To Split or Not to Split

Characterizing Chemical Pollution Impacts in Aquatic Ecosystems with Species Sensitivity Distributions for Specific Taxonomic Groups

Oginah, Susan Anyango; Posthuma, Leo; Hauschild, Michael; Slootweg, Jaap; Kosnik, Marissa; Fantke, Peter

Published in:
Environmental Science and Technology

Link to article, DOI:
10.1021/acs.est.3c04968

Publication date:
2023

Document Version
Publisher's PDF, also known as Version of record

Environmental Science & Technology

pubs.acs.org/est

Article

To Split or Not to Split: Characterizing Chemical Pollution Impacts in Aquatic Ecosystems with Species Sensitivity Distributions for Specific Taxonomic Groups

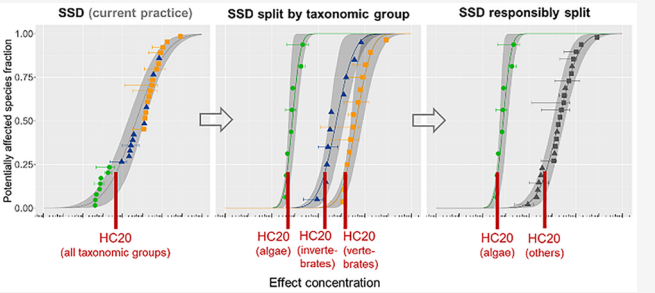
Susan Anyango Oginah,* Leo Posthuma, Michael Hauschild, Jaap Slootweg, Marissa Kosnik, and Peter Fantke*

Cite This: *Environ. Sci. Technol.* 2023, 57, 14526−14538

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Bridging applied ecology and ecotoxicology is key to protect ecosystems. These disciplines show a mismatch, especially when evaluating pressures. Contrasting to applied ecology, ecotoxicological impacts are often characterized for whole species assemblages based on Species Sensitivity Distributions (SSDs). SSDs are statistical models describing per chemical across-species sensitivity variation based on laboratory toxicity tests. To assist in the aligning of the disciplines and improve decision-support uses of SSDs, we investigate taxonomic-group-specific SSDs for algae/cyanobacteria/aquatic plants, invertebrates, and vertebrates for 180 chemicals with sufficient test data. We show that splitting improves pollution impact assessments for chemicals with a specific mode of action and, surprisingly, for narcotic chemicals. We provide a framework for splitting SSDs that can be applied to serve in environmental protection, life cycle assessment, and management of freshwater ecosystems. We illustrate that using split SSDs has potentially large implications for the decision-support of SSD-based outputs around the globe.



KEYWORDS: *freshwater ecosystems, mode of action, ecotoxicity, Water Framework Directive, water quality, life cycle impact assessment*

INTRODUCTION

Characterizing ecotoxicity effects, whether as part of chemical safety assessment, evaluating the environmental performance of products and services in a life cycle perspective, or environmental quality characterization, requires addressing different chemicals' potential to cause harm on different species,[1] while bridging the disciplines of applied ecology and ecotoxicology.[2] This can be achieved using chemical-specific species sensitivity distributions (SSDs). SSDs are classically used to describe variations in sensitivity across multiple species and are commonly derived from collections of laboratory toxicity test endpoints, such as no-observed effect concentrations (NOECs) or the effect concentration causing a response in 50% of the exposed individuals (EC50s).[3,4] Field-based Species Sensitivity Distributions (fSSDs) have been proposed as they are considered more ecologically relevant. However, they present challenges, e.g., the isolation of the effect of a single chemical from combined effects of multiple stressors. Recognizing the regulatory and other practical uses of current laboratory-data-based SSDs, we focus on the "classical", laboratory-data-based SSDs in the present paper.

Laboratory-toxicity data-based SSDs are practically used for regulatory purposes and Life Cycle Impact Assessment (LCIA), e.g., to derive protective standards (threshold concentrations) or expected impact levels of ambient chemical pollution.[5,6] Recently, their use has expanded to the comprehensive diagnosis of the role of chemical pollution as a driver for biodiversity loss in polluted ecosystems by using SSD-based mixture toxic pressure information (expressed as msPAF, the multisubstance Potentially Affected Fraction of species) as pressure metric, as this resulted in reduced parameters numbers and thus improved statistical power in diagnostic analyses.[2,7,8] The choice of required input data and the statistical distribution methods vary among jurisdictions. Models commonly used to fit SSDs include log-normal, log−logistic, or other models that fit the available data well, and commonly, confidence intervals or other metrics of variability and uncertainty are reported.[3,9] Crucial to acknowledge is that SSDs are commonly fitted to all available test data per chemical—following the principles developed by the earliest users—where it is assumed that the SSD describes the exposure-impact relationship for whole field species

Received: June 26, 2023
Revised: September 7, 2023
Accepted: September 8, 2023
Published: September 21, 2023

assemblages. Today, two key motives support the derivation of SSDs for distinct taxonomic groups, i.e., split SSDs.

- First, as common in applied ecology, environmental assessment practices focus on separate taxonomic groups (rather than whole assemblages).[10]
- Second, as recognized in applied ecotoxicology, different compound groups can have different (specific) modes of action (MoAs) (such as insecticides affecting insects most), which implies that the currently used, nonsplit SSDs may show poor statistical fits to data across taxonomic groups.[3,11]

The latter argument was already given in an outlook on developments in SSDs in 2002, arguing that a split in taxonomically distinct SSDs and accounting for the mode of action knowledge would be beneficial statistically and conceptually, with improved interpretation for the decision-support uses of SSD-outputs,[12] with further discussions by Fox et al. 2021.[3] Commonly, the derived toxic thresholds, i.e., protective environmental concentrations (HC5) and Predicted No Effect Concentration (PNEC) for aquatic communities and regulatory applications, could be improved by splitting the SSDs regardless of the SSD distributions used.[3,13]

The optional splitting of SSDs comes with a potential trade-off. They become statistically less robust because split SSDs are based on data per taxonomic group. Decision support applications require robust SSDs, defined by their insensitivity to changes in the available collections of input data. Robustness can be characterized with a statistical approach, enabling categorization of optional splitting as "responsible splits" (robust resulting SSDs per taxonomic group) or not (nonrobust outcomes, statistical trade-offs).

In the context of the use of SSDs in setting environmental quality standards, for life cycle assessments, and for diagnostic assessments of causes of global change, the present study's main goal was to investigate whether and how the splitting of SSDs—here, according to taxonomic groups—can be systematically undertaken, that is for chemicals with and without a specific mode of action while accounting for statistical trade-offs and to evaluate whether such splitting yields an improved impact characterization of exposure to chemical contaminants. To achieve this goal, we defined four specific objectives: (i) To derive a harmonized ecotoxicity database from available, curated freshwater ecotoxicity test data and to enrich this database with taxonomic and mode of action information. (ii) To propose a generically applicable framework for deriving split SSDs and demonstrate the utility of the framework for a set of chemicals with sufficient available freshwater test data. (iii) To evaluate whether using the proposed framework would result in different decision-support outcomes, i.e., improved characterization of the expected ecological impacts of chemical pollution, given that many chemicals have limited available test data. (iv) To, finally, derive practical and broadly applicable rules describing when and how chemical-specific ecotoxicity data can be split responsibly. In order to illustrate the potential relevance of splitting SSDs, we apply these rules to a set of chemicals and illustrate the effects for the derivation of protective standards and for the derivation of chemical-specific ecotoxicity effect factors for use in LCIA.[14,15]

METHODS

Overview. The study on splitting SSDs was developed with a selection of test data, the choice for the log-normal model to create SSDs, and following a stepwise approach for broader applications when splitting SSDs would be considered, where example outputs focus on both protective criteria and LCIA. These steps are elaborated below. We emphasize that the findings on splitting SSDs are generic and can be applied to other data selection criteria, statistical models, and SSD decision-support outputs. We followed the recommendations derived in the Global Life Cycle Impact Assessment Method (GLAM) effort under the auspices of the United Nations Environment Programme to derive metrics for assessing ecotoxicity impacts in LCIA as discussed by Owsianiak et al. 2023.[15] Main reasons are that the derived effect threshold of 20th percentile is close to the domain of environmentally relevant concentrations and that this threshold requires only a minimum of 5 species to have 1 tested species falling within the range below the HC20.

Experimental Test Data Curation and Harmonization. We started from a database of experimental ecotoxicity effect test data for 9868 chemicals,[4] from which we selected experimental data (i.e., we removed read-across data) for freshwater species for data harmonization and curation. This process included harmonization of species names, classification of species into taxonomic groups, and calculation of average effect test values across data points per combination of species and chemical (see Table A1). We then selected chemicals with data for three or more distinct species and taxonomic groups. The taxonomic groups used in developing the split-SSD argument were pragmatically inspired by the European Union (EU)-Water Framework Directive, which discerns various Biological Quality Elements—from which we selected three groups (here: Algae, cyanobacteria, and aquatic plants (A), Invertebrates (I), and Vertebrates (V)). The resulting data set for 180 chemicals is provided in the Supporting Information (Excel file Table S1).

Chemicals were first classified according to a systematic taxonomy based on the ClassyFire approach and assigned a mode of action (MoA) based on classifications from different pesticide resistance action committees, the Verhaar scheme,[16,17] and reported MoA information,[18] before mapping the chemical MoA to taxonomic groups.[19] If MoA information was lacking, the MoA reported for the chemical class was used. Chemical use categories were derived from prescribed use information (pesticidecompendium.bcpc.org).

As a next step, recognizing that the available raw species sensitivity data are of diverse kinds (acute NOEC, chronic NOEC, acute EC50, etc.), we used a set of data-driven extrapolation formulas to "translate" the diverse raw data types into a set of harmonized endpoint data, to enable derivation of SSDs from those. Since we illustrate the methodology for the context of LCIA, we extrapolated the available test endpoint data (e.g., acute and chronic NOEC, EC50, and EC10) to chronic EC10 equiv (the recommended starting point for deriving ecotoxicity impacts in LCIA). This was done by applying the formulas of Table 1, based on data-driven patterns recognized and described by Aurisano et al. 2019.[20]

Chemicals were then classified as "data-rich" (chemicals with ≥3 distinct species from the ≥3 taxonomic groups) for the present study or as "data-poor" (<3 distinct species per taxonomic group and/or <3 taxonomic groups) in line with, e.g., Müller et al. 2017.[21] The set of "data-rich" chemicals was kept for further analysis.

Systematic Decision Tree to Evaluate the Splitting of SSDs. All SSDs in the present study were derived as log-normal

https://doi.org/10.1021/acs.est.3c04968
Environ. Sci. Technol. 2023, 57, 14526−14538

Table 1. Overview of Regression Equations Derived Based on All Available Freshwater Test Data for 9868 Chemicals Used to Extrapolate Laboratory-Test Derived Species Sensitivity EndPoints (the Diverse Set of Reported Sensitivity Metrics: Column 1) to Chronic EC10 equiv, Used in the Present Study to Derive SSD-EC10eq Based on the Extrapolated EC10-Equivalent Data for the 180 Chemicals

Endpoints	Extrapolation equation
Acute NOEC	$\log EC10_{chronic} = 0.816 \times \log NOEC_{acute} + 0.021$
Chronic NOEC	$\log EC10_{chronic} = 0.965 \times \log NOEC_{chronic} - 0.144$
Acute EC50	$\log EC10_{chronic} = 0.869 \times \log EC50_{acute} - 0.508$
Chronic EC50	$\log EC10_{chronic} = 0.872 \times \log EC50_{chronic} + 0.733$
Acute EC10	$\log EC10_{chronic} = 0.813 \times \log EC10_{acute} + 0.967$

distribution of species sensitivity (here: chronic EC10 equiv) data. Note that other SSD models may be fitted to the data, and those can be split for taxonomic groups, and the reasoning below can be specifically adapted if needed for those models.

A log-normal distribution is characterized by its mean and standard deviation, which are also used as moments of the log-normal SSDs (as μ and σ, respectively). A systematic evaluation decision tree was designed to distinguish between alternative outcomes (full split into three SSDs, if not: partial split, if not: no split). Given the two main arguments for splitting the data in taxonomic-group-specific SSDs, the first decision point is a statistical test series to evaluate whether some taxonomic subsets of data differ significantly from other subsets. The statistical tests for inter-SSD comparisons can conveniently be based on generally applied statistical test methods given the underlying distribution model (log-normal). The systematic evaluation decision tree for splitting SSDs thus considers among other evaluations of (dis)similarities in slopes and means, as depicted in Figure 1.

The following tests were executed, starting from the raw data and the relevant descriptive statistics (μ and σ), as illustrated in Figure 1. Levene's test was run to check the homogeneity of variances (considering thus σ = slope differences among SSDs). If variances are not significantly different, a one-way analysis of variance (ANOVA) (parametric) was run to evaluate the (dis)similarity of means (μ, the position parameter of the SSDs) among subsets of the three taxonomically grouped test data. Nonhomogeneous variance is a signal of differences across subsets of the data; in this case, the Kruskal–Wallis test (nonparametric) was applied to evaluate differences in μ. *A posteriori* multiple comparisons tests followed if there were significant differences between compared taxonomic groups (that is, Tukey's HSD (parametric) and Dunn's (non-parametric) tests, with the p-value set at 0.05). If full split (A $\neq$ I $\neq$ V) is *not* supported, the independent t test (parametric) and Mann–Whitney U test (nonparametric) were used to compare the mean of one taxonomic group versus the other two groups merged, e.g., A $\neq$ I + V. For further confirmation, the conclusion drawn from these tests on splitting SSDs was evaluated by deriving a linear regression model, whereby one group is considered an "anchor" to test whether others differ from the anchor (see Table 2).

The set of results (on differences in σ's and/or μ's) were collated, together with the *a posteriori* test results, to draw a conclusion on statistical motives to employ a full or partial split based on the whole assemblage of test data for a chemical (minimum: 3 × 3 = 9). The resulting SSDs may be three SSDs based on (as a minimum) three data points each, which can thus produce nonrobust outcomes based on the identified subsets of the data. Therefore, the decision tree proceeds with an evaluation of the robustness of the resulting SSDs, whereby nonrobust outcomes are identified. Robust SSDs were defined by quantifying the confidence interval around the derived LCIA ecotoxicity impact metric (i.e., HC20). A robust SSD yields an HC20 with a narrow confidence interval, whereby we pragmatically applied 5 squared geometric standard deviation as the boundary below which we identify the HC20 of a split-, partially split, or no-split as robust. Where a split caused a thus-defined nonrobust SSD, it was investigated whether partial remerging (i.e., A+V, I+V, or A+I) or full-remerging (A+I+V) resulted in robust SSDs (following the same robustness test).

All the statistical analyses and the building of split SSDs were performed in R version 4.1.2,[22] modifying and expanding existing code to construct SSDs (edild.github.io/ssd). Figures were generated using the ggplot R package, version 3.4.1.[23]

Evaluating Mode of Action and Use Category Information. The results of the splitting procedure were evaluated *vis-a-vis* information on MoA and use categories, whereby it was expected that specific modes of action (e.g., insecticidal action) or use category (e.g., labeling as an insecticide) would imply splitting off (at least) the target taxonomic group as a separate SSD. Data were plotted in different subgroups and combinations to verify associations between these two aspects and the results of the splitting approach.

Derivation of User-Oriented Impact Metrics (HC20) Values for LCIA. The SSDs that resulted from the split-assessments are (in the case study) SSDs based on chronic EC10 equiv (SSD-EC10eq) of a compound, which are in turn used to define the ecotoxicological effect factor of that compound at the HC20-level (the 20th percentile of that SSD).[15,24] Thus, for sufficiently robust SSDs, we derived those values for all of the studied compounds. In addition to this, we also illustrate the impacts of splitting on protective regulatory standards (related to HC5, PNEC, and similar concepts) for some selected compounds. Finally, the uncertainty assessment around taxonomic group-specific HC20s was quantified by combining intraspecies and interspecies variability (see Supporting Information Uncertainty analysis section for details).

RESULTS

Harmonized Ecotoxicity Test Data for Freshwater Species. We start with 120,835 species-specific toxicity test data, totaling 9868 chemicals, 1123 species, and 234 test endpoints, distributed as shown in Figure 2. Because of nonsystematic global testing practices, the data set does not equally cover the taxonomic groups, test durations, and endpoint types. Invertebrates are the primary taxonomic group (78%) in the data set. Likewise, acute toxicity data dominate the data set (71% of the total data across taxonomic groups), mainly acute EC50s for invertebrates (44,077 acute EC50s and 20,000 acute NOECs), with fewer chronic EC50s (n = 5555), of which only very few (n = 82) data are for vertebrates.

Species sensitivities span many orders of magnitude for both short-term peak and longer-term chronic exposures (Figure 2). For example, acute EC50s range between 7.3×10^{-8} and 3.3×10^{-9} μg/L for invertebrates, and chronic EC50s range from 1.6×10^{-5} to 1.0×10^{-8} μg/L for invertebrates. Likewise, acute NOECs range between 8.0×10^{-7} and 1.0×10^{-9} μg/L, and chronic NOECs range between 6.0×10^{-6} and 2.5×10^{-8} μg/L.

https://doi.org/10.1021/acs.est.3c04968
Environ. Sci. Technol. 2023, 57, 14526−14538

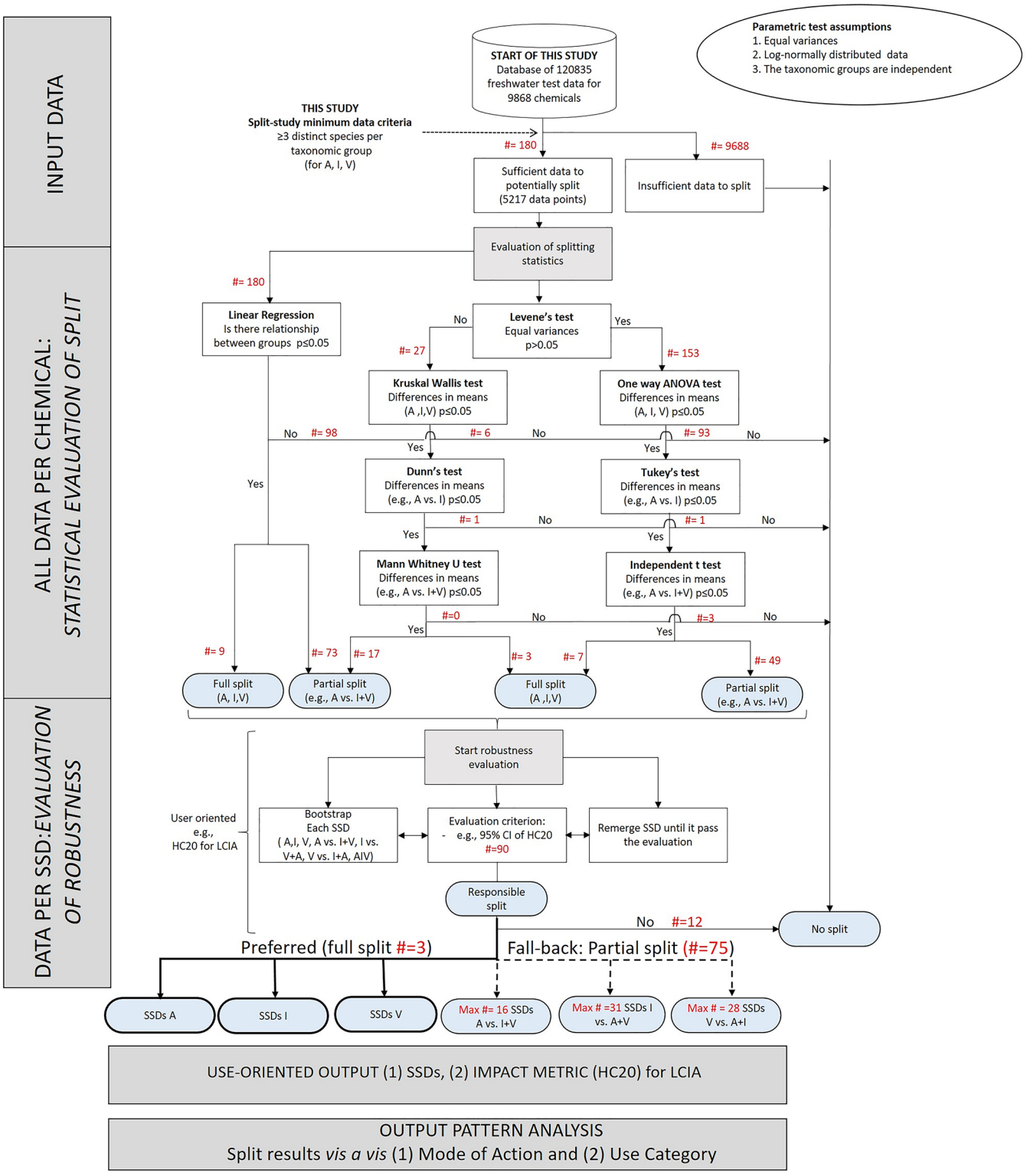


Figure 1. Schematic summary of a decision tree for evaluating whether the assemblage of available ecotoxicity data for a chemical can be subdivided into (here) three taxonomic groups based on three specific Water Framework Directive defined Biological Quality Elements: Algae, cyanobacteria, and aquatic plants (A), Invertebrates (I), and Vertebrates (V). Numbers of cases involved in each step (as result) are shown in red. (top) Description and selection of input data for the start of the splitting approach. (middle) Analysis steps to evaluate statistical motives to split using all available data per chemical (fully or partially). (bottom) Evaluating whether the split results in a trade-off of a nonrobust species sensitivity distribution (SSD) for one or more subgroups. (bottom gray blocks) Summarizing results for practical Life Cycle Impact Assessment (20th percentile values) and evaluating whether statistical split covaries with the mode of action and use category information. Note that the no-split approach is the historically best-known form and most frequently applied format of constructing SSDs and using those for environmental decision support purposes, prescribed in various policy guidance documents (details vary across jurisdictions).

https://doi.org/10.1021/acs.est.3c04968
Environ. Sci. Technol. 2023, 57, 14526−14538

Table 2. List of Statistical Tests Used to Evaluate the Potential of Split SSDs for Different Taxonomic Groups: Algae, Cyanobacteria, and Aquatic Plants (A), Invertebrates (I), and Vertebrates (V)

Step	Statistical test	Description	Interpretation
1	Levene's test	-Tests the null hypothesis that the variances (related to the SSD σ-parameter) of different taxonomic groups (A, I, V) are equal	$p > 0.05$ $\sigma_1 \approx \sigma_2 \approx \sigma_3$
		-If $p > 0.05$, the test does not (fails to) reject the null hypothesis that the variances of different taxonomic groups are equal. If $p < 0.05$, then taxonomic groups have different slopes (SSDs are different)	
1.1	One way ANOVA test	-Tests the null hypothesis that the mean (related to the SSD μ-parameter) of different taxonomic groups is equal (A, I, V)	$p \leq 0.05$: $\mu_1 \neq \mu_2 \neq \mu_3$
		-If $p > 0.05$, then taxonomic groups have the same mean. If $p \leq 0.05$, then taxonomic groups have different means (SSDs have different position parameters)	
1.1.1	Tukey's test	-Test multiple pairwise comparisons between groups' means (μ) to identify which groups have a different mean (e.g., A vs I)	$p \leq 0.05$: $\mu_A \neq \mu_I$
		If $p > 0.05$, then taxonomic groups have the same mean	
		-If $p \leq 0.05$, then taxonomic groups have different means (specific split of SSDs).	
1.1.2	Independent t test	-Tests the null hypothesis that the mean (μ) of different taxonomic groups is equal (e.g., V vs I+A)	$p \leq 0.05$: $\mu_V \neq \mu_{I+A}$
		-If $p > 0.05$, then taxonomic groups have the same mean	
		-If $p \leq 0.05$, then taxonomic groups have different means. (specific split of SSDs)	
1.2	Kruskal–Wallis test	-Nonparametric equivalent of ANOVA test, tests the null hypothesis that the mean (μ) of different taxonomic groups is equal (A, I, V)	$p \leq 0.05$: $\mu_1 \neq \mu_2 \neq \mu_3$
		-If $p > 0.05$, then taxonomic groups have the same mean	
		-If $p \leq 0.05$, then taxonomic groups have different means.	
1.2.1	Dunn's test	-Nonparametric post hoc test similar to Tukey's test. Test multiple pairwise comparisons between groups' mean to identify which taxonomic groups have different means (e.g., A vs I)	$p \leq 0.05$: $\mu_A \neq \mu_I$
		-If $p > 0.05$, then taxonomic groups have the same mean	
		-If $p \leq 0.05$, then taxonomic groups have different means.	
1.2.2	Mann–Whitney U test	-Nonparametric equivalent to independent t test, test the null hypothesis that the mean (μ) is equal across different taxonomic groups (V vs I+A)	$p \leq 0.05$: $\mu_V \neq \mu_{I+A}$
		-If $p > 0.05$, then taxonomic groups have the same mean	
		-If $p \leq 0.05$, then taxonomic groups have different means.	
2.0	Linear Regression	-Linear model with categorical predictors.	$p \leq 0.05$: $\mu_A \neq \mu_I$
		-Test group-level differences between groups (e.g., A vs I). If $p \leq 0.05$: the test rejects the null hypothesis that the mean (μ) is equal across different taxonomic groups; thus, the taxonomic groups are not related	

From the curated data, 180 chemicals were selected as data-rich (chemicals with data for ≥3 distinct species from ≥3 taxonomic groups), yielding 5217 test end point data for developing and testing the SSD-splitting framework. Upon deriving the chronic-EC10-values from these data, some taxa dominate the final data-rich subset for further study steps, with 47% invertebrates, 33% vertebrates, and 20% algae, cyanobacteria, and aquatic plants. Note that only 1.81% of the chemicals have sufficient data for a potential full split into three taxonomic group-specific SSDs.

Split SSDs for Different Taxonomic Groups and Relation to Mode of Action. Statistically significant support for full or partial splitting into SSDs representing Algae, cyanobacteria, and aquatic plants (A), Invertebrates (I), and Vertebrates (V), or the combinations of AI, AV, or IV, was found for 3 (<2%) and 75 (42%) out of 180 data-rich chemicals, respectively, based on statistical tests comparing the mean and standard deviation of SSDs (procedure in Figure 1). Notably, the available data support the use of a split-SSD modeling approach for some narcotic chemicals (see Figure 3), which is visually indicated by nonoverlapping means (dots) and standard deviations for different taxonomic groups. This latter outcome was highly unexpected, given three decades of accepted no-split SSD practices and the expectation that split-SSDs would be found for only chemicals with a specific mode of action. However, a split can also be warranted for narcotic chemicals, provided that the available data set is sufficiently rich. The Supporting Information Excel file (Excel Table S2) presents all statistical details and characteristics of the resulting SSDs.

We found a clear and consistent pattern of higher sensitivity of the targeted taxonomic group for chemicals with a specific MoA (63% of the 180 chemicals; see Supporting Information Figure S1), with, as expected, the more sensitive taxonomic groups predominantly represented at the lower end of the shown sensitivity distribution patterns (See Figure 3). For example, I and A are the most sensitive groups in the panels shown on insecticides (especially AChE (Acetyl Choline Esterase) inhibition) and photosynthesis inhibition, respectively. The nontargeted groups are often not statistically different, resulting in a partial split, likely partly attributable to the absence of an MoA-related mechanism that would induce a split or possibly also due to lower numbers of test data for the nontargeted taxonomic groups.

We summarized the data further, using the "working point" on impacts that are used to derive LCIA Effect Factors based on global consensus recommendations (the Hazardous Concentration for 20% of the species, HC20, derived from chronic EC10 equiv). This yielded similar results, here illustrated for the gross chemical use categories. Again, there was a match between the expected and observed sensitive taxonomic groups, e.g., insecticides and I and herbicides and A; Figure 4a. Unexpectedly, one herbicide (tributyltin-cation, CAS 36643-28-4) with an endocrine-disrupting MoA affected invertebrates, which appeared as the most sensitive taxonomic group. Fungicides showed general toxicity, with the most sensitive

https://doi.org/10.1021/acs.est.3c04968
Environ. Sci. Technol. 2023, 57, 14526−14538

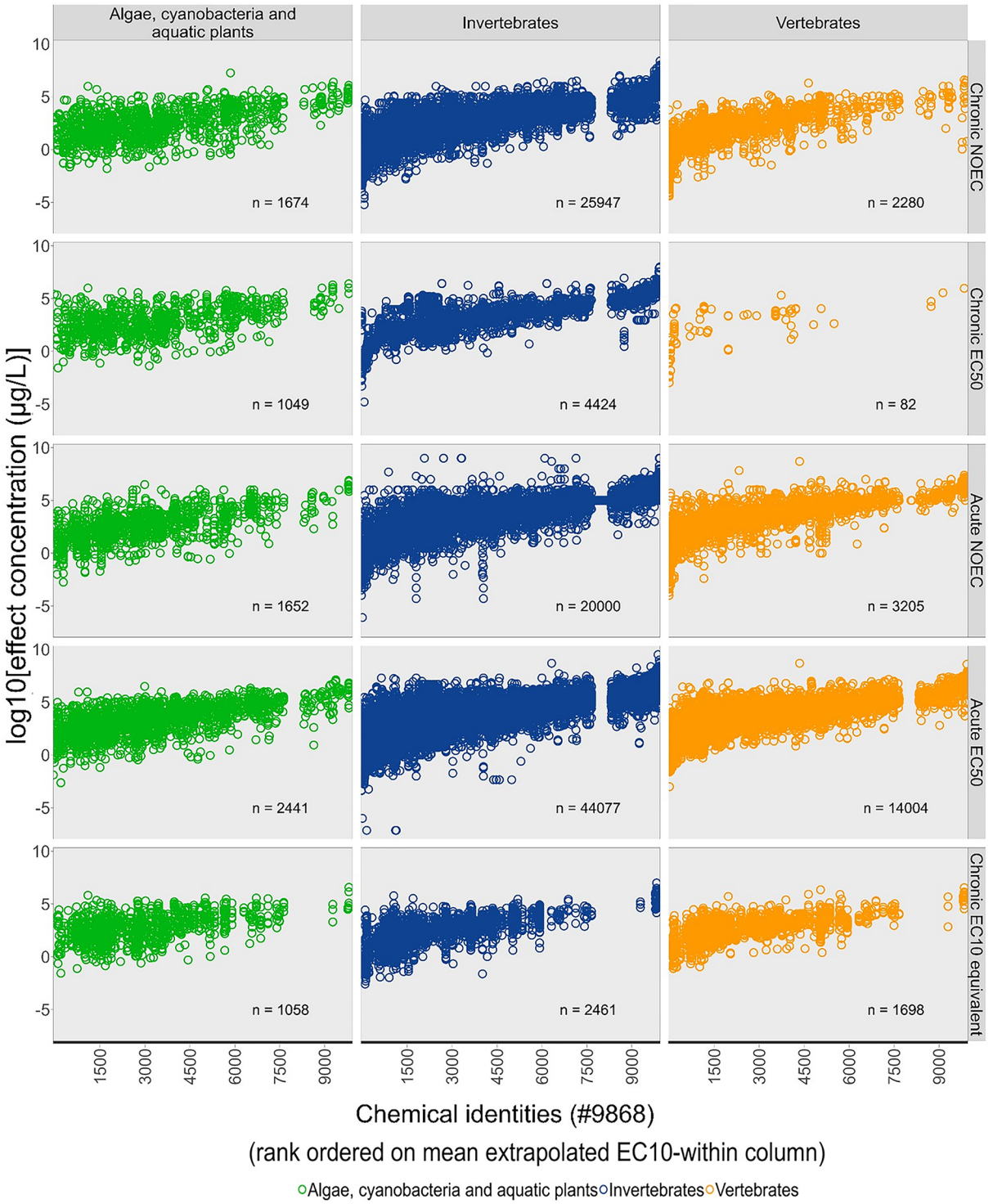


Figure 2. Distribution of 120,835 species sensitivity endpoints (Y) for 9868 rank-ordered chemicals (X), distinguishing taxonomic groups (colors) and endpoint types (measured: rows 1 to 4; extrapolated split-research data: row 5). Rank order was based on mean extrapolated chronic EC10-equivalent values per chemical for each column. On the *x*-axis, value gaps lack data for taxonomic group × endpoint combinations (e.g., vertebrates × NOEC) for the gap intervals.

taxonomic groups differing across chemicals, with algae, cyanobacteria, and aquatic plants tending to be less sensitive than other taxonomic groups. Moreover, chemicals with no specified use category ("Other uses") showed no clear pattern of any taxonomic group being affected most.

Specific Regulatory and Decision-Relevance Issues. The relevance of our findings for the various contemporary decision-support uses of SSDs is illustrated first by the fact that our analyses cover 15 Water Framework Directive (WFD) Priority Substances (marking chemicals of current Europe-wide concern, black stars in Figure 4) and two chemicals listed under the fourth WFD Watch List (black crossed dots in Figure 4, marking chemicals of emerging concern for water quality policies).[25] Second, the relevance of splitting for the outcomes of

https://doi.org/10.1021/acs.est.3c04968
Environ. Sci. Technol. 2023, 57, 14526−14538

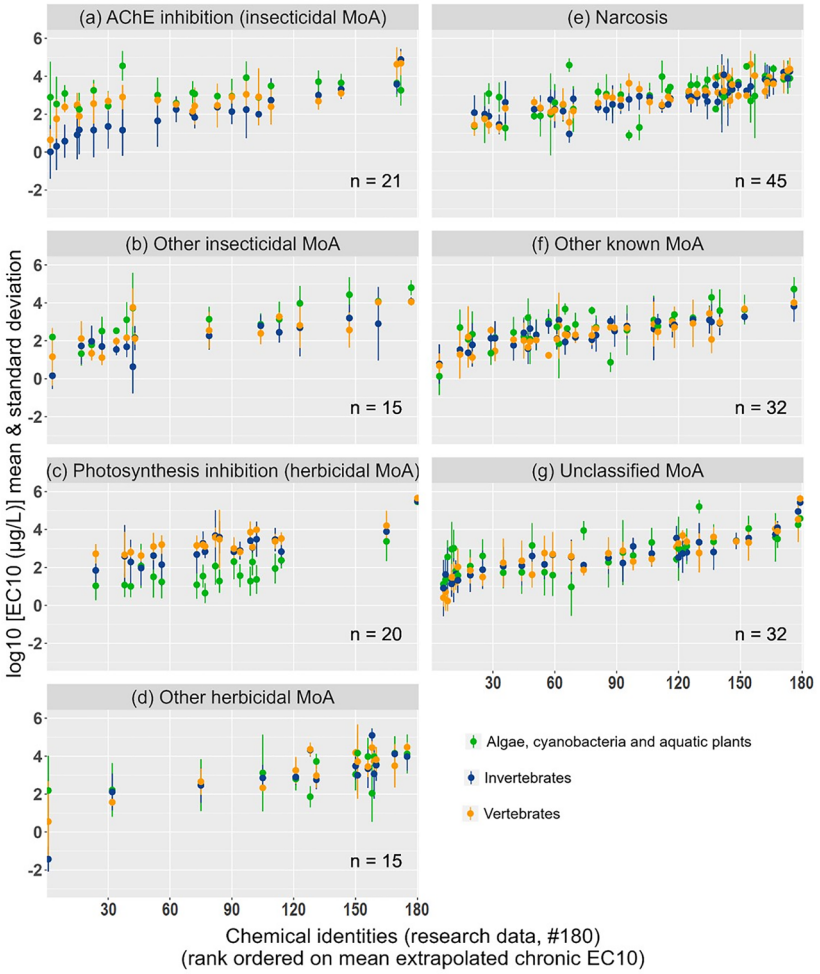


Figure 3. Distributions of the species sensitivity impact metric (Y, chronic EC10 equiv) for 180 rank-ordered chemicals (X). Taxonomic groups are colored. Chemicals within each panel are rank-ordered based on the mean impact metric values calculated across all data per chemical. MoA = Mode of Action, *n* = number of chemicals in each panel. SSD-splitting grossly coincides (visually) with nonoverlapping dots and standard deviations for different taxonomic groups (colors). Letters in each panel: For chemicals targeting invertebrates: (a) AChE inhibition (insecticidal MoA; chemical class with most data for this MoA); (b) other insecticidal MoA (lumped); for chemicals targeting algae, cyanobacteria, and aquatic plants as primary producers; (c) photosynthesis inhibition (herbicidal MoA with most data); (d) other herbicidal MoA (lumped); for chemicals with baseline toxicity: (e) Narcosis. (f) "Other known MoA" includes chemicals for which MoA was provided but for which a specific targeted taxonomic group is unknown. (g) "Unclassified MoA" includes chemicals for which MoA information was lacking.

using SSDs for decision making shows as substantial differences between SSD-based insights on protective criteria and/or impacts of chemicals generated without (classic approach) and with applying our proposed (partial) split, elaborated in the next section further.

Third, based on currently available data, the split concerns a relatively large proportion of the chemicals. The use category "herbicides" involved 56 chemicals, among which data for 29 chemicals (statistical output) supported a full or partial split, with a suggested partial splitting of algae, cyanobacteria, and aquatic plants from the other taxonomic groups for 25 chemicals (A vs I+V). The use category "insecticides" involved 27 chemicals, among which data for 20 chemicals (statistical output) supported a full or partial split, with a suggested splitting of I from the other groups for 14 chemicals. The "fungicide" category involved 32 chemicals. Data for 14 of these chemicals supported a partial split; data for 14 of these chemicals supported a partial split, whereby 12, 8, and 6 were separated from the rest of the group for primary producers, invertebrates, and vertebrates, respectively.

The "Other uses" category involved 65 chemicals; data for 27 (statistical output) chemicals supported partial splitting, with data for 15, 14, and 17 chemicals showing a splitting of A, I, and V from the rest, respectively. This indicates that there are certain chemicals in industrial or other (nonagricultural) uses to which particular taxonomic groups are more sensitive than others. For instance, algae, cyanobacteria, and aquatic plants appeared sensitive to 2,4-dinitrotoluene, and vertebrates appeared sensitive to phenol, both relevant in polymer and plastic production. Again, note that splitting SSDs is not limited to chemicals with a specific known MoA or specified use category.

Quantitative Implications of Split SSDs for Decision Support. The relevance of the proposed split-SSD approach in decision support is shown quantitatively in Figure 5, illustrating all potential outcomes (no split, full split, and partial split). Vertical black and red lines show that critical concentrations (at the fifth and 20th percentile levels) derived with the SSDs differ substantially between the classical no-split approach and the full- or partial-split approaches. That is shown for the derivation of protective environmental standards (black vertical lines, where a

https://doi.org/10.1021/acs.est.3c04968
Environ. Sci. Technol. 2023, 57, 14526−14538

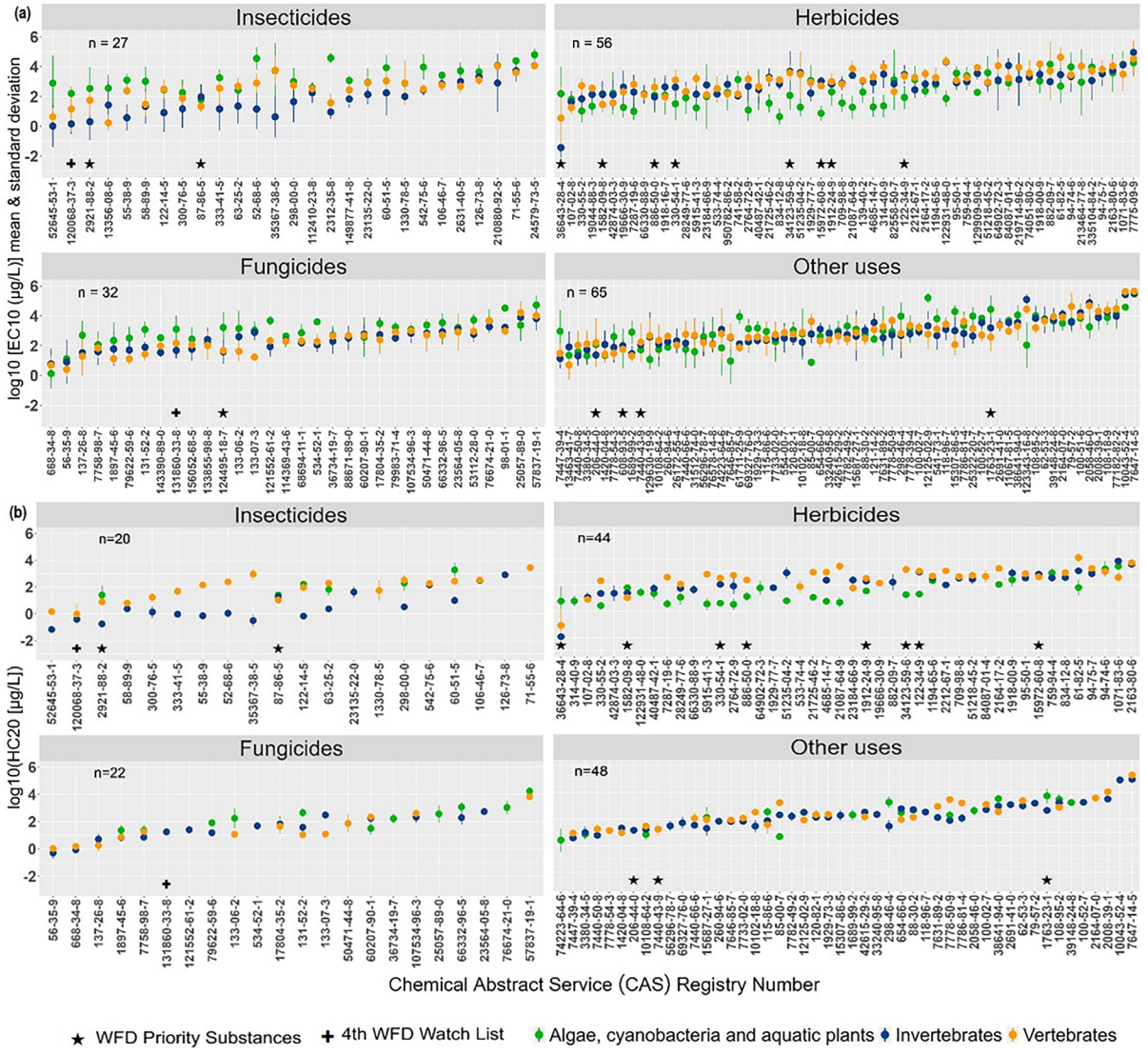


Figure 4. (a) Distributions of the species sensitivity impact metric (Y, chronic EC10 equiv) for chemical use categories (panels), rank ordered as in panel (b). Use categories: Insecticides are chemicals used for targeting invertebrates; Herbicides are chemicals used for targeting algae, cyanobacteria and aquatic plants as primary producers; Fungicides are chemicals targeting fungi. "Other uses" are chemicals for which the targeted taxonomic group is not specified (e.g., industrial chemicals). (b). Similar results present the HC20-metric that is practically used in Life Cycle Impact Assessment global consensus approaches (20th percentile of the distribution of the above impact metric across tested species) for only robust SSDs. Percentile values of SSDs are expressed in LCIA as HC = Hazardous Concentration. HC20-estimates (with standard deviations) are summarized for 134 chemicals with ≥6 data points (low uncertainty) for 261 chemical-taxonomic group combinations.

regulatory-defined low-end, fifth percentile value of a chronic-data SSD is used to express a critical protective concentration, X) and for the estimation of the impact of pollution, by quantifying the potentially affected fraction of species (Y) at an ambient exposure level (X), here at the 20th percentile of the SSD (red vertical lines). Figure 5 also illustrates the influence of splitting for confidence intervals around the SSDs (gray bands), which are wider for data-poor SSDs. In these examples, the HC5 or HC20 can shift by more than 1 order of magnitude as a consequence of splitting, as compared to the whole-assemblage (classical) SSD.

The first 4 rows of the plots (i.e., simazine, fenthion, trichlorfon, and sodium pentachlorophenate) illustrate SSD patterns for chemicals with nonoverlapping 95% confidence interval (CI) ranges of their split SSDs, supporting a partial or full split (supported by statistical tests, see Methods). In comparison, pyraflufen-ethyl showed overlapping 95% CI ranges, and statistical tests did not support splitting. In all cases where a responsible split was supported by statistical evaluation, the whole-species assemblage SSDs showed that the observed test data for one or more particular taxonomic groups were unevenly distributed over the SSD, with, e.g., the sensitive taxonomic group clustering toward the lower tail (Figure 5).

Decision Tree and Resulting HC20s. Figures 3−5 summarize and illustrate results and decision-supporting relevance of splitting to derive taxonomic group-specific SSDs. These results reflect outcomes of a systematic decision tree (Figure 1), in which test data, statistical testing to motivate a split, and user-required SSD-robustness considerations are combined. In our analysis, the assessment of all available data reveals that a full or partial split would be supported for 90 of the 180 chemicals. However, judgment of the trade-off effect on SSD robustness showed that three chemicals resulted in robust full split SSDs, 75 in partial split SSDs as a fallback option, and

https://doi.org/10.1021/acs.est.3c04968
Environ. Sci. Technol. 2023, 57, 14526−14538

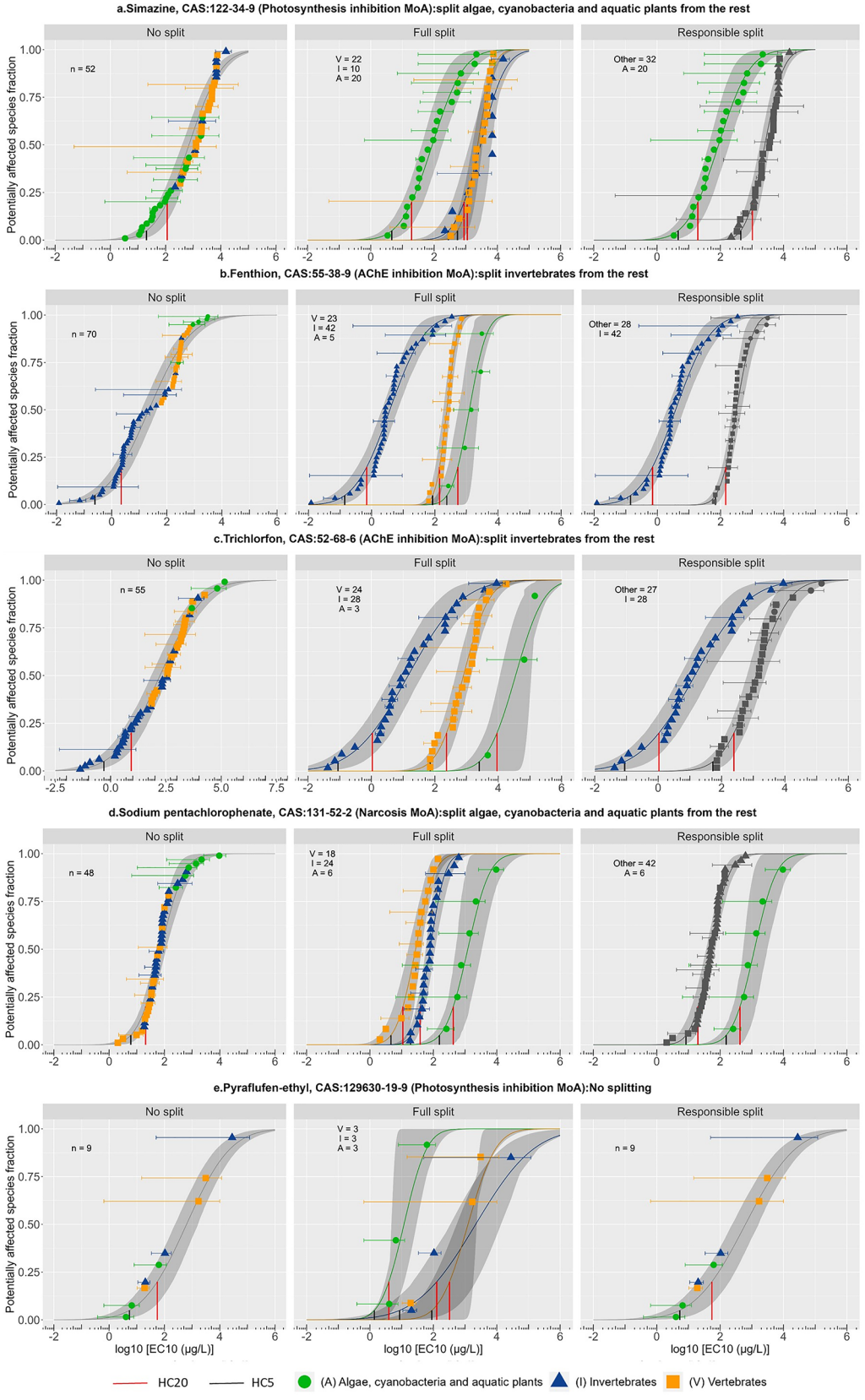


Figure 5. Illustration of deriving no-split, full-split, or partially split (taxonomic group-specific) species sensitivity distributions and its consequences for deriving protective environmental quality standards and for use in Life Cycle Impact Assessment (LCIA). There is no consequence for these decision support applications if vertical black or red lines, respectively, are similar for full- or partial-split SSDs compared to no-split SSDs. Black or red lines for a chemical show differences of up to more than 2 orders of magnitude. Columns: no-split (historically the common SSD use, left), full split (middle), and responsible split (right) SSDs. Rows are selected chemicals. Panels show sensitivity endpoints across species (dots: chronic EC10-equivalents), sigmoid curves (log-normal fitted SSDs), and 95% confidence intervals (shaded). Protective standards (maximum concentrations) are

https://doi.org/10.1021/acs.est.3c04968
Environ. Sci. Technol. 2023, 57, 14526−14538

Figure 5. continued

derived from a uniform policy-chosen Y = 0.05, with protective threshold concentrations (Hazardous Concentration for 5% of the species, HC5) derived on X (black lines). Impact magnitudes in LCIA are derived from a uniform Y = 0.2 with similar consequences derived on X (red lines). Rows from top to bottom illustrate the results of different MoA-related splitting situations and results from data-richer to data-poorer chemicals. Row 1: data-rich herbicide (simazine); rows 2 and 3 data-rich insecticides (fenthion and trichlorfon); row 4 data-rich chemical with baseline toxicity (sodium pentachlorophenate); row 5: data-poor herbicide (pyraflufen-ethyl). The error lines of the data points show intraspecies variability in the test data set for each chemical. Interspecies variability is represented by 95% confidence intervals (gray bands) based on bootstrapping (1000 iterations). With sufficient data, SSD splitting is supported for chemicals with specific MoA (rows 1, 2, 3, and 5) but (surprisingly) also for baseline toxicity (rows 4).

102 in nonrobust split SSDs, of which the latter are therefore represented by the classical no-split SSD (Excel Table S2). The final result on SSDs robustness involved a pragmatic decision to include only SSDs with a squared geometric standard deviation (GSD^2) $\leq$ 5 around the log-mean as a cutoff point. The sequence of steps and analyses of the shown decision tree can be used to judge whether available test data on a chemical can be (partially) split SSDs into robust results for chemicals other than the 180 study chemicals included in this study (shown for the log-normal model, but applicable in a similar way to other data and model choices).

DISCUSSION

Triggered by a need to set regulatory, protective environmental quality benchmarks, the global regulatory use of SSDs started with recognizing that log-transformed ecotoxicity test data collected for various species appeared to follow a bell-shaped distribution.[4] Thereupon, the first broad use of SSDs was to derive protective environmental quality standards from all available data, based on the assumption that the distribution of sensitivities of tested species resembles that of the nontested field species assemblage.[6] This subsequently provided the basis for a wide array of uses of SSDs in environmental quality protection, assessment, and management. However, almost all of the (regulatory) applications still derive SSD from all data, aiming to represent the whole species assemblage following initial decisions. Various jurisdictions prescribe different minimum requirements to the underlying effect data, recognizing (i) that different taxonomic groups should be represented to fulfill the adopted assumption and (ii) that more test data generally imply more robust SSD models. The present study reconsidered all this and evaluated from "first-principles" (the two ecological and statistical motives) whether the derivation of a (partially) split SSD is warranted and, if so, how that would operate and work out for decision support uses. Illustrated by chemicals with a minimum of necessary test data, we demonstrate that an improved fit of SSDs to the underlying data can be found even for some chemicals with a narcotic (nonspecific) MoA, in contrast to common expectations and our own initial beliefs.

Although the results show how the splitting of data into taxonomic group-specific SSDs improves the fit of the models to the data and the final decision-support interpretations caused by that, it should be recognized that the use of SSDs derived from laboratory toxicity data is not a panacea for all environmental problems with chemicals. The complexity and diversity of the chemical pollution problem has, so far, resulted in additional useful methods to characterize hazards, among which we include the derivation and use of field SSDs and bioassays—which are both methods that do not require laboratory-to-field extrapolation (as in classical SSDs). Environmental policies and bridging applied ecology and ecotoxicology can be served by investigating multiple lines of evidence to disentangle the effects of multiple stressors and unintended ambient mixtures.

The present study provides the scientific answer on "to split or not to split"; it is scientifically better to split SSDs according to taxonomic grouping (better fit of the models to the data and associated decision-support implications) unless the number and quality of available test data limits that. Whether this is implemented in practice and how far this would provide improved environmental assessments and management depend on the jurisdiction, the available data per chemical, and the difference in SSDs between taxonomic groups. If implemented in practice, there would be consequences for data collection (seeking to add data that appear missing for specific taxonomic groups) as well as for derivation of protective standards (such as HC5, PNEC, and similar terms) for LCIA and environmental quality assessment (potentially affected fraction of species). The split-approach undoubtedly results in fewer data per (split-)SSD, which may have various consequences in practice. Such SSDs may be statistically robust (as in the present study results). However, the splitting also relates to the debate on the "minimum number of species (or taxonomic groups) per SSD" and using an optional Safety Factor on an HC5 derived from a split-SSD. The minimum-number debate would need to ascertain that a (likely lower) minimum number of tests should represent the sensitivity variation within a taxonomic group. The Safety Factor debate—which was triggered by uncertainties—would need to consider that the method likely lowers the HC5 because of accounting for the sensitive taxonomic group, addressing part of the uncertainties of the "classical" approach. Such debates can start upon adopting splitting based on the methodology laid out in the present study.

Our results suggest that "always-splitting" is warranted as a starting point for any assessment but that splitting may be limited, and has trade-offs, given the characteristics of the available data. Specifically, an exploration of the available base data (120,835 toxicity endpoint values) shows that invertebrates are frequently tested and evaluated,[26−29] which results in a higher likelihood of finding robust split SSDs for invertebrates. The lowest number of data was available for bacteria and fungi, which hampers assessment of risks for these groups as well as for important functions of microorganisms in ecosystems, even without the opportunity to derive a split SSD for this group, where relevant. This points to the need for more tests of microorganisms in freshwater ecosystems. Despite limitations due to data scarcity, it is evident that a partial split may result when assessing chemicals other than the 180 studied chemicals.

The Role of Mode of Action and Use Category Information. Splitting is principally warranted from the viewpoint of applied ecology, where distinctive bioassessment approaches focusing on taxonomic groups are common. However, the motives for splitting have always been expected to be stronger for chemicals with a specific MoA or more grossly defined chemical use categories. In general, the results of both

https://doi.org/10.1021/acs.est.3c04968
Environ. Sci. Technol. 2023, 57, 14526−14538

mechanistic MoA and chemical use category considerations closely matched expectations on this, supported by patterns shown in Figure 3 and Figure 4. For instance, invertebrates appear at the lower level of the sensitivity distribution (lowest SSD mean) for chemicals designed to operate via the AChE-inhibition and labeled as insecticides, while primary producers fall into the lower tails of the distributions for photosynthesis inhibitors used as herbicides. Our results confirm that it is key to consider MoA and/or the use category (where this applies, such as for pesticides) information on chemicals to trigger, considering the use of better-fitting models for separate taxonomic groups.[30−32] The better principles and the better fit have implications for practical uses of SSDs, both for deriving protective standards and for use in environmental impact assessments and LCIA (Figure 5).

We made two additional notable observations. First, confirming previous studies and theory, the lowest variation in sensitivity across taxonomic groups was found in chemicals with nonspecific MoA, i.e., narcosis (Figure 3, vertical spread). This supports the hypothesis that even among not closely related species, toxicity through nonpolar narcosis is associated with relatively lower interspecies sensitivity differences.[33] Although there may be an applied ecology and a statistical reason to consider splitting data-rich chemicals with a narcotic MoA, the improved fit mainly implies improved SSD-based outputs for chemicals with specific MoA, making the latter the primary focus for aiming at splitting SSDs in practice.[3] Second, some outcomes are not predictable by MoA or use category information. For example, the herbicide tributyltin-cation (CAS 36643-28-4) is relevant in chemical formulations designed to control weeds, but there are potential side effects on invertebrates, more than for other herbicides. Most of the observed lower sensitivities in our SSDs are evidence of unwanted side effects, which are visible despite the diversity of the side effects, a low number of tests, and a few chemicals designed to control vertebrates.

The application of splitting or not has various implications for decision support. Scientifically, it is likely that a study on the calibration between data on the predicted msPAF values for an array of sampling sites and the observed effects on a particular taxonomic group at those sites is more meaningful when based on split SSDs, given the more accurate impact assessment upon splitting (Figure 5). For example, observed impacts on invertebrate species derived from ecological monitoring can be better calibrated to the msPAF derived from the SSD-Invertebrates than from the classical overall SSD, with a misfit of the SSD to the test data. This calibration (higher PAF implies higher risk, proven with calibration data or not) is often used as a basis for the decision-supporting uses of SSDs. Specific calibration work can now be undertaken to quantify the Potentially Disappearing Fraction (PDF) of species due to chemical exposure, given the possibility of quantifying impacts in terms of PAF from the split SSDs. PDF is a biodiversity damage metric used in LCIA for all impact categories related to ecosystem quality.[34,35]

In general, the decision-support uses of SSDs will be conceptually improved and numerically altered with split SSDs, provided that those are robust. This holds for data sets, model choices, or SSD-based output metrics that differ from the data, model, and metrics used in the present paper. The finding that splitting SSD is relevant in any case holds without prejudice to either of these matters. Upon splitting, conceptual numeric improvements may result in more accurate protective standards and better quantitative impact assessment of predicted or observed ambient pollution.[36] The black lines in Figure 5 (subfigures a, b, c, and d) illustrate that a protective (no impact) environmental quality standard, estimated as fifth percentile from an SSD of chronic data, is lower when a split SSD is employed. That is understandable, as HC5 for the most sensitive group now represents a protection of 95% of the species in that group. In turn, this has (similar) implications for regulatory-adopted criteria based on these estimated HC 5s, such as the PNEC. The red lines in Figure 5 (subfigures a, b, c, and d) show similar results for the LCIA-employed impact metric at the 20th percentile, leading to different impact estimates of the use of chemicals in products based on splitting.

After evaluation of the decision tree (Figure 1), our observation confirms that chemicals with more data (e.g., simazine and fenthion) and a specific MoA provide the strongest basis for responsible splitting. Thus, the more data, the more robust the SSDs after the (partial) split, even for narcotic chemicals (e.g., sodium pentachlorophenate). However, species selection bias during laboratory testing (i.e., for chemicals with specific MoA) currently limits to often find full-split SSDs. For example, trichlorfon, an insecticide operating via the AChE inhibition, statistically supports a full-split SSD. In contrast, few data points for a nontarget taxonomic group (i.e., Algae, cyanobacteria, and aquatic plants; $n = 3$) result only in partial splitting based on the SSD robustness check, indicating the need to include more tests for nontarget species to derive taxonomic group-specific split SSDs where appropriate. Thus, for one to have a full split SSD, all the taxonomic groups require sufficient data. At the other end of the spectrum, avoiding a negative trade-off for prediction accuracy with nonrobust SSDs that occur for data-poor chemicals is essential. For example, although the statistical tests on all available test data suggest that the primary producers could be split off from the rest of the groups for pyraflufen-ethyl, the broad and overlapping confidence intervals render the whole-assemblage SSD statistically more robust as compared to the partial-split alternative. Statistical assessments may be used to decide on split SSDs, or not, but not solely. It is also important to evaluate whether splitting is better in practice, yielding better decision support based on conceptual principles and trade-off effects. Relatively data-poor chemicals may result in split SSDs that are not robust for one or more taxonomic groups because the process of splitting counters the statistical rule that "more data result in more robust SSDs" (see Figure S2).

We selected 180 data-rich chemicals to develop a decision tree for splitting SSDs (Figure 1), which can be employed (or adapted) for all chemicals and taxonomic groups or other ways to group chemicals, species, or statistical models. In the exploration of data for a chemical, it is likely that a partial split (e.g., Algae, cyanobacteria, and aquatic plants versus Invertebrates and Vertebrates together) is found for many chemicals or other taxonomic groups than investigated in the present study. Applications to split according to specific traits or trophic positions would follow the same decision tree logic, resulting in risk information on chemical effects on specific traits or trophic positions.

Overall, our study highlights that splitting is a better approach in deriving SSDs and using the models for decision support, provided that the resulting SSDs are sufficiently robust. Robust split results improve the fit of the models to the data and, therefore, the interpretation of SSDs in the discussed uses. The relevance for decision support may potentially be further increased when a split would consider different service-providing units (SPUs), a concept used in the context of

https://doi.org/10.1021/acs.est.3c04968
Environ. Sci. Technol. 2023, 57, 14526−14538

ecosystem services research.[34] This is because it is key not only to protect and restore biodiversity in terms of structural characteristics of ecosystems (the present use) but also in terms of functional characteristics and provided services.[37,38] Assessments that would consider ecological information, such as functional groups or trait characteristics, may help to identify the SPU and ecosystem services that are both valuable and potentially impacted.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.est.3c04968.

A protocol for deriving split SSDs, including an R code used for split SSDs statistical analysis based on the process flow in Table A1 and uncertainty analysis in Table A2 (PDF)

Data used to evaluate the splitting of SSDs (XLSX)

■ AUTHOR INFORMATION

Corresponding Authors

Peter Fantke − *Quantitative Sustainability Assessment, Department of Environmental and Resource Engineering, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark;* orcid.org/0000-0001-7148-6982; Phone: +45 45254452; Email: pefan@dtu.dk; Fax: +45 45933435

Susan Anyango Oginah − *Quantitative Sustainability Assessment, Department of Environmental and Resource Engineering, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark;* orcid.org/0000-0002-0448-524X; Phone: +45 71664818; Email: sanog@dtu.dk

Authors

Leo Posthuma − *National Institute for Public Health and the Environment, 3720 BA Bilthoven, The Netherlands; Department of Environmental Science, Radboud University Nijmegen, 6525 AJ Nijmegen, The Netherlands;* orcid.org/0000-0003-0399-5499

Michael Hauschild − *Quantitative Sustainability Assessment, Department of Environmental and Resource Engineering, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark*

Jaap Slootweg − *National Institute for Public Health and the Environment, 3720 BA Bilthoven, The Netherlands*

Marissa Kosnik − *Quantitative Sustainability Assessment, Department of Environmental and Resource Engineering, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark;* orcid.org/0000-0002-2609-5816

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.est.3c04968

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank José De Sousa Jorge Ferreira for statistical advice, Dick de Zwart for providing and curating the initial ecotoxicity data set, and Kerstin Johanna Felicitas von Borries for debugging/checking the analysis code. This project has received funding from the Prorisk project financed by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant No. 859891. L.P. and J.S. were supported by SPR (Strategic Program RIVM) project No. S030006 "CE-TRANSIT," as run under the auspices of RIVM's Director General and RIVM's Scientific Advisory Board.

■ REFERENCES

(1) Fantke, P.; Aurisano, N.; Bare, J.; Backhaus, T.; Bulle, C.; Chapman, P. M.; De Zwart, D.; Dwyer, R.; Ernstoff, A.; Golsteijn, L.; Holmquist, H.; Jolliet, O.; McKone, T. E.; Owsianiak, M.; Peijnenburg, W.; Posthuma, L.; Roos, S.; Saouter, E.; Schowanek, D.; van Straalen, N. M.; Vijver, M. G.; Hauschild, M. Toward Harmonizing Ecotoxicity Characterization in Life Cycle Impact Assessment. *Environ. Toxicol. Chem.* **2018**, *37* (12), 2955−2971.

(2) Sigmund, G.; Ågerstrand, M.; Antonelli, A.; Backhaus, T.; Brodin, T.; Diamond, M. L.; Erdelen, W. R.; Evers, D. C.; Hofmann, T.; Hueffer, T.; Lai, A.; Torres, J. P. M.; Mueller, L.; Perrigo, A. L.; Rillig, M. C.; Schaeffer, A.; Scheringer, M.; Schirmer, K.; Tlili, A.; Soehl, A.; Triebskorn, R.; Vlahos, P.; vom Berg, C.; Wang, Z.; Groh, K. J. Addressing Chemical Pollution in Biodiversity Research. *Glob. Chang. Biol.* **2023**, *29* (12), 3240−3255.

(3) 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. Recent Developments in Species Sensitivity Distribution Modeling. *Environ. Toxicol. Chem.* **2021**, *40* (2), 293−308.

(4) Posthuma, L.; van Gils, J.; Zijp, M. C.; van de Meent, D.; de Zwart, D. Species Sensitivity Distributions for Use in Environmental Protection, Assessment, and Management of Aquatic Ecosystems for 12 386 Chemicals. *Environ. Toxicol. Chem.* **2019**, *38* (4), 905−917.

(5) Iwasaki, Y.; Sorgog, K. Estimating Species Sensitivity Distributions on the Basis of Readily Obtainable Descriptors and Toxicity Data for Three Species of Algae, Crustaceans, and Fish. *Peers J.* **2021**, *9*, e10981.

(6) Posthuma, L.; de Zwart, D. Species Sensitivity Distributions. In *Encyclopedia of Toxicology,* 3rd ed.; 2014; Vol. *4*, pp 363−368. DOI: 10.1016/B978-0-12-386454-3.00580-7.

(7) Lemm, J. U.; Venohr, M.; Globevnik, L.; Stefanidis, K.; Panagopoulos, Y.; van Gils, J.; Posthuma, L.; Kristensen, P.; Feld, C. K.; Mahnkopf, J.; Hering, D.; Birk, S. Multiple Stressors Determine River Ecological Status at the European Scale: Towards an Integrated Understanding of River Status Deterioration. *Glob. Chang. Biol.* **2021**, *27* (9), 1962−1975.

(8) Schneeweiss, A.; Juvigny-Khenafou, N. P. D.; Osakpolor, S.; Scharmüller, A.; Scheu, S.; Schreiner, V. C.; Ashauer, R.; Escher, B. I.; Leese, F.; Schäfer, R. B. Three Perspectives on the Prediction of Chemical Effects in Ecosystems. *Glob. Chang. Biol.* **2023**, *29* (1), 21−40.

(9) Sorgog, K.; Kamo, M. Quantifying the Precision of Ecological Risk: Conventional Assessment Factor Method vs. Species Sensitivity Distribution Method. *Ecotoxicol. Environ. Saf.* **2019**, *183* (May), 109494.

(10) Schäfer, R. B.; Jackson, M.; Juvigny-Khenafou, N.; Osakpolor, S. E.; Posthuma, L.; Schneeweiss, A.; Spaak, J.; Vinebrooke, R. Chemical Mixtures and Multiple Stressors: Same but Different? *Environ. Toxicol. Chem.* **2023**, *42* (9), 1915−1936.

(11) Aldenberg, T.; Jaworska, J. S.; Traas, T. P. *Normal Species Sensitivity Distributions and Probabilistic Ecological Risk Assessment* **2001**, DOI: 10.1201/9781420032314.ch5.

(12) Posthuma, L.; Suter, G. W.; Traas, T. P. *Species Sensitivity Distributions in Ecotoxicology;* CRC Press: Boca Raton, FL, 2002; pp 1−581.

(13) ECETOX. Estimating Toxicity Thresholds for Aquatic Ecological Communities from Sensitivity Distributions, February 11−13, 2014, Amsterdam, Netherlands, 2014; p 98. https://www.ecetoc.org/wp-content/uploads/2021/10/ECETOC_WR_28.pdf.

(14) Fantke, P.; Huang, L.; Overcash, M.; Griffing, E.; Jolliet, O. Life Cycle Based Alternatives Assessment (LCAA) for Chemical Substitution. *Green Chem.* **2020**, *22* (18), 6008−6024.

(15) Owsianiak, M.; Hauschild, M. Z.; Posthuma, L.; Saouter, E.; Vijver, M. G.; Backhaus, T.; Douziech, M.; Schlekat, T.; Fantke, P.

Ecotoxicity Characterization of Chemicals: Global Recommendations and Implementation in USEtox. *Chemosphere* **2023**, *310*, 136807.
(16) 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. Creation of a Curated Aquatic Toxicology Database: EnviroTox. *Environ. Toxicol. Chem.* **2019**, *38* (5), 1062−1073.
(17) Enoch, S. J.; Hewitt, M.; Cronin, M. T. D.; Azam, S.; Madden, J. C. Classification of Chemicals According to Mechanism of Aquatic Toxicity: An Evaluation of the Implementation of the Verhaar Scheme in Toxtree. *Chemosphere* **2008**, *73* (3), 243−248.
(18) Busch, W.; Schmidt, S.; Kühne, R.; Schulze, T.; Krauss, M.; Altenburger, R. Micropollutants in European Rivers: A Mode of Action Survey to Support the Development of Effect-Based Tools for Water Monitoring. *Environ. Toxicol. Chem.* **2016**, *35* (8), 1887−1899.
(19) Djoumbou Feunang, Y.; Eisner, R.; Knox, C.; Chepelev, L.; Hastings, J.; Owen, G.; Fahy, E.; Steinbeck, C.; Subramanian, S.; Bolton, E.; Greiner, R.; Wishart, D. S. ClassyFire: Automated Chemical Classification with a Comprehensive, Computable Taxonomy. *J. Cheminform.* **2016**, *8* (1), 1−20.
(20) Aurisano, N.; Albizzati, P. F.; Hauschild, M.; Fantke, P. Extrapolation Factors for Characterizing Freshwater Ecotoxicity Effects. *Environ. Toxicol. Chem.* **2019**, *38* (11), 2568−2582.
(21) Müller, N.; de Zwart, D.; Hauschild, M.; Kijko, G.; Fantke, P. Exploring REACH as a Potential Data Source for Characterizing Ecotoxicity in Life Cycle Assessment. *Environ. Toxicol. Chem.* **2017**, *36* (2), 492−500.
(22) R Core Team. A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria, 2021. https://www.r-project.org/.
(23) Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis,* 2nd ed.; Gentleman, R., Hornik, K., Parmigiani, G., Eds.; Springer-Verlag: New York, 2016; Vol. 35.
(24) Sala, S.; Biganzoli, F.; Mengual, E. S.; Saouter, E. Toxicity Impacts in the Environmental Footprint Method: Calculation Principles. *Int. J. Life Cycle Assess.* **2022**, *27* (4), 587−602.
(25) Cortes, L. G.; Marinov, D.; Sanseverino, I.; Cuenca, A. N.; Niegowska, M.; Rodriguez, E. P.; Lettieri, T. *Selection of Substances for the 4th Watch List under the Water Framework Directive;* Publication Office of the European Union: Luxembourg, 2022. DOI: 10.2760/01939.
(26) Lagadic, L.; Caquet, T. Invertebrates in Testing of Environmental Chemicals: Are They Alternatives? *Environ. Health Perspect.* **1998**, *106* (supp2), 593−611.
(27) Berger, E.; Haase, P.; Schäfer, R. B.; Sundermann, A. Towards Stressor-Specific Macroinvertebrate Indices: Which Traits and Taxonomic Groups Are Associated with Vulnerable and Tolerant Taxa? *Sci. Total Environ.* **2018**, *619−620*, 144−154.
(28) Berger, E.; Haase, P.; Oetken, M.; Sundermann, A. Field Data Reveal Low Critical Chemical Concentrations for River Benthic Invertebrates. *Sci. Total Environ.* **2016**, *544*, 864−873.
(29) Khamis, K.; Hannah, D. M.; Brown, L. E.; Tiberti, R.; Milner, A. M. The Use of Invertebrates as Indicators of Environmental Change in Alpine Rivers and Lakes. *Sci. Total Environ.* **2014**, *493*, 1242−1254.
(30) Maltby, L.; Blake, N.; Brock, T. C. M.; Van Den Brink, P. J. Insecticide Species Sensitivity Distributions: Importance of Test Species Selection and Relevance to Aquatic Ecosystems. *Environ. Toxicol. Chem.* **2005**, *24* (2), 379−388.
(31) Van Den Brink, P. J.; Blake, N.; Brock, T. C. M.; Maltby, L. Predictive Value of Species Sensitivity Distributions for Effects of Herbicides in Freshwater Ecosystems. *Hum. Ecol. Risk Assess.* **2006**, *12* (4), 645−674.
(32) Maltby, L.; Brock, T. C. M.; Van Den Brink, P. J. Fungicide Risk Assessment for Aquatic Ecosystems: Importance of Interspecific Variation, Toxic Mode of Action, and Exposure Regime. *Environ. Sci. Technol.* **2009**, *43* (19), 7556−7563.
(33) Robinson, A.; Lahive, E.; Short, S.; Carter, H.; Sleep, D.; Pereira, G.; Kille, P.; Spurgeon, D. Chemicals with Increasingly Complex Modes of Action Result in Greater Variation in Sensitivity between Earthworm Species. *Environ. Pollut.* **2021**, *272*, 115914.
(34) Oginah, S. A.; Posthuma, L.; Maltby, L.; Hauschild, M.; Fantke, P. Linking Freshwater Ecotoxicity to Damage on Ecosystem Services in Life Cycle Assessment. *Environ. Int.* **2023**, *171*, 107705.
(35) Verones, F.; Bare, J.; Bulle, C.; Frischknecht, R.; Hauschild, M.; Hellweg, S.; Henderson, A.; Jolliet, O.; Laurent, A.; Liao, X.; Lindner, J. P.; Maia de Souza, D.; Michelsen, O.; Patouillard, L.; Pfister, S.; Posthuma, L.; Prado, V.; Ridoutt, B.; Rosenbaum, R. K.; Sala, S.; Ugaya, C.; Vieira, M.; Fantke, P. LCIA Framework and Cross-Cutting Issues Guidance within the UNEP-SETAC Life Cycle Initiative. *J. Clean. Prod.* **2017**, *161*, 957−967.
(36) Kosnik, M. B.; Hauschild, M. Z.; Fantke, P. Toward Assessing Absolute Environmental Sustainability of Chemical Pollution. *Environ. Sci. Technol.* **2022**, *56* (8), 4776−4787.
(37) Edens, B.; Maes, J.; Hein, L.; Obst, C.; Siikamaki, J.; Schenau, S.; Javorsek, M.; Chow, J.; Chan, J. Y.; Steurer, A.; Alfieri, A. Establishing the SEEA Ecosystem Accounting as a Global Standard. *Ecosyst. Serv.* **2022**, *54*, 101413.
(38) Vysna, V.; Maes, J.; Petersen, J..; La Notte, A.; Vallecillo, S.; Aizpurua, N.; Ivits, E.; Teller, A. *Accounting for Ecosystems and Their Services in the European Union (INCA). Final Report from Phase II of the INCA Project Aiming to Develop a Pilot for an Integrated System of Ecosystem Accounts for the EU,* 2021 ed.; Statistical Report; Publications Office of the European Union: Luxembourg, 2021. DOI: 10.2785/19790.