SSDs) estimated by fitting a statistical distribution to ecotoxicity data are indispensable tools used to derive the hazardous concentration for 5 % of species (HC5) and thereby a predicted no-effect concentration in environmental risk assessment. Whereas various statistical distributions are available for SSD estimation, the fundamental question of which statistical distribution should be used has received limited systematic analysis. We aimed to address this knowledge gap by applying four frequently used statistical distributions (log-normal, log-logistic, Burr type III, and Weibull distributions) to acute and chronic SSD estimation using aquatic toxicity data for 191 and 31 chemicals, respectively. Based on the differences in the corrected Akaike's information criterion (AICc) as well as visual inspection of the fitting of the lower tails of SSD curves, the log-normal SSD was generally better or equally good for the majority of chemicals examined. Together with the fact that the ratios of HC5 values of other alternative SSDs to those of log-normal SSDs generally fell within the range 0.1–10, our findings indicate that the log-normal distribution can be a reasonable first candidate for SSD derivation, which does not contest the existing widespread use of log-normal SSDs.

1. Introduction

Chemical pollution is a global driver that causes adverse environmental impacts as well as human health problems (Johnson et al., 2020; Persson et al., 2022; Steffen et al., 2015). Indeed, Persson et al. (2022) have reasoned that the safe operating space that lies within the planetary boundary of chemical pollution, highlighted by plastic pollution, has likely been exceeded. Conducting ecological risk assessments, including hazard and exposure assessments, is a fundamental step towards addressing chemical pollution.

Species sensitivity distributions (SSDs) are crucial tools for deriving the hazardous concentration for 5 % of species (HC5) that is fundamental to predict environmentally "safe" concentrations in ecological risk assessments (Posthuma et al., 2019, 2002). SSDs are statistical distributions that express the different sensitivities (toxicity values) of species to chemicals, typically obtained by fitting a statistical distribution to data from ecotoxicity tests. For example, the HC5 estimated from an SSD based on toxicity metrics such as no observed effect concentrations (NOECs) and 10 % effect concentrations is frequently used to derive a predicted no-effect concentration (PNEC) or environmental benchmark values such as water quality standards. Similarly, HC5 values derived from SSDs based on acute toxicity measures, such as 50 % effect concentration (EC50) and median lethal concentration (LC50), are primarily used in pesticide risk assessment (EFSA Panel on Plant Protection Products and their Residues (PPR), 2013). Historically, the choice of the statistical distribution for an SSD has been based on practicality rather than scientifically defensible evidence. Several distributions are commonly used and are recommended in extant guidelines. For example, the log-normal distribution is considered a pragmatic choice and is recommended in several guidelines (European Commission., 2011; Wheeler et al., 2002). Because the mathematical properties of the normal distribution have been extensively investigated

* Corresponding author at: KWR Water Research Institute, Groningenhaven 7, Nieuwegein 3433 PE, the Netherlands.
** Corresponding author.
E-mail addresses: Miina.Yanagihara@kwrwater.nl, yanagihara.miina.lk@ehime-u.ac.jp (M. Yanagihara), hiki.kyoshiro@nies.go.jp (K. Hiki), yuichi-iwasaki@aist.go.jp, yuichiwsk@gmail.com (Y. Iwasaki).

https://doi.org/10.1016/j.ecoenv.2024.116379
Received 29 October 2023; Received in revised form 16 April 2024; Accepted 21 April 2024
Available online 6 May 2024

(Aldenberg and Jaworska, 2000; Kamo et al., 2022), it is possible to derive the magnitude of the assessment factor required to achieve protection of 95 % of species with a probability of 95 % based on the sample size and the variation of sensitivity (Kamo et al., 2022). Similarly, the log-logistic distribution, which has wider tails and employs a more conservative assumption than the log-normal distribution (Aldenberg and Slob, 1993), is commonly used for SSD estimation (Warne et al., 2018; Wheeler et al., 2002). In Australia and New Zealand, the Burr type III distribution, which is the specific case of the generalized Burr distribution (Burr, 1942; Shao, 2000), has been adopted as one of the standard distributions to estimate SSDs (Warne et al., 2018).

Previous studies (Grist et al., 2002; He et al., 2014; Newman et al., 2000; Wheeler et al., 2002; Xu et al., 2015; Zhao and Chen, 2016) have compared the consequences of choosing a particular statistical distribution for an SSD by using several different distributions as well as bootstrap methods to estimate SSDs. However, most of the studies failed to incorporate distributions other than log-normal and log-logistic distributions, and comparisons using chronic toxicity data have been limited. For example, Newman et al. (2000) used NOECs to estimate SSDs for eight of 23 chemicals, but other studies (Grist et al., 2002; He et al., 2014; Wheeler et al., 2002; Xu et al., 2015; Zhao and Chen, 2016) relied on acute toxicity data, such as EC50 and LC50 for a maximum of 32 chemicals. Furthermore, the relationship between the choice of distribution and HC5 remains unclear because there was a lack of consistency in the choice of the distributions or methods used to estimate SSDs across the studies (Grist et al., 2002; He et al., 2014; Newman et al., 2000; Wheeler et al., 2002; Xu et al., 2015; Zhao and Chen, 2016). Therefore, a more systematic understanding of how the choice of a distribution affects the derivation of HC5 is crucial. To tackle this challenge, it is essential to employ larger datasets containing a wide range of test species and chemicals, as well as more types of statistical distributions.

As an alternative approach to choosing a single distribution for SSD estimation, model averaging based on information criteria such as Akaike's information criterion (AIC) has been implemented in several software tools (see (Fox et al., 2021)) for more details) and utilized at the national level (Canadian Council of Ministers of the Environment, 2019) and several recent studies (e.g., (Hamoutene et al., 2023; Wang et al., 2023)). While the model-averaging-based SSD estimation is appealing because it skips the selection process of a single distribution to be applied and incorporates the uncertainty about which model is correct (Schwarz and Tillmanns, 2019), it is important to note that model averaging does not necessarily guarantee a reduction in prediction error (see (Dormann et al., 2018) for a more comprehensive discussion). Additionally, the adoption of the model averaging for SSD estimation in many jurisdictions worldwide is not yet common. Therefore, addressing which distribution to choose for the SSD estimation is of continuous value. Findings obtained from such attempts should also provide valuable insights into e.g., the identification of candidate statistical distributions for the model averaging.

In this study, we analyzed both acute and chronic toxicity data for a wide range of chemicals to address the question of which statistical distribution is better suited for the SSD estimation by comparing the relative goodness of models in terms of prediction as well as HC5 values derived from SSDs. We used four different statistical distributions: log-normal, log-logistic, Burr Type III, and Weibull. The acute and chronic ecotoxicity data obtained from a curated database (Connors et al., 2019) were used to examine a wide range of chemicals. The four distributions were selected because they are considered candidate distributions or have been proposed as standard methods in several guidelines by the European Commission, Australia, and New Zealand (European Commission., 2011; Fox et al., 2021; Warne et al., 2018). Because the log-normal distribution is one of the most common distributions used to estimate SSDs (European Commission., 2011; Posthuma et al., 2019; Wheeler et al., 2002), we regarded it as a reference model in this study. Our approach was to compare the predictive goodness and HC5 values of the log-normal SSD with those of SSDs based on the other three distributions.

2. Materials and methods

2.1. Ecotoxicity data compilation

An ecotoxicity dataset was collected from the EnviroTox database version 2.0.0 (https://envirotoxdatabase.org; (Connors et al., 2019)). The EnviroTox is a recently developed ecotoxicity database curated from existing databases, such as USEPA's ECOTOXicology Knowledgebase and OECD QSAR Toolbox (Connors et al., 2019), and it has previously been used in research to compare acute and chronic SSDs (Hiki and Iwasaki, 2020) as well as freshwater and saltwater SSDs (Yanagihara et al., 2022). The dataset obtained from the EnviroTox database included a total of 67,019 test records (55,598 acute and 11,421 chronic) for 1426 species and 4125 chemicals. Chemicals were classified based on Consensus Mode of Action (MoA) categories derived by combining four existing frameworks that focus mainly on fish toxicity: Verhaar framework, Assessment Tool for Evaluating Risk, Toxicity Estimation Software Tool, and OASIS (Connors et al., 2019; Kienzler et al., 2019). The Consensus MoA, hereafter referred to simply as the MoA, classifies chemicals into three categories: narcotic, specifically acting, and unclassified. The MoA was used to explore its possible influences on the comparison of the four different distributions, as examined similarly in previous studies (Hiki and Iwasaki, 2020; Oginah et al., 2023; Yanagihara et al., 2022). Both freshwater and saltwater toxicity data were used without distinction in our analysis because no systematic differences were found between SSDs estimated separately from the two types of data (Yanagihara et al., 2022). Furthermore, a preliminary analysis indicated that the inclusion of elements such as zinc and copper in our SSD analysis did not materially affect our conclusions. However, considering the challenges involved in accurately accounting for the influence of water chemistry parameters such as pH and water hardness on the bioavailability and toxicity of these elements (Adams et al., 2020), they were excluded from our SSD analysis.

To select records and chemicals for estimating SSDs, we used the following criteria, which were slightly modified from previous studies (Hiki and Iwasaki, 2020; Yanagihara et al., 2022): (1) the acute effect measures were EC50 or LC50, and the chronic effect measures were NOEC or the no-observed-effect level (NOEL); (2) the effect concentrations did not exceed five times the water solubility of the tested chemical (Connors et al., 2019); (3) the SSDs were estimated based on toxicity data of 10 or more species from at least three of four taxonomic groups (referred to as trophic groups in ecological risk assessment): algae, invertebrates, amphibians, and fish (Yanagihara et al., 2022); and (4) the bimodality coefficient, which was derived from the sample size, skewness, and excess kurtosis, did not exceed 0.555, a benchmark value suggested by previous studies (Fox et al., 2021; Freeman and Dale, 2013; Pfister et al., 2013). We excluded chemicals with bimodal toxicity data based on this bimodality coefficient to concentrate on our primary goal of examining the model selection in terms of fitting single statistical distributions. This is because chemicals with a certain mode of action are likely to exhibit bi- or multimodal SSDs (Oginah et al., 2023), and such data may require different approaches such as estimating SSDs for separate sensitivity groups (Fox et al., 2021). Based on these criteria, we selected 7176 out of 55,598 acute test records and 745 out of 11,421 chronic test records. These records included results for 268 and 42 test chemicals in acute and chronic toxicity tests, respectively.

2.2. Data analysis

The data analysis was conducted using R version 4.1.3 (R Core Team, 2022) with the R packages "tidyverse" (Wickham et al., 2019) and "ssdtools" (Thorley and Schwarz, 2018). We used "ggplot2" (Wickham, 2016) to visualize data. For SSD estimation, we utilized the

"ssd_fit_dists" function in the "ssdtools" package to fit four distributions (log-normal, log-logistic, Burr type III, and Weibull distributions) using maximum likelihood estimation. If multiple effect concentrations were available for a specific combination of species and a chemical, the geometric mean of these concentrations was calculated and employed in the SSD estimation. Notably, for 77 chemicals with acute data and 11 chemicals with chronic data, the Burr type III distribution could not be fitted because the calculations failed to converge, and these chemicals were removed from further analyses. It is known that estimating the parameters of the Burr type III distribution using maximum likelihood may be susceptible to issues of numerical stability and convergence (Fox et al., 2021). The HC5 values for each SSD model were estimated using the "ssd_hc" function in the "ssdtools" package via parametric bootstrapping with 1000 iterations.

Subsequently, the corrected Akaike's information criterion for small sample sizes (AICc) and the HC5 values of each SSD model were compared with the results from the log-normal SSD. Akaike's information criterion (AIC) is a measure of goodness of fit and model complexity (the number of parameters included) and is used to rank fitted models in terms of prediction (Burnham et al., 2011). In our analysis, the AICc was used for model comparison because the available toxicity data were generally limited for individual chemicals. The difference in AICc values (AICc differences) can be used to determine the relative goodness of models in terms of prediction. As a rule of thumb, models with AICc differences greater than 9–11 or greater than 20 have little or essentially no empirical support from the data, respectively (Burnham et al., 2011). Here, to compare the log-normal SSD with SSDs based on other distributions for each chemical, we calculated AICc differences by subtracting the AICc value of the log-normal SSD from those of other SSDs. In addition, according to Burnham et al. (2011), we used absolute AICc differences of >10 and >20 to operationally evaluate the relative importance of models. We also compared HC5 ratios, calculated by dividing the HC5 values of the alternative SSDs by that of the log-normal SSD. These HC5 comparisons were carried out for SSDs based on all the four distributions, regardless of the resulting AICc values (i.e., predictive accuracy), to examine the deviations from the log-normal SSDs. In addition, we calculated the kurtosis and skewness of toxicity data using the "kurtosis" and "skewness" functions in the "e1071" package (Meyer et al., 2023). The R code for the SSD modeling, as well as an example dataset, can be found on the GitHub website at https://github.com/miinay/SSD_distribution_comparison.

3. Results and discussion

3.1. Overview of estimated SSDs

We estimated acute and chronic SSDs for 191 and 31 chemicals, respectively. Table 1 shows the summary information on the chemicals and the number of species per SSD (see Supplementary Material for more detailed information). Based on the MoA (Connors et al., 2019; Kienzler et al., 2019), the 77, 42, and 72 chemicals for acute SSDs were classified into narcotic, specifically acting, and unclassified groups, respectively. Among the chemicals for chronic SSDs, 10, 9, and 12 chemicals were classified as narcotic, specifically acting, and unclassified groups, respectively. As expected, toxicity data for more species were available for the acute SSDs than the chronic SSDs (Table 1). Toxicity data used for the estimation of acute SSDs included three and four taxonomic groups for 142 and 49 chemicals, respectively. For all 31 chronic SSDs, the toxicity data were available for only three taxonomic groups: algae, invertebrates, and fish. More specifically, for acute toxicity data, the average compositions of the taxonomic groups in each SSD (algae, invertebrates, fish, and amphibians) were 16.8 %, 36.1 %, 45.0 %, and 2.0 %, respectively. For chronic toxicity data, the compositions were 46.7 %, 30.8 %, 22.5 %, and 0 %, respectively.

Table 1
Summary information about the number of chemicals examined and number of species in the species sensitivity distributions (SSDs) used in the present study (mean ± standard deviation; minimum–maximum).

		Total	Mode of action		
			Narcotic	Specifically acting	Unclassified
Acute SSD	Number of chemicals	191	77	42	72
	Number of species per SSD	30 ± 32 (10–215)	24 ± 30 (10–215)	43 ± 42 (10–199)	29 ± 23 (10–154)
Chronic SSD	Number of chemicals	31	10	9	12
	Number of species per SSD	18 ± 8 (10–51)	15 ± 5 (10–23)	21 ± 12 (10–51)	19 ± 7 (10–31)

3.2. Comparisons of the models obtained with the four distributions

Comparisons between the log-normal and three other SSDs (i.e., log-logistic, Burr type III, and Weibull distributions) for 191 and 31 chemicals based on acute and chronic toxicity data, respectively, revealed a limited number of cases of large AICc differences (Fig. 1). No specific MoA was related to the larger AICc differences (Fig. 1). The absolute values of the AICc differences were smaller than 10 for 542 (95 %) and 92 (99 %) paired comparisons of acute and chronic SSDs, respectively. The numbers of cases wherein the absolute value of the AICc difference exceeded 10 or 20 were 31 and 6 among a total of 573 paired comparisons of acute SSDs, and 1 and 0 among a total of 93 paired comparisons of chronic SSDs, respectively. These results showed that, in most cases, it was not straightforward to choose the single best statistical distribution to estimate an SSD for prediction purposes based on the AICc value.

Among the cases where the absolute AICc differences exceeded 10 (31 cases for acute SSDs and one case for a chronic SSD), 23 cases exhibited AICc differences >10 and the remaining nine cases showed AICc differences below -10. In the nine cases of AICc differences below -10, the HC5 differences between the log-normal SSDs and the other SSDs were smaller than a factor of 10. The implication was that the use of log-normal SSDs did not result in large HC5 differences, despite not being selected as the best SSD based on the AICc (see the next section for a discussion of HC5 ratios). These nine cases were derived from five chemicals (Figure S1 and Table S1), and the distributions that showed smaller AICc values than the log-normal distribution (i.e., the distributions that were better supported by the AICc) varied between chemicals. Among the five chemicals, the Shapiro-Wilk test rejected the assumption of normality for the $\log_{10}$-transformed toxicity data of three chemicals: pentachlorophenol, 3,4-dichloroaniline, and 2,4-dichlorophenol. This result at least indicates that deviations from normality, as assessed by the Shapiro-Wilk test, can help identify cases where the use of the log-normal SSDs is less suitable than SSDs based on other distributions from the AICc perspective.

The AICc differences greater than 10 occurred mainly in the comparisons between log-normal and Weibull SSDs (19 out of 22 cases for acute SSDs and one case for chronic SSD), indicating stronger support for the log-normal SSD in these cases. Because we found some characteristic deviations at the tails of the distributions (Figure S2), we compared the skewness and kurtosis of the toxicity data with AICc differences (Figure S3). Chemicals with positively skewed data—longer tails at higher toxicity values—tended to have larger AICc differences for the Weibull distribution. In contrast, chemicals with negatively skewed data did not have larger AICc differences, suggesting that the Weibull distribution could be fitted satisfactorily to negatively skewed but not to positively skewed data. These observations were consistent across both acute and chronic toxicity data (Figure S3).

By simply focusing on the best model with the smallest AICc value, the log-normal, log-logistic, Burr type III, and Weibull distributions were

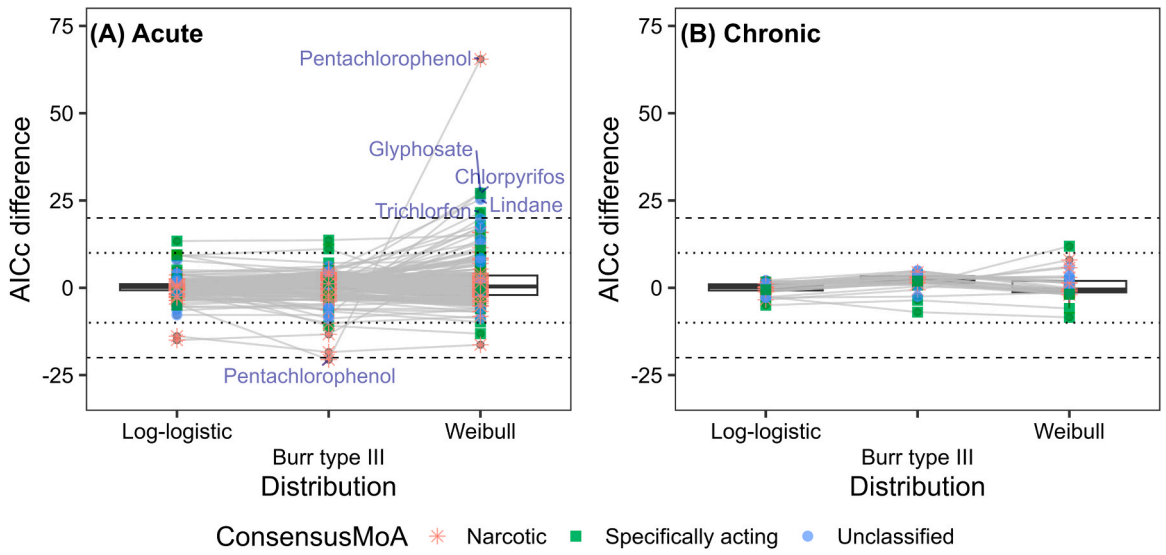


Fig. 1. Comparison of the corrected Akaike's Information Criterion (AICc) difference between the log-normal model and the alternative models for acute species sensitivity distributions (SSDs) (A) and chronic SSDs (B). The dotted and dashed lines in each panel indicate AICc differences with absolute values of 10 and 20, respectively. Box plots show the distributions of AICc differences for the three models. The bold horizontal line, box, error bar, and plot indicate the median, interquartile range, 1.5× (interquartile range), and outlier, respectively. The chemical names have been added where the AICc differences were >20 or <−20.

selected as the best model for 33 %, 20 %, 6 %, and 40 % of the 191 chemicals for acute SSDs. For chronic SSDs, the corresponding percentages were 29 %, 23 %, 0 %, and 48 % of the 31 chemicals. The proportions of Weibull SSDs selected as the best models were somewhat higher than those of log-normal SSDs. However, when the Weibull SSD was not selected as the best model (i.e., the sub-optimal model), large differences in AICc (e.g., > 10) between the best models and the Weibull SSD models were more frequently observed (18 % and 12.5 % of the cases where the Weibull SSD was the sub-optimal model in acute and chronic SSDs, respectively), compared to the differences between the best model and the sub-optimal log-normal SSD (4 % and 0 %, similarly as above). Note that the emphasis on only best models can be misleading in cases where the AICc differences between the best model and other sub-optimal models are not substantial (e.g., AICc difference <10). A few studies have demonstrated that SSDs corresponding to a specific distribution (e.g., Burr type III distribution or Gompertz distribution) were most often selected as the best models for multiple chemicals (He et al., 2014; Newman et al., 2000), but making rigorous comparisons is difficult because of variations in the sets of statistical distributions examined and the methodologies used for selecting the "best" model (see Supplementary Material and Table S2).

3.3. Comparisons of HC5 values estimated by the four distributions

Comparison of the ratios of HC5 values of alternative SSDs to those of log-normal SSDs revealed that, for both acute and chronic SSDs, 98 % of

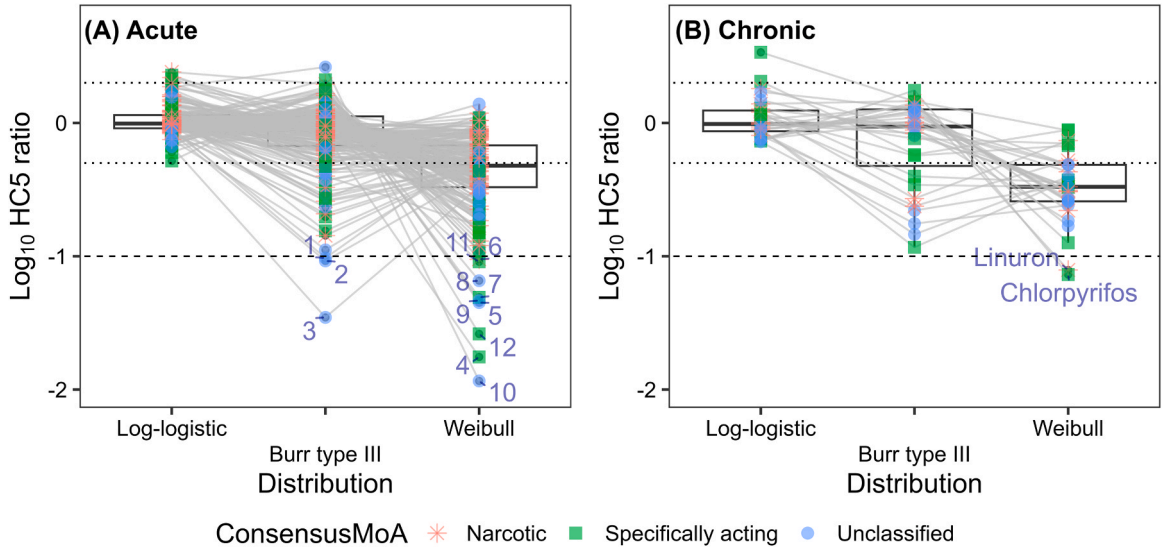


Fig. 2. Comparison of the differences in hazardous concentration for 5% of species (HC5) between the log-normal model and alternative models for acute species sensitivity distributions (SSDs) (A) and chronic SSDs (B). The dotted and dashed lines in each panel show HC5 ratios of two and 10, respectively. Box plots show the distributions of HC5 ratios for the three models. The bold horizontal line, box, error bar, and plot indicate the median, interquartile range, 1.5× (interquartile range), and outlier, respectively. The labels in panel A represent specific chemicals as follows: 1) ametryn, 2) zineb, 3) chlorsulfuron, 4) trichlorfon, 5) bis(tributyltin)oxide, 6) 2,4-dichlorophenol, 7) phorate, 8) glyphosate, 9) potassium permanganate, 10) calcium hypochlorite, 11) 2-butenoic acid, 3-[(dimethoxyphosphinyl)oxy]-, methyl ester, and 12) terbufos.

the HC5 ratios fell within the range 0.1–10 (Fig. 2). Specifically, out of a total of 573 and 93 comparisons made for acute and chronic SSDs, respectively, the HC5 ratios were found to be outside the range 0.1–10 in only 12 and 2 cases, respectively. In all 14 of those cases, the HC5 ratios were less than 0.1. The implication was that the HC5 values of the other SSDs were 0.1 times lower than those of log-normal SSDs. No common characteristic of these 14 chemicals was identified in terms of their MoA, AICc differences, or the number of species per SSD (Table 2). In addition, the average compositions of the taxonomic groups (algae, invertebrates, fish, and amphibians) among the chemicals listed in Table 2 were 18.3 %, 36.0 %, 44.7 %, and 1.0 %, respectively, for acute toxicity data, and 41.4 %, 45.2 %, 13.4 %, and 0 %, for chronic toxicity data. These compositions were similar to the averaged compositions among all the chemicals examined (as mentioned above), suggesting the unclear influence of taxonomic composition on the deviated HC5 ratios.

The median HC5 ratios of log-logistic SSDs to log-normal SSDs were 0.99 and 0.98 for acute and chronic SSDs, respectively, and the ratios fell within the range 0.5–2 for 97 % of acute SSDs and 94 % of chronic SSDs. This result was in line with the findings of previous studies, which have often reported trivial differences in fitting between the log-normal and log-logistic distributions (Xu et al., 2015; Zhao and Chen, 2016). Similarly, the median HC5 ratios of Burr type III SSDs to log-normal SSDs were 0.88 and 0.94 for acute and chronic SSDs, respectively, and in 84% of acute SSDs and 74% of chronic SSDs, the HC5 ratios were within the range 0.5–2. However, the median HC5 ratios of Weibull SSDs to log-normal SSDs were smaller: 0.51 for acute SSDs and 0.38 for chronic SSDs. The HC5 ratios fell within the range 0.5–2 in 45% of acute SSDs and 19% of chronic SSDs, indicating that Weibull SSDs resulted in lower HC5 values than log-normal SSDs more frequently than log-logistic or Burr type III SSDs did.

The differences in the HC5s between the Weibull and log-normal SSDs that exceeded a factor of 10 (Fig. 2) were primarily found when the AICc differences exceeded 10 (six acute cases and one chronic case out of nine and two cases, respectively). The implication was that the empirical data did not support these Weibull SSDs. The HC5 values based on these Weibull SSDs should therefore be handled with considerable caution in ecological risk assessments, despite their conservative nature. Furthermore, as was the case for the AICc differences between the Weibull and log-normal SSDs, the HC5 ratios were clearly related to skewness (Figure S4A). The implication was that the presence of higher toxicity values on the right side of the distribution contributed to the lower HC5 estimates in the Weibull SSDs. The differences between the Weibull and log-normal SSDs were also clear in two examples, trichlorfon and glyphosate (Figure S2 A, B), where the HC5 ratios fell outside the range 0.1–10, and the AICc differences exceeded 20. The toxicity data for these two chemicals showed positive skewness, and when the highest toxicity values were excluded, the HC5 ratios for both chemicals fell within the range 0.1–10.

Other cases where the HC5 ratios were outside the range 0.1–10 used the Burr type III distribution for SSD estimation, but the AICc differences were less than 10 (Table 2). The relationships between the HC5 difference and skewness or kurtosis of the toxicity data were unclear in the case of the Burr type III model (Figure S4). However, these Burr type III SSDs produced poor fits in the lower tails of the SSDs (Figure S2C–E). As partly discussed in the European Food Safety Authority guidance (EFSA Panel on Plant Protection Products and their Residues (PPR), 2013), these results emphasize the importance of visually inspecting the fitting of the lower tails of SSDs when selecting a statistical distribution to derive an SSD, and thereby an HC5 value.

3.4. Implications for choice of distribution to derive HC5 values

Based on AICc differences as well as visual inspection of SSD curves, the log-normal SSD was generally better or equally good for the majority of acute toxicity estimates of 191 chemicals and chronic toxicity estimates of 31 chemicals. Consequently, together with the fact that the ratios of HC5 values of other alternative SSDs to those of log-normal SSDs generally fell within the range 0.1–10, our findings indicate that the log-normal distribution can be a reasonable first choice for SSD derivation and do not contest the existing widespread use of log-normal SSDs (Aldenberg et al., 2001; Wheeler et al., 2002). However, it should be emphasized that our results do not preclude the use of other distributions for SSD derivation. In addition, the present study has examined the four distributions that have been frequently used for SSD estimation in many previous studies as well as ecological risk assessments and regulatory processes, but several other distributions, such as arctangent distribution, have been recently employed to estimate SSDs (Huang et al., 2022). Further studies would be valuable to test our findings by incorporating these other distributions.

Recent advances, particularly in computational techniques, have enabled the application and selection of multiple statistical distributions for SSD derivation (Fox et al., 2021). This development has further led to the incorporation of a model-averaging approach, wherein an HC5 value is estimated using a weighted average of SSDs based on multiple statistical distributions (Thorley and Schwarz, 2018). While our study compared the application of individual statistical distributions to estimate an SSD and investigating differences between this approach and the model-averaging approach is beyond the scope of this study, the

Table 2
Chemicals with hazardous concentration for 5% of species (HC5) ratios of alternative species sensitivity distributions (SSDs) to log-normal SSDs outside the range 0.1–10.

A). Acute SSDs

HC5 ratio	Distribution	Chemical name	AICc difference	Number of species per SSD	MoA
0.097	Burr type III	Ametryn	−8.81	37	Unclassified
0.095	Weibull	2,4-Dichlorophenol	10.28	30	Narcotic
0.092	Burr type III	Zineb	3.05	21	Unclassified
0.091	Weibull	2-Butenoic acid, 3-[(dimethoxyphosphinyl)oxy]-, methyl ester	4.23	24	Specifically acting
0.066	Weibull	Glyphosate	27.37	37	Unclassified
0.049	Weibull	Phorate	8.9	27	Specifically acting
0.047	Weibull	Potassium permanganate	19.74	35	Unclassified
0.045	Weibull	Bis(tributyltin)oxide	17.41	50	Unclassified
0.035	Burr type III	Chlorsulfuron	0.87	19	Unclassified
0.026	Weibull	Terbufos	6.44	15	Specifically acting
0.018	Weibull	Trichlorfon	21.56	79	Specifically acting
0.012	Weibull	Calcium hypochlorite	13.84	25	Unclassified

B). Chronic SSDs

HC5 ratio	Distribution	Chemical name	AICc difference	Number of species per SSD	MoA
0.079	Weibull	Linuron	7.99	19	Narcotic
0.073	Weibull	Chlorpyrifos	11.9	51	Specifically acting

results clearly suggested the inclusion of a log-normal distribution as the candidate distribution for model averaging.

Whereas the optimal methods for estimating SSDs remain an open question even with the computational advances, the use of the log-normal distribution (i.e., the normal distribution for log10-transformed data) offers several inherent advantages, including the ease of calculation and mathematical properties that have been thoroughly investigated in terms of SSDs (see previous studies (Aldenberg and Jaworska, 2000; Kamo et al., 2022) for more details). Given the fact that there has been a call for the harmonization of the derivation of environmental quality benchmarks (Batley and Warne, 2017), our findings provide a valuable basis for a discussion about what statistical distributions should be fitted.

4. Conclusion

To address the fundamental question of which statistical distribution should be applied for SSD estimation, acute and chronic SSDs for 191 and 31 chemicals, respectively, were estimated by fitting the four statistical distributions to aquatic toxicity data. SSDs based on the different statistical distributions were then compared, focusing mainly on AICc and HC5 values. The comparison of models based on AICc differences revealed that log-normal SSDs exhibited better or equally good performance compared to other models in 95% of acute and 99% of chronic SSD estimation cases. In addition, in cases where the AICc values of the log-normal SSDs were substantially higher than those of other models (indicating less support for log-normal SSDs by AICc), the HC5 ratios between the log-normal and other models consistently remained below a factor of 10. Consequently, these results indicate that the log-normal distribution can be regarded as a reasonable first candidate for estimating SSDs.

CRediT authorship contribution statement

Miina Yanagihara: Writing – review & editing, Writing – original draft, Software, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Kyoshiro Hiki:** Writing – review & editing, Software, Methodology, Investigation, Conceptualization. **Yuichi Iwasaki:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

All data used in this study are available in the EnviroTox database at https://envirotoxdatabase.org. The R code for the SSD modeling and an example dataset can be found on the GitHub website at https://github.com/miinay/SSD_distribution_comparison.

Acknowledgements

We are grateful for helpful comments from Renske Hoondert and anonymous reviewers. This study was supported by the Environment Research and Technology Development Fund (JPMEERF20231001) of the Environmental Restoration and Conservation Agency of Japan and by the Ministry of Education, Culture, Sports, Science and Technology, Japan (MEXT) to a project on Joint Usage/Research Center– Leading Academia in Marine and Environment Pollution Research (LaMer). M.Y. acknowledges funding from the Japan Society for the Promotion of Science (JSPS KAKENHI, Grant Number 21J00885).

Supplementary material

(i) Table S1, AICc differences and HC5 ratios of the chemicals with AICc differences less than −10; (ii) Figure S1, comparison of the acute SSDs estimated by the four models for five chemicals; (iii) comparisons with previous studies selecting the "best" models for SSDs; (iv) Table S2, comparison of best models by reanalysis of the data in He et al. (2014); (v) Figure S2, comparisons of acute SSDs estimated by the four models for five chemicals; (vi) Figure S3, comparison between the AICc difference and the skewness and the kurtosis of the toxicity data; (vii) Figure S4, comparison between HC5 ratio and skewness and kurtosis of toxicity data. (PDF) (viii) List of the chemicals examined in SSD comparisons (XLSX)

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ecoenv.2024.116379.

References

Adams, W., Blust, R., Dwyer, R., Mount, D., Nordheim, E., Rodriguez, P.H., Spry, D., 2020. Bioavailability assessment of metals in freshwater environments: a historical review. Environ. Toxicol. Chem. 39, 48–59.

Aldenberg, T., Jaworska, J.S., 2000. Uncertainty of the hazardous concentration and fraction affected for normal species sensitivity distributions. Ecotoxicol. Environ. Saf. 46, 1–18.

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

Aldenberg, T., Jaworska, J.S., Traas, T.P., 2001. Normal species sensitivity distributions and probabilistic ecological risk assessment. Species Sensit. Distrib. Ecotoxicol. 49–102.

Batley, G.E., Warne, M.S.J., 2017. Harmonization of water and sediment quality guideline derivation. Integr. Environ. Assess. Manag. 13, 458–459.

Burnham, K.P., Anderson, D.R., Huyvaert, K.P., 2011. AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. Behav. Ecol. Sociobiol. 65, 23–35.

Burr, I.W., 1942. Cumulative frequency functions. Ann. Math. Stat. 13, 215–232.

Canadian Council of Ministers of the Environment, 2019. Canadian water quality guidelines for the protection of aquatic life: manganese. Canadian Council of Ministers of the Environment, Winnipeg, MB.

Connors, K.A., Beasley, A., Barron, M.G., Belanger, S.E., Bonnell, M., Brill, J.L., de Zwart, D., Kienzler, A., Krailler, J., Otter, R., Phillips, J.L., Embry, M.R., 2019. Creation of a curated aquatic toxicology database: EnviroTox. Environ. Toxicol. Chem. 38, 1062–1073.

Dormann, C.F., Calabrese, J.M., Guillera-Arroita, G., Matechou, E., Bahn, V., Bartoń, K., Beale, C.M., Ciuti, S., Elith, J., Gerstner, K., Guelat, J., Keil, P., Lahoz-Monfort, J.J., Pollock, L.J., Reineking, B., Roberts, D.R., Schröder, B., Thuiller, W., Warton, D.I., Wintle, B.A., Wood, S.N., Wüest, R.O., Hartig, F., 2018. Model averaging in ecology: a review of Bayesian, information-theoretic, and tactical approaches for predictive inference. Ecol. Monogr. 88, 485–504.

EFSA Panel on Plant Protection Products and their Residues (PPR), 2013. Guidance on tiered risk assessment for plant protection products for aquatic organisms in edge-of-field surface waters. Wiley. ⟨https://doi.org/10.2903/j.efsa.2013.3290⟩.

European Commission., 2011. Common Implementation Strategy for the Water Framework Directive (2000/60/EC): Guidance Document No. 27. Technical guidance document for deriving environmental quality standards (No. Technical Report-2011-055). https://doi.org/10.2779/43816.

Fox, D.R., van Dam, R.A., Fisher, R., Batley, G.E., Tillmanns, A.R., Thorley, J., Schwarz, C.J., Spry, D.J., McTavish, K., 2021. Recent developments in species sensitivity distribution modeling. Environ. Toxicol. Chem. 40, 293–308.

Freeman, J.B., Dale, R., 2013. Assessing bimodality to detect the presence of a dual cognitive process. Behav. Res. Methods 45, 83–97.

Grist, E.P.M., Leung, K.M.Y., Wheeler, J.R., Crane, M., 2002. Better bootstrap estimation of hazardous concentration thresholds for aquatic assemblages. Environ. Toxicol. Chem. 21, 1515–1524.

Hamoutene, D., Marteinson, S., Kingsbury, M., McTavish, K., 2023. Species sensitivity distributions for two widely used anti-sea lice chemotherapeutants in the salmon aquaculture industry. Sci. Total Environ. 857, 159574.

He, W., Qin, N., Kong, X., Liu, W., Wu, W., He, Q., Yang, C., Jiang, Y., Wang, Q., Yang, B., Xu, F., 2014. Ecological risk assessment and priority setting for typical toxic pollutants in the water from Beijing-Tianjin-Bohai area using Bayesian matbugs calculator (BMC). Ecol. Indic. 45, 209–218.

Hiki, K., Iwasaki, Y., 2020. Can we reasonably predict chronic species sensitivity distributions from acute species sensitivity distributions? Environ. Sci. Technol. 54, 13131–13136.

Huang, P., Liu, S.-S., Wang, Z.-J., Ding, T.-T., Xu, Y.-Q., 2022. Deriving the predicted no effect concentrations of 35 pesticides by the QSAR-SSD method. Chemosphere 298, 134303.

Johnson, A.C., Jin, X., Nakada, N., Sumpter, J.P., 2020. Learning from the past and considering the future of chemicals in the environment. Science 367, 384–387.
Kamo, M., Hayashi, T.I., Iwasaki, Y., 2022. Revisiting assessment factors for species sensitivity distributions as a function of sample size and variation in species sensitivity. Ecotoxicol. Environ. Saf. 246, 114170.
Kienzler, A., Connors, K.A., Bonnell, M., Barron, M.G., Beasley, A., Inglis, C.G., Norberg-King, T.J., Martin, T., Sanderson, H., Vallotton, N., Wilson, P., Embry, M.R., 2019. Mode of action classifications in the EnviroTox database: development and implementation of a consensus MOA classification. Environ. Toxicol. Chem. 38, 2294–2304.
Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., 2023. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien.
Newman, M.C., Ownby, D.R., Mézin, L.C.A., Powell, D.C., Christensen, T.R.L., Lerberg, S. B., Anderson, B.A., 2000. Applying species-sensitivity distributions in ecological risk assessment: assumptions of distribution type and sufficient numbers of species. Environ. Toxicol. Chem. 19, 508–515.
Oginah, S.A., Posthuma, L., Hauschild, M., Slootweg, J., Kosnik, M., Fantke, P., 2023. To split or not to split: characterizing chemical pollution impacts in aquatic ecosystems with species sensitivity distributions for specific taxonomic groups. Environ. Sci. Technol. 57, 14526–14538.
Persson, L., Carney Almroth, B.M., Collins, C.D., Cornell, S., de Wit, C.A., Diamond, M.L., Fantke, P., Hassellöv, M., MacLeod, M., Ryberg, M.W., Søgaard Jørgensen, P., Villarrubia-Gómez, P., Wang, Z., Hauschild, M.Z., 2022. Outside the safe operating space of the planetary boundary for novel entities. Environ. Sci. Technol. 56, 1510–1521.
Pfister, R., Schwarz, K.A., Janczyk, M., Dale, R., Freeman, J.B., 2013. Good things peak in pairs: a note on the bimodality coefficient. Front. Psychol. 4, 700.
Posthuma, L., Suter II, G.W., Traas, T.P., 2002. Species Sensitivity Distributions in Ecotoxicology. CRC Press.
Posthuma, L., van Gils, J., Zijp, M.C., van de Meent, D., de Zwartd, D., 2019. Species sensitivity distributions for use in environmental protection, assessment, and management of aquatic ecosystems for 12 386 chemicals. Environ. Toxicol. Chem. 38, 703–711.
R Core Team, 2022. R: A Language and Environment for Statistical Computing.
Schwarz, C.J., Tillmanns, A.R., 2019. Improving statistical methods to derive species sensitivity distributions. Province of British Columbia: Victoria, BC, Canada..
Shao, Q., 2000. Estimation for hazardous concentrations based on NOEC toxicity data: an alternative approach. Environmetrics 11, 583–595.
Steffen, W., Richardson, K., Rockström, J., Cornell, S.E., Fetzer, I., Bennett, E.M., Biggs, R., Carpenter, S.R., de Vries, W., de Wit, C.A., Folke, C., Gerten, D., Heinke, J., Mace, G.M., Persson, L.M., Ramanathan, V., Reyers, B., Sörlin, S., 2015. Sustainability. Planetary boundaries: guiding human development on a changing planet. Science 347, 1259855.
Thorley, J., Schwarz, C., 2018. ssdtools: an R package to fit species sensitivity distributions. J. Open Source Softw. 3, 1082.
Warne, M., Batley, G.E., Van Dam, R.A., Chapman, J.C., Fox, D.R., Hickey, C.W., Stauber, J.L., 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. Australian and New Zealand Governments and Australian state and territory governments, Canberra..
Wang, Z., Qu, C., Zhang, J., Zhi, L., Tang, T., Yao, H., Li, W., Shi, C., Qi, S., 2023. Constructing model-averaging species sensitivity distributions of Phenanthrene based on reproductive fitness: Implications for assessing ecological risk in urban watershed. J. Hazard. Mater. 443, 130296.
Wheeler, J.R., Grist, E.P.M., Leung, K.M.Y., Morritt, D., Crane, M., 2002. Species sensitivity distributions: data and model choice. Mar. Pollut. Bull. 45, 192–202.
Wickham, H., 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York.
Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.L., Miller, E., Bache, S.M., Müller, K., Ooms, J., Robinson, D., Seidel, D.P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., Yutani, H., 2019. Welcome to the {tidyverse}. J. Open Source Softw. 4, 1686.
Xu, F.-L., Li, Y.-L., Wang, Y., He, W., Kong, X.-Z., Qin, N., Liu, W.-X., Wu, W.-J., Jorgensen, S.E., 2015. Key issues for the development and application of the species sensitivity distribution (SSD) model for ecological risk assessment. Ecol. Indic. 54, 227–237.
Yanagihara, M., Hiki, K., Iwasaki, Y., 2022. Can chemical toxicity in saltwater be predicted from toxicity in freshwater? A comprehensive evaluation using species sensitivity distributions. Environ. Toxicol. Chem. 41, 2021–2027.
Zhao, J., Chen, B., 2016. Species sensitivity distribution for chlorpyrifos to aquatic organisms: model choice and sample size. Ecotoxicol. Environ. Saf. 125, 161–169.