Environmental Toxicology and Chemistry—Volume 43, Number 8—pp. 1914–1927, 2024
Received: 19 January 2024 | Revised: 2 March 2024 | Accepted: 11 May 2024

Hazard/Risk Assessment

Characterizing Freshwater Ecotoxicity of More Than 9000 Chemicals by Combining Different Levels of Available Measured Test Data with In Silico Predictions

Mélanie Douziech,[a,b,*] Susan Anyango Oginah,[c] Laura Golsteijn,[d] Michael Zwicky Hauschild,[c,e] Olivier Jolliet,[c] Mikołaj Owsianiak,[c] Leo Posthuma,[f,g] and Peter Fantke[c,e,*]

[a]Agroscope, Life Cycle Assessment Research Group, Zurich, Switzerland
[b]Centre of Observations, Impacts, Energy, MINES Paris Tech, PSL University, Sophia Antipolis, France
[c]Quantitative Sustainability Assessment, Department of Environmental and Resource Engineering, Technical University of Denmark, Lyngby, Denmark
[d]PRé Sustainability, Amersfoort, The Netherlands
[e]Centre for Absolute Sustainability, Technical University of Denmark, Lyngby, Denmark
[f]Department of Environmental Science, Radboud Institute for Biological and Environmental Science, Radboud University, Nijmegen, The Netherlands
[g]National Institute for Public Health and the Environment, Centre for Sustainability, Environment and Health, Bilthoven, The Netherlands

Abstract: Ecotoxicological impacts of chemicals released into the environment are characterized by combining fate, exposure, and effects. For characterizing effects, species sensitivity distributions (SSDs) estimate toxic pressures of chemicals as the potentially affected fraction of species. Life cycle assessment (LCA) uses SSDs to identify products with lowest ecotoxicological impacts. To reflect ambient concentrations, the Global Life Cycle Impact Assessment Method (GLAM) ecotoxicity task force recently recommended deriving SSDs for LCA based on chronic EC10s (10% effect concentration, for a life-history trait) and using the 20th percentile of an EC10-based SSD as a working point. However, because we lacked measured effect concentrations, impacts of only few chemicals were assessed, underlining data limitations for decision support. The aims of this paper were therefore to derive and validate freshwater SSDs by combining measured effect concentrations with in silico methods. Freshwater effect factors (EFs) and uncertainty estimates for use in GLAM-consistent life cycle impact assessment were then derived by combining three elements: (1) using intraspecies extrapolating effect data to estimate EC10s, (2) using interspecies quantitative structure–activity relationships, or (3) assuming a constant slope of 0.7 to derive SSDs. Species sensitivity distributions, associated EFs, and EF confidence intervals for 9862 chemicals, including data-poor ones, were estimated based on these elements. Intraspecies extrapolations and the fixed slope approach were most often applied. The resulting EFs were consistent with EFs derived from SSD-EC50 models, implying a similar chemical ecotoxicity rank order and method robustness. Our approach is an important step toward considering the potential ecotoxic impacts of chemicals currently neglected in assessment frameworks due to limited test data. *Environ Toxicol Chem* 2024;43:1914–1927. © 2024 The Author(s). *Environmental Toxicology and Chemistry* published by Wiley Periodicals LLC on behalf of SETAC.

Keywords: Ecological risk assessment; Environmental toxicology; Freshwater toxicology; Hazard/risk assessment; In silico methods; Predictive toxicology; Uncertainty

INTRODUCTION

Chemicals released into the environment can affect various ecosystems. Different assessment frameworks have been developed to evaluate and prevent or reduce these impacts. In one of those, life cycle assessment (LCA), ecotoxicological impacts of chemicals used in and emitted along product, service, and technology life cycles are characterized by their environmental fate, exposure, and effects in the life cycle impact assessment phase (Fantke et al., 2018; Jolliet et al., 2006). This enables selection of products with lowest expected impact of their use and

This article includes online-only Supporting Information.

* Address correspondence to melanie.douziech@agroscope.admin.ch and pefan@dtu.dk
Published online 11 June 2024 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/etc.5929

is an important metric to design benign products, evaluate product environmental footprints, and monitor the environmental improvements following the use of safer chemicals over time. It is thus key to try to cover all chemicals in LCA and avoid neglecting data-poor chemicals in practice because that could severely bias conclusions on most benign chemicals.

Successive harmonization efforts have led to a scientific global consensus model to characterize (eco-)toxic impacts of chemicals in LCA, the USEtox model (Henderson et al., 2011; Rosenbaum et al., 2008; Westh et al., 2015). Another framework aims to evaluate the absolute sustainability of chemical pressure based on this pressure in itself but also on the capacity of the ecosystem to withstand it (Kosnik et al., 2022). These methods complement others that start with the setting of protective environmental standards per chemical (e.g., the European Union's Registration, Evaluation, Authorisation and Restriction of Chemicals [REACH] regulation) and that can eventually be expressed as regional chemical footprints of unintended ambient mixtures (Bjørn et al., 2014; Zijp et al., 2014).

Common to most of these assessment frameworks is that the potential toxicity to species exposed to chemicals can be characterized using species sensitivity distributions (SSDs; Posthuma & De Zwart, 2012). An SSD describes the variation in sensitivity for a chemical for a set of tested species and allows for the estimation of the toxic pressure of a predicted or observed exposure concentration, expressed as the potentially affected fraction (PAF) of species. Various studies in aquatic ecosystems have shown that higher toxic pressure levels increase the difficulty of maintaining biodiversity at nonpolluted reference levels (see Posthuma et al., 2020).

For the ecotoxicity assessment, the required compound-specific SSDs are derived from measured or estimated test effect data for a set of species exposed to a certain chemical under laboratory test conditions. Recently, the Global Life Cycle Impact Assessment Method (GLAM) task force recommended that the effect factor (EF) used for ecotoxicity characterization in the impact assessment phase of LCA should be based on a concentration reflecting the chronic aspects of true ambient exposure levels. This concentration defines the "working point" on the SSD. That working point should thus be close to the domain of ambient concentrations. The 20th percentile response level of an SSD of measured or extrapolated chronic 10% effective concentration (EC10) values, $HC20_{EC10}$, with HC denoting the hazardous concentration, was therefore adopted as the current "working point." A chronic EC10 value is the concentration at which tested individuals of a species show an adverse effect of 10% on a common life-history trait (e.g., growth or reproduction) in comparison to nonexposed individuals (Owsianiak et al., 2023). This GLAM recommendation differs from the approach followed so far, where the 50th percentile response level of an SSD based on measured or extrapolated EC50 values was used as the previous working point, that is $HC50_{EC50}$, mostly derived from acute EC50 test data (Müller et al., 2017). That previous practice was based on the consideration that the amount of EC50 data is relatively high and that the estimation of the EC50 as an impact metric from a concentration–response curve is statistically most robust. Over time, it has been recognized that this "old" working point bears little relationship with ambient exposure levels.

To ensure that the SSD derived from test data sets for species from the studied ecosystem are both statistically robust and sufficiently representative of the species assemblages in field ecosystems and that the EFs derived from them are ecologically representative and statistically robust, various scholars have proposed minimum data requirements. The prescribed number of required species and species groups can vary between four and 10 data points per chemical for different use contexts (Nugegoda & Kibria, 2013). While different research teams derived different definitions for optimal approaches to derive SSDs, SSDs for more data-rich compounds are invariantly statistically more robust because they are less sensitive to adding or removing test data and likely more representative of field species assemblages.

A recent case study illustrated the use of the current GLAM recommendations for freshwater ecosystem–based SSDs on a minimum of five species from at least three distinct species groups (Owsianiak et al., 2023). Following these recommendations, $HC20_{EC10}$-based EFs could only be derived for 31 out of the 115 chemicals from that case study's inventory. This example demonstrates that currently available ecotoxicity data would allow assessment of ecotoxicity effects in LCA for relatively few compounds and would in turn lead to the unjustified neglect of the potential ecotoxicity of many chemicals, as highlighted in several other publications (Oginah et al., 2023; Posthuma et al., 2019; Saouter et al., 2019). The limited availability of measured effect concentrations also limits the use of SSDs in the other assessment frameworks (Dyer et al., 2008). We posit that it is better to avoid neglecting chemicals in comparative LCA contexts and provide EFs that can be (in part) based on additionally generated ecotoxicity data with their confidence intervals.

Filling data gaps with measured ecotoxicological effect data is not always desirable in the context of reduced animal testing and not always necessary. Alternatives to measured ecotoxicological effect data exist in the form of so-called in silico methods (Von Borries et al., 2023). The quantitative structure–activity relationship (QSAR) is an example of an in silico method that estimates the effect concentration based on the chemical's physicochemical properties. Regression equations extrapolating from one effect concentration to another either for the same species (*intraspecies*) or toward a different species (*interspecies*) are another example. With the LCA context as an example, the consequences for the derived SSD (and its decision support–related metrics, e.g., $HC20_{EC10}$) when the underlying SSD is based on a combination of measured data and in silico effect data are, however, unclear. Studies exist which evaluated single in silico tools for SSDs (see Douziech et al., 2020; Hoondert et al., 2019), but a systematic comparison of different in silico approaches and the consequences for the $HC20_{EC10}$ in terms of predictive power and uncertainty are currently missing. Furthermore, guidance on how to use these in silico methods to derive EFs for as many compounds as possible for supporting the impact assessment phase of practical LCA assessments is lacking. Finally, insights into the consequences of implementing the $HC20_{EC10}$ approach for scoring and ranking of chemicals as compared to the $HC50_{EC50}$ approach are also lacking, despite

them being needed to evaluate the novel consensus method and its adoption in practice. That is, if the change in working point and the addition of in silico data represent a robust method to characterize ecotoxicity, then the application of both the old and the new methods for an array of chemicals should yield a similar rank order of relative toxicity differences across an array of chemicals.

The aim of the present study was therefore to define a methodological approach for deriving compound-specific freshwater SSDs and $HC20_{EC10}$-based EFs for a large number of chemicals, based on different levels of available measured effect concentrations and different in silico methods. The method can be applied similarly to the various frameworks in which SSDs are used for other decision support practices. To illustrate how the use of extrapolation methods for addressing data-poor conditions works out, the developed approach was applied to derive a consistent set of freshwater ecotoxicity GLAM-compliant EFs for use in contemporary LCA. The resulting novel $HC20_{EC10}$ estimates were compared with older $HC50_{EC50}$ estimates, derived under more data-rich conditions, to evaluate the implications of the new GLAM recommendations when combined with in silico methods.

MATERIALS AND METHODS

Steps followed

The steps followed to derive EFs from different levels of available measured effect concentrations can be summarized in five steps (Figure 1).

First, the database of measured effect concentrations (DatabaseRaw) to be used was chosen, and two subsets were created. DatabaseRich included only data-rich chemicals with measured effect data for at least five species from three species groups. Reference SSDs were derived from these data-rich chemicals based on mean and standard deviation and $HC20_{EC10}$ estimated based on Owsianiak et al. (2023) and used as a comparative "anchor." DatabaseAll included all chemicals based on the steps detailed below (see *Database of measured effect concentrations*). Second, in silico methods to handle data-poor chemicals were identified (see section *In silico methods*). Third, the in silico methods were tested on DatabaseRich, by replacing measured values with predicted values, and thereupon comparing resulting SSDs, hypothesizing that a good in silico method would yield similar SSDs. The $HC20_{EC10}$ values based on different combinations of measured and in silico–based effect concentrations were therefore compared to the "anchor" $HC20_{EC10}$ derived from data-rich chemicals (see section *Evaluations across SSDs*). In brief, the usefulness of intraspecies extrapolation equations, QSARs, interspecies correlation estimation (ICE) equations, and an approach using an average slope of the SSDs was evaluated. Based on the findings of this third step, an approach to characterize data-poor chemicals was derived in the fourth step. Finally, $HC20_{EC10}$ values and EFs were estimated based on the defined approach for all chemicals in DatabaseAll, compared to $HC50_{EC50}$ based on the previous impact assessment of LCA (more data-rich) consensus models; and the uncertainty of the derived $HC20_{EC10}$ was estimated (see section *Estimating freshwater ecotoxicity EFs and their uncertainty*).

Database of measured effect concentrations

We used a harmonized and enriched version of the database of ecotoxicity test data published by Posthuma et al. (2019). This database was chosen because it is the most consistently curated set of ecotoxicity data for the purpose of SSD derivation available to date (Owsianiak et al., 2019). For the present evaluation, only measured effect data on freshwater species with a harmonized endpoint were used, representing 9862 chemicals and 118,720 effect data (=DatabaseRaw). Peer-reviewed test data and European Chemicals Agency (ECHA)–dossier data, representing industry-based chemical safety assessment test data, were kept to increase the number of species-specific measured effect data available per chemical. Ongoing research suggests that the subsets of peer-reviewed test data and ECHA-dossier data do not systematically differ, apart from differences in the composition of the tested species. The two data sources were therefore combined to improve the statistical SSD robustness (which increases with increasing data richness). Because the available ECHA data obtained in 2016 were only specified per species group, we assigned each dossier test to the most tested species, meaning *Daphnia magna* for tests on daphnids, *Pimephales promelas* for tests on fishes, and *Raphidocelis subcapitata* for tests on algae. Further, in case several effect data were available for the same chemical and the same species, the geometric mean of the available measured effect data was calculated to have a single effect data point for SSD derivation for one

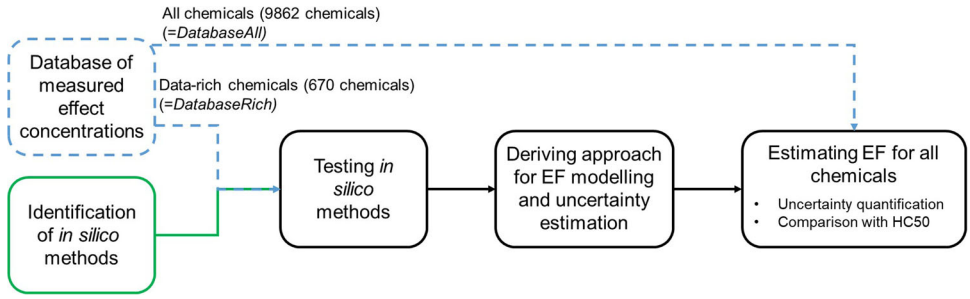

FIGURE 1: Steps followed to derive freshwater ecotoxicity effect factors based on different levels of available measured effect concentrations and their inclusion in a database: DatabaseRich and Database. All are subsets of the original database, as explained in the text. The number of chemicals included in each database is also specified. EF = effect factor; HC50 = 50% hazardous concentration.

chemical and species. The hierarchy followed in this case is explained in Supporting Information, S1, and is in line with current GLAM recommendations. The resulting database used to apply the approach developed in this paper (=DatabaseAll) had 71,605 data points for 9862 chemicals. The majority of the chemicals in DatabaseAll (90%, meaning 8894 chemicals) had fewer than five measured effect data for fewer than three species groups, and 95% (9395) of the chemicals in DatabaseAll had fewer than 10 species-specific measured effect data from three species groups (Supporting Information, S2). Thus, the number of compounds for which the GLAM-recommended assessment can be made is diminished by a lack of measured, compound-specific ecotoxicity test data for a sufficient variety (and number) of species.

For DatabaseRich, only the measured effect data from peer-reviewed sources were kept from DatabaseRaw, to ensure the reproducibility of the estimated $HC20_{EC10}$ and the representativeness for freshwater ecosystems. Peer-reviewed test data and ECHA-dossier data were excluded in DatabaseRich because it was not clear which species these effect data points referred to, so their reproducibility was not ensured. Data-rich chemicals were pragmatically defined for the present study as having at least five measured effect data from at least three species groups (Fantke et al., 2018; Müller et al., 2017; Rosenbaum et al., 2008). This minimum number was chosen, first, because $HC20_{EC10}$ values derived from SSDs composed of at least five data points do not require extrapolation of the $HC20_{EC10}$ beyond the available data (Owsianiak et al., 2019). Second, this threshold was chosen because use of five data points is not too limiting on the number of chemicals meeting this minimum number, while experience has shown that SSDs based on them are relatively robust (Oginah et al., 2023; Posthuma et al., 2019). Third, in the context of the impact assessment of freshwater ecotoxicity, the three species groups should cover algae, crustaceans, and fish (see Dong et al., 2016). In total, DatabaseRich included 10,144 data points for 670 chemicals based on measured effect data and was the starting point for the comparison of the effects of using ecotoxicity data generated with in silico methods. Using a more conservative approach to describe SSDs as data-rich when based on measured effect data for, for example, at least 10 species from three species groups would have nearly halved the size of DatabaseRich ($N = 341$; Supporting Information, S3). Requiring a higher number of data in this step would thus have had a substantial trade-off on the judgment of the effects of the extrapolation methods and was therefore not preferred. This observation further supports the pragmatic choice of using five species-specific effect data from three species groups as a threshold to derive reproducible and representative $HC20_{EC10}$ values for the evaluation of adding in silico effect data.

In silico methods

Four types of in silico methods were identified and tested in the present study. First, intraspecies endpoint extrapolations estimating another endpoint from one given endpoint, for example, a chronic EC10 estimated from a chronic no-observed-effect concentration (NOEC), were evaluated. The intraspecies endpoint extrapolation Equations (1) to (5), derived from the same database used in this paper and described in Oginah et al. (2023; Equations [1]–[5], all variables in micrograms per liter), were used.

$$\text{Log } EC10_{chronic} = (0.816 \times \log NOEC_{acute}) + 0.021 \quad (1)$$

$$\text{Log}EC10_{chronic} = (0.965 \times \log NOEC_{chronic}) - 0.144 \quad (2)$$

$$\text{Log}EC10_{chronic} = (0.869 \times \log EC50_{acute}) - 0.508 \quad (3)$$

$$\text{Log}EC10_{chronic} = (0.872 \times \log EC50_{chronic}) + 0.733 \quad (4)$$

$$\text{Log}EC10_{chronic} = (0.813 \times \log EC10_{acute}) + 0.967 \quad (5)$$

These endpoint extrapolation methods were chosen because they were based on a scientifically peer-reviewed methodology that used a recent, curated set of underlying effect data (Oginah et al., 2023). In addition, the ecotoxicity data were close to the chemical space of the chemicals included in DatabaseRaw and represented chemicals with measured effect data for at least three distinct species from at least three taxonomic groups.

Second, interspecies in silico equations or species extrapolations, meaning equations estimating an effect data for one species from the effect data of another species, were tested as an option to generate additional input data for deriving SSDs. More specifically, we used the ICE equations published by Raimondo et al. (2009). All available ICE equations were used because Bejarano et al. (2017) showed that including ICE equations with lower predictive accuracy barely influenced the representativity of the derived SSD. In fact, no applicability domain is provided for the ICE equations in the literature. The acute EC50 values estimated from the ICE equations were extrapolated to chronic EC10 values using Equation (3).

Third, we evaluated the representativeness of $HC20_{EC10}$ derived from an SSD based on a fixed slope derived as the average SSD slope across data-rich chemicals, $SSD_{\sigma \log EC10}$. The $SSD_{\sigma \log EC10}$ in Equation (6) was fixed to 0.7. Posthuma et al. (2002, 2019) have shown that the standard deviation of SSDs across chemicals that all had a larger number of underlying data points available tends to stabilize toward a slope of 0.7, whereas the slope under (far) more data-poor conditions can be very flat or steep simply by coincidence. As an example, the SSD derived from three highly similar test species can by coincidence be very steep, but this calculated value will change to less steep or steeper values by adding one or more test data of dissimilar species. Likewise, a coincidental flat slope tends to change toward a slope of 0.7 when adding species. For the fixed-slope approach, we employed this data-driven observed pattern on slope values as a function of the number of data points. That is, a log-normal distribution is hereby assumed, with the mean as the average of all measured effect data available and a fixed standard deviation of 0.7.

$$\text{Log } HC20_{EC10} = SSD_{\mu \log EC10} + (z_{0.2} \times SSD_{\sigma \log EC10}) \quad (6)$$

Finally, QSARs predicting the ecotoxicity of a chemical based on its physicochemical properties or molecular descriptors were evaluated for their ability to complement SSDs based

on limited measured effect concentrations. The QSARs used in our study were chosen based on criteria from previously published studies (Gramatica, 2007; Netzeva et al., 2007; Organisation for Economic Co-operation and Development [OECD], 2007). We focused the comparison on QSARs from the T.E.S.T consensus model estimating EC50 for *D. magna* and *P. promelas* (Martin, 2020) and from the European Union LIFE COMBASE project for *R. subcapitata* (Istituto di Ricerche Farmacologiche Mario Negri, 2018). The chosen QSAR methods have specific applicability domains. Only EC50s estimated for chemicals within the applicability domains of the QSARs were kept. To apply the QSARs, a simplified molecular-input line-entry system (SMILES) notation was automatically retrieved per chemical using the webchem R package (Szöcs et al., 2023). It should be noted that SMILES notations, required for some in silico methods, could not be retrieved for 43 chemicals out of the 670 reference compounds. Equation (3) was subsequently used to extrapolate the QSAR-based EC50 to chronic EC10 values. $HC20_{EC10}$ values were also calculated for combinations of QSARs and ICE-based effect concentrations. The acute EC50s estimated from the QSARs and ICE equations were extrapolated to chronic EC10s using Equation (3).

The uncertainty brought along by ICE- and QSAR-based effect concentrations in comparison to the uncertainty related to the limited amount of measured data points was thoroughly discussed in Douziech et al. (2020). They showed that adding sufficient in silico–based effect data to SSDs of data-poor chemicals reduced the overall uncertainty of the resulting SSDs, despite the additional modeling uncertainty brought along by these two in silico methods.

Evaluations across SSDs

For defined data-rich chemicals, an SSD from in vivo test data was derived using its mean across species-specific log EC10 values for this chemical as detailed in Owsianiak et al. (2023). Species sensitivity distributions for the data-rich chemicals were used as reference (see entry line 2 in Table 1) to comparatively characterize the (dis)similarity between the data-rich SSDs and the SSDs defined, wholly or partly, by systematically removing measured ecotoxicity data and adding ecotoxicity data generated by the aforementioned in silico methods. This procedure was only applied to the 670 chemicals in DatabaseRich, to provide insights into the effects of the investigated options to amend (replace original with extrapolated data) ecotoxicity data. We thereupon judged the effect of the change from original to extrapolated data because a "good" extrapolation method is expected to have a minimal effect on the SSD. A justification for using the $HC20_{EC10}$ based on measured chronic EC10 and intraspecies extrapolated measured effect data to chronic EC10 as reference for the comparisons (no. 2 in Table 1) is detailed in the *Results* section.

It was hypothesized that additions of predicted effect data for chemicals with insufficient measured test data for missing species would improve the representativeness of the resulting SSDs. This was judged by calculating a representativeness ratio (R_{repr} in Equation [7]) per chemical by comparing $HC20_{reference}$ (no. 2 in Table 1), based on at least five measured or extrapolated chronic EC10s, to $HC20_{estimated}$ based on a systematically reduced number of measured effect data complemented with in silico–based effect data (no. 3–9 in Table 1). This relies on the assumption that $HC20_{reference}$ of the data-rich chemicals is a good anchor to judge the $HC20_{EC10}$ of the extrapolation methods, just for the present (comparative) purpose. Because the addition of test data for a compound can pertain to any value of the novel data point, we removed data points and replaced them with extrapolated data of any of the methods a repeated amount of times, referred to as *iterations*, to get insight into patterns that can be realistically expected when adding data from novel measurements or from extrapolation.

TABLE 1: Summary of the 20% hazardous concentration ($HC20_{estimated}$) derived with the different in silico approaches and the HC20s used as reference values for comparisons (1 and 2)

No.	Name	Description
1	$HC20_{EC10,measured}$	Based only on measured chronic EC10 for at least the minimum required data set (≥5 test data from ≥3 taxonomic groups). Number of compounds = 10.
2	$HC20_{EC10,extrap}$	Based on measured chronic EC10 and measured effect data intraspecies extrapolated from selected endpoints toward chronic EC10. *Reference-value for all comparisons.* Number of compounds = 670.
SSDs derived upon data enrichment with various techniques (below) compared with the reference (no. 2).		
3	$HC20_{EC10,slope}$	Based on measured data and enriched with intraspecies endpoint extrapolated chronic EC10 and fixed slope $SSD_{\sigma \log\log EC10}$.
4	$HC20_{EC10,ICE}$	Based on measured data and enriched with extrapolated chronic EC10 and all ICE-based acute EC50 extrapolated to chronic EC10 using Equation (3).
5,6,7	$HC20_{EC10,QSAR}$	Based on measured data and enriched with extrapolated chronic EC10 and one, two, or three QSAR-based acute EC50 values extrapolated to chronic EC10 using Equation (3).
8	$HC20_{EC10,QSAR,ICE}$	Based on measured data and enriched with extrapolated chronic EC10 and three, two, or one QSAR-based acute EC50 values depending on applicability domain and all applicable ICE-based acute EC50s. The acute EC50s were extrapolated to chronic EC10 using Equation (3).
9	$HC20_{QSAR,ICE}$	Not based on measured data for any species but fully based on three, two, or one QSAR-based acute EC50 values depending on applicability domain and all applicable ICE-based acute EC50s using the acute EC50 based on QSAR. The acute EC50s were extrapolated to chronic EC10 using Equation (3).

EC10 = 10% effective concentration; ICE = interspecies correlation estimation; QSAR = quantitative structure–activity relationship; SSD = species sensitivity distribution.

$$R_{repr} = \frac{HC20_{reference}}{HC20_{estimated}} \quad (7)$$

A ratio of 1 implies a perfect match between the original (data-rich) $HC20_{EC10}$ and the ones resulting from removal of data and replacement by an extrapolated value(s), a ratio <1 a larger $HC20_{EC10}$ of the estimated versus the reference $HC20_{EC10}$ (lower hazard estimated), and a ratio >1 a smaller $HC20_{EC10}$ of the estimated versus the reference $HC20_{EC10}$ (higher hazard estimated). We compared $HC20_{EC10}$ based only on measured chronic EC10 (no. 1 in Table 1) to $HC20_{EC10}$ derived for the same chemicals but using two extrapolated effect data per species group for the chemicals in DatabaseRich. This was done to test the influence of extrapolated effect data on the $HC20_{EC10}$. Per iteration, extrapolated EC10s from measured data were randomly drawn, and an $HC20_{EC10}$ was calculated from them per chemical, as well as an R_{repr}. The 25th, 50th, and 75th percentiles and the minimum and maximum of the R_{repr} were then computed over the R_{repr} derived for the 10 chemicals with $HC20_{EC10,measured}$ over 1000 iterations.

Regarding the other in silico approaches, an increasing number of measured effect data was removed randomly per chemical, followed by adding extrapolated values generated by any of the methods (no. 3–9 in Table 1). An $HC20_{EC10,slope}$ was derived using the constant slope for the reduced set of measured effect data (no. 3 in Table 1). Further, $HC20_{EC10,ICE}$ values were estimated based on the reduced set of measured effect data systematically complemented by either all ICE estimates available for the remaining effect data (no. 4 in Table 1), one QSAR estimate (no. 5 in Table 1), two QSAR estimates (no. 6 in Table 1), three QSAR estimates (no. 7 in Table 1), or all ICE estimates available for the remaining effect data and the QSAR estimates (no. 8 in Table 1). In other words, the number of measured effect data was systematically reduced and complemented with in silico–based effect data, whereby ideal replacements would not change the SSD. In case several effect data were available for the same chemical and species, an average effect data point was derived according to the steps mentioned in Supporting Information, S1. Approaches 1 to 7 in Table 1 work only if at least one measured effect data point is available. Alternatively, we evaluated how $HC20_{EC10}$ derived using QSAR estimates as a basis for ICE equations (no. 9 in Table 1) compare to the $HC20_{EC10}$ of the reference data set. This procedure was repeated for all 670 chemicals 40 times to limit the computational expenses while still evaluating the robustness of the results. In addition, we quantified the applicability of each in silico approach as a percentage of chemicals to which the approach could be applied compared to the total of 670 and present the outcomes in the *Results* section.

Estimating freshwater ecotoxicity EFs and their uncertainty

The outcomes of the evaluations of the in silico methods were used to identify whether and which of these methods result in SSDs that resemble the data-rich ones and, if so, to define the approach that would be most suitable to estimate $HC20_{EC10}$ for data-poor chemicals. As an extra validation step, we compared the past ($HC50_{EC50}$, more data-rich) and novel ($HC20_{EC10}$, more data-poor and extrapolated) consensus methods for deriving EFs, which should result ideally in similar hazard rankings of the chemicals. That is, the $HC20_{EC10}$ values calculated for all chemicals of DatabaseAll were then compared to the former $HC50_{EC50}$ values derived following the approach in USEtox (Ver. 2.12). This was possible for 629 chemicals with at least five measured chronic or acute EC50s from three different species groups and a corresponding estimated $HC20_{EC10}$. Finally, once the $HC20_{EC10}$ values were calculated for all chemicals, the EFs were estimated using Equation (8) (Owsianiak et al., 2023).

$$EF = \frac{0.2}{HC20_{EC10}} \quad (8)$$

In addition to the calculation of the EFs, we estimated their uncertainty based on the approach presented in Oginah et al. (2023). The calculated uncertainty of the EFs can be used by end users to decide on using the various EFs in an LCA. The overall uncertainty around the calculated HC20 values is derived by combining the geometric standard deviation (GSD) for the interspecies variability, thus across available effect values, and for the intraspecies variability, thus around the effect values. We did not determine the GSD for data-poor chemicals (fewer than six records) because the limited number of data could imply strong bias. Instead, we used a fixed interspecies GSD calculated as the 97.5th percentile of estimated values across chemicals with six records. Similarly, we used a fixed intraspecies GSD for chemicals with one record, calculated as the 97.5th percentile of the estimated intraspecies GSD across chemicals with two records. To evaluate the robustness of the $HC20_{EC10}$ values obtained through various in silico methods, we set a practical upper limit criterion of total GSD approximately 2.23 based on experience with commonly accepted SSDs. The $HC20_{EC10}$ values were deemed robust if the uncertainty metric was below this criterion.

RESULTS

Performance of the in silico approaches

From the 670 chemicals kept in DatabaseRich, only 10 had measured chronic EC10s for at least five species from three species groups. The other chemicals had variable numbers of measured effect data for other endpoints. The ecotoxicity assessments method recently recommended in GLAM would therefore only be feasible for 10 out of 670 compounds if strictly based on the recommended measured chronic EC10 data (Owsianiak et al., 2023). Implicitly, this would neglect 660 compounds in LCA, even if those compounds would clearly cause a toxic pressure based on insights from the available ecotoxicity data. This provides a clear motive for the present study, to explore whether the extrapolation methods can be used to expand on the available ecotoxicity data, associated SSDs, and resulting EFs (including representing their uncertainty).

Intraspecies endpoint extrapolations. As a first step, we tested the usefulness of intraspecies extrapolation approaches (Equations [1]–[5]). The R_{repr} between the $HC20_{EC10}$ values based only on measured chronic EC10 and the $HC20_{EC10}$ values based also on extrapolated chronic EC10 derived from 1000 iterations over the 10 chemicals with sufficient measured chronic EC10 varied between 1.7 and 2.54. It should be noted that these values, despite relying on several iterations, only consider 10 chemicals. The range of R_{repr} (e.g., over the 25th and 75th percentiles of the iteration results) decreased with an increasing number of initial measured data points from 7 to 2 (Supporting Information, S4). These results resemble the pattern described by Posthuma et al. (2002), here represented by R_{repr}, closer to one with an increasing number of input data. These observations provide a data-driven insight, and reconfirm literature patterns, that the SSDs for chemicals with more effect data are more (statistically) robust (insensitive) to adding or removing test data. The R_{repr} values were >1, which implies that adding extrapolated effect data (line 2) to sets of data-poor chemicals (line 1) generally yielded a lower $HC20_{EC10}$ estimate. This suggests a slightly higher ecotoxicity potential at the exposure level considered in impact assessment in LCA. Apparently, the data sets for the data-poor chemicals are characterized by tests with species with relatively lower sensitivity, whereby the added (extrapolated) species data are apparently more sensitive. A comparison at the endpoint level is provided in Supporting Information, S4.

Based on these findings, the following comparative results rely on $HC20_{EC10}$ values based on measured chronic EC10 combined with intraspecies extrapolated measured effect data to chronic EC10 as anchors (cf. $HC20_{EC10,extrap}$, no. 2 in Table 1), expanding the number of data for evaluating the use of extrapolation methods from 10 to 670 chemicals. Although the use of the intraspecies extrapolation methods itself is an extrapolation step, we consider it key to help evaluate the effects of the other extrapolation methods for a large number of chemicals. Therefore, we utilized the latter data as reference.

Comparative outcomes of the in silico approaches. Figure 2 compares the R_{repr} obtained from $HC20_{EC10}$ values derived from the four in silico approaches, namely the average slope of 0.7 ($HC20_{EC10,slope}$); a combination of measured effect data and ICE estimates ($HC20_{EC10,ICE}$); or a combination of measured and QSAR-based estimates for a fish, a crustacean, and an algae species ($HC20_{EC10,QSAR}$).

The results are shown for chemicals with 1 to 20 measured data points. Note that applying in silico approaches for chemicals that are already composed of more than 20 measured data points had specific effects (Supporting Information, S5). While at a lower number of initial data points the addition of new species reduces the uncertainty band and likely improves the robustness of the SSD, this was not observed at a higher number of data points. This is likely attributable to the fact that the extrapolation does add new data but (at higher initial numbers of data) much less likely also new species with a specific higher or lower sensitivity. This implies that outcomes for higher numbers of initial species therefore will vary, dependent on the iteration and the original composition of the data.

Over the range of originally measured data points from 1 to 20, the resulting median R_{repr} for the extrapolated

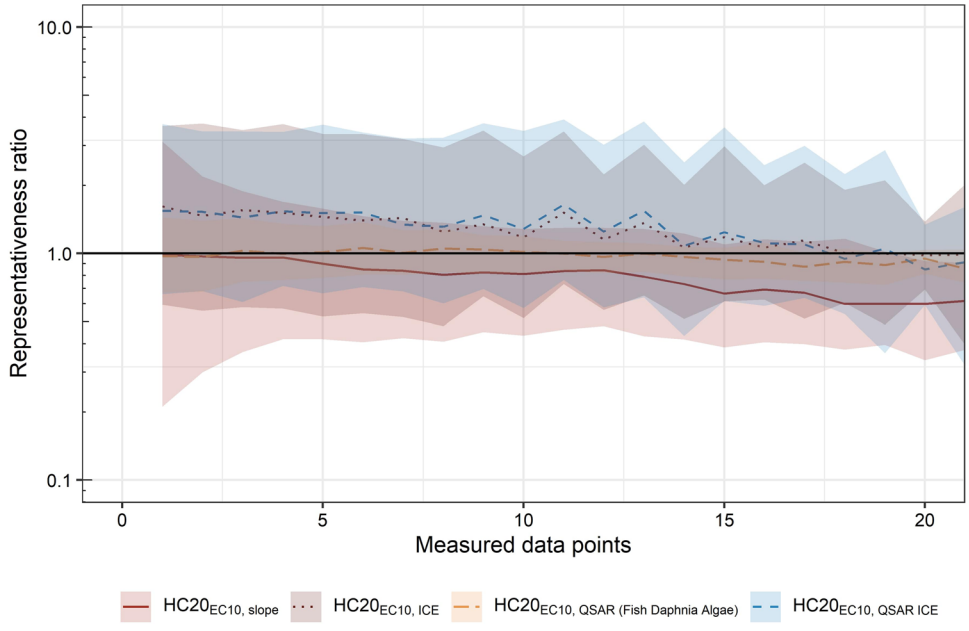

FIGURE 2: Median representativeness ratio between 20% hazardous concentration 10% effect concentration ($HC20_{EC10,extrap}$; the baseline or reference) and four alternative ways to estimate $HC20_{EC10}$ for data-poor chemicals, tested with data for 670 chemicals (lines and dotted lines). The baseline is derived from measured chronic EC10 and measured effect data extrapolated to EC10 (no. 2 in Table 1). The *y*-axis is a log-scale, and a perfect representativeness ratio corresponds to $y = 1$, marked as a horizontal black solid line. The *x*-axis shows the number of measured data points available per chemical prior to the remove and extrapolate/amend approaches. The shaded area corresponds to the 25th and 75th percentiles of 40 iterations. The average-slope approach was applied to 670, the interspecies correlation estimation (ICE) equations approach to 598, all quantitative structure–activity relationships (QSARs) approach to 142, and the combined QSAR and ICE approach to 417 chemicals.

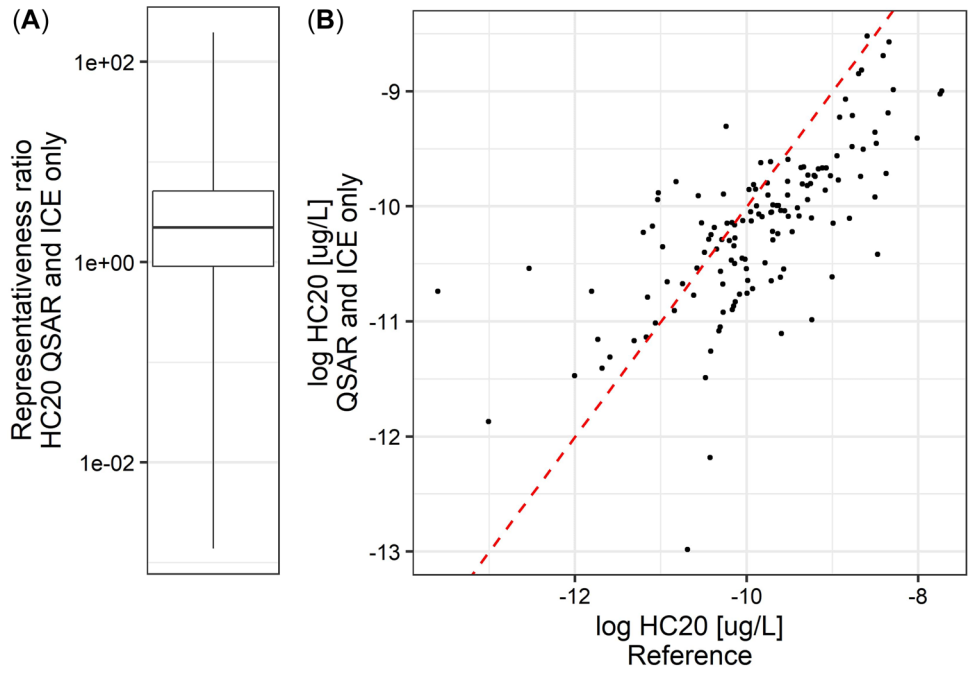

FIGURE 3: (**A**) Boxplot of the representativeness ratio between the 20% hazardous concentration (HC20)of the 10% effect concentration (EC10) derived from measured and extrapolated effect data (the reference anchor $HC20_{EC10,extrap}$) and derived only from in silico–based effect data (quantitative structure–activity relationship and interspecies correlation estimation equations; $n = 137$ chemicals) and (**B**) correspondence plot of the estimated (*y*-axis) versus reference (*x*-axis) $HC20_{EC10}$. QSAR = quantitative structure–activity relationship; ICE = interspecies correlation estimation.

$HC20_{EC10,slope}$ ranges between 0.59 and 0.98 (tendency for lowered hazard after extrapolation), for $HC20_{EC10,ICE}$ between 0.98 and 1.62 (tendency for higher hazard after extrapolation), for $HC20_{EC10,QSAR}$ between 0.88 and 1.06 (both lower and slightly higher hazard), and for $HC20_{EC10,QSAR,ICE}$ between 0.85 and 1.65 (tendency for higher hazard). In comparison, the R_{repr} between the $HC20_{EC10}$ based only on measured chronic EC10 and the $HC20_{EC10}$ based on extrapolated chronic EC10 derived from 1000 iterations over the 10 chemicals with sufficient measured chronic EC10 varied between 1.7 and 2.54.

As for the range around R_{repr} and the original number of ecotoxicty data, if only one measured data point was available, the ratio of the 75th to the 25th percentile for $HC20_{EC10,slope}$ was 14.8, while it was 6.2 and 5.6 for $HC20_{EC10,ICE}$ and $HC20_{EC10,QSAR,ICE}$, respectively. Consistently, but most for the constant-slope approach, addition of extrapolated data to very data-poor chemicals tends to result in a higher hazard estimate based on the complemented data. This ratio stayed rather stable with increasing number of measured data points for $HC20_{EC10,ICE}$ and $HC20_{QSAR,ICE}$ but reduced to approximately 3.2 for $HC20_{EC10,slope}$. That is, starting with more initial data, the addition of extrapolated data yields far more similar hazard insights (seven times lowered R_{repr} as compared to one initial data point). Starting with more data yields consistently better results.

Among the three methods relying on EC10 estimates, the $HC20_{EC10}$ that is estimated from a combination of measured and QSAR-based effect data ($HC20_{EC10,QSAR}$) leads to a small deviation of the R_{repr} from 1 and a relative constant spread, which only slightly reduces with an increasing number of measured initial data points. This indicates that this method, if applied to data-poor compounds, appears to result in the most consistent insights in the hazards of chemicals.

For $HC20_{EC10}$ based on measured and ICE-based effect data ($HC20_{EC10,ICE}$) and a combination of QSAR and ICE-based effect data ($HC20_{EC10,QSAR,ICE}$), the median R_{repr} $HC20_{EC10,extrap}$ is approximately 2 and reduces with an increasing number of measured data points. The opposite is the case for HC20 based on the mean of the measured effect data and a fixed slope of 0.7 ($HC20_{EC10,slope}$): the median R_{repr} is approximately 0.8 for fewer than five data points but increases with an increasing number of initial measured data points. This is important to note because the aim of the present study is to propose an approach for data-poor chemicals, whereby its application would yield good hazard insights.

The R_{repr} of the $HC20_{EC10,slope}$ is the only one consistently below 1, meaning an increasing overestimation of extrapolated $HC20_{EC10}$ values as compared to the reference $HC20_{EC10,extrap}$ with increasing number of measured effect data. Adding information on "missing species" on the basis of a fixed slope influences the SSDs differently than for the other in silico approaches, where the added information is not fixed by any systematic pattern apart from "more initial data is better."

Using one, two, or three species-specific QSARs to estimate $HC20_{EC10,QSAR}$ had little influence on the bias and spread thereof (Supporting Information, S6). The pattern was similar to the one observed in Figure 2 where lower initial numbers of ecotoxicity data implied a wider bandwidth of the R_{repr}, in turn implying less robust SSDs at a lower number of ecotoxicity data per chemical and no further reduction in the bandwidth above 10 measured effect data.

Figure 3 illustrates the R_{repr} resulting from combining QSARs and ICE-based effect data to estimate the $HC20_{EC10}$ for chemicals with no measured effect data ($HC20_{QSAR,ICE}$) as well as a scatterplot of the reference anchor ($HC20_{EC10,extrap}$) and the estimated $HC20_{EC10}$ ($HC20_{QSAR,ICE}$). The limited applicability domain of the used QSARs implied that an $HC20_{QSAR,ICE}$ based on effect data from at least five species and three species groups could only be estimated for 137 chemicals out of the 605 chemicals with <35 measured data points and available SMILES notation. The median bias is approximately 2, while the fraction of the 75th to the 25th percentile is approximately 5.6. The underestimation of estimated versus measured $HC20_{EC10}$ also appears from Figure 3.

Applicability of the in silico approaches. As shown in Table 2, using a default slope to derive an $HC20_{EC10}$ ($HC20_{EC10,slope}$) is the only method applicable to technically use this extrapolation method for all 670 chemicals (100%). The other in silico methods evaluated in the present study do not all reach this level of applicability. For example, the algae QSAR approach was only applicable to <25% of the 670 chemicals. The ICE-based approach is the method that adds the highest number of effect data to the total amount of effect data available to derive the $HC20_{EC10}$.

Approach to characterize data-poor chemicals

The comparative results, generated by studying 670 compounds with sufficient measured chronic EC10s, amended with intraspecies extrapolated values, was used to derive a proposal for generally applicable data processing steps to derive $HC20_{EC10}$-based EFs for data-rich and data-poor chemicals. Indeed, the comparative results showed a larger median R_{repr} for the approaches using ICE-based effect data as well as a larger spread in the computed R_{repr}, represented with the 25th and 75th percentiles (Figure 2). While the median R_{repr} closest to 1 was obtained from $HC20_{EC10,QSAR\ algae,fish,daphnia}$, its limited applicability (21.5%; Table 2) limits its use for a large range of chemicals. The average-slope approach with median R_{repr} of the $HC20_{EC10,slope}$ close to 1 and a spread comparable to the approaches using ICE-based effect data seems a promising alternative to estimate $HC20_{EC10,slope}$ for data-poor chemicals.

Based on these findings, the following generally applicable method is proposed to avoid neglecting ecotoxic chemicals in LCA and therefore to characterize the $HC20_{EC10}$ for data-poor chemicals in line with the GLAM3 recommendations:

1. In cases where several endpoints for the same chemical and species are available, chronic EC10s should be preferred, or alternatively the hierarchical approach presented in Supporting Information, S1, should be used.
2. Per chemical, the unique species-specific effect data should be extrapolated to chronic EC10s based on extrapolation equations.
3. Per chemical and species, in cases where several values for the chronic EC10 are available, meaning also potentially extrapolated values, their geometric mean should be calculated.
4. For chemicals with at least five effect data from at least three species groups, the $HC20_{EC10}$ can be calculated based on the approach in Owsianiak et al. (2023).
5. For chemicals with effect data for fewer than five species and three species groups:
   a. In case the SMILES notation is available and the application of QSAR can help reach effect data for exactly five species and three species groups, the approach in Owsianiak et al. (2023) should be applied for this extended data set to derive the $HC20_{EC10}$.
   b. In case the SMILES notation is not available or the application of QSAR cannot help reach effect data for exactly five species and three species groups, the default slope approach should be applied to derive $HC20_{EC10}$.

The uncertainty of the resulting EFs is further to be noted as a key part of the end result of the extrapolation process so that end users can decide on applying the extrapolation methods for the conditions of an LCA study.

These recommendations are assumed to be valid for all QSARs and extrapolation equations meeting the criteria specified in the *Materials and Methods* section (Gramatica, 2007; Netzeva et al., 2007; OECD, 2007) because those

TABLE 2: Maximum technical applicability (percentage of chemicals for which the in silico approach could be applied to judge its accuracy) of each in silico method together with the number of data points each in silico method adds to the available measured effect data

HC20	Technical applicability (%; for evaluating extrapolation performance)	Average no. of data points added (when applicable)
$HC20_{EC10,slope}$	100	Only the available measured effect data are used.
$HC20_{EC10,ICE}$	90.3	39 (Supporting Information S1, S7)
$HC20_{EC10,QSAR\ algae}$	22.2	1
$HC20_{EC10,QSAR\ fish}$	52.7	1
$HC20_{EC10,QSAR\ daphnia}$	55.7	1
$HC20_{EC10,QSAR\ algae,fish}$	21.5	2
$HC20_{EC10,QSAR\ algae,daphnia}$	21.6	2
$HC20_{EC10,QSAR\ fish,daphnia}$	52.1	2
$HC20_{EC10,QSAR\ algae,fish,daphnia}$	21.5	3

EC10 = 10% effective concentration; HC20 = 20% hazardous concentration; ICE = interspecies correlation estimation; QSAR = quantitative structure–activity relationship.

data-driven methods themselves are likely to further develop and improve over time. The use of the ICE in this hierarchical approach is not recommended at this stage.

Derived EFs

The recommended approach presented above in *Approach to characterize data-poor chemicals* was applied to the entire set of chemicals available in DatabaseAll (=9862). Including all measured effect data and ECHA-dossier data, 701 chemicals in DatabaseAll had effect data for at least five species from three species groups so that Step 4 of the approach in *Approach to characterize data-poor chemicals* could be applied. For 285 chemicals, QSAR estimates were added to the pool of measured effect data so that the approach in Owsianiak et al. (2023) could be used to derive the $HC20_{EC10}$ (=Step 5a of the approach in *Approach to characterize data-poor chemicals*). Finally, the derived EFs for the vast majority of the chemicals, meaning 8876 chemicals, could be based on the default slope approach (=Step 5b of the approach in *Approach to characterize data-poor chemicals*) because adding QSAR-based effect data was not sufficient to reach the minimum of five measured effect data from three species groups for these chemicals.

The EFs were expressed as PAF (cubic meters per kilogram; Figure 4), describing the volume of the freshwater compartment affected at the toxic pressure levels characterized as $HC20_{EC10}$ upon emitting 1 kg of chemical. The EFs span various orders of magnitude, implying that the ecotoxic potential of different chemicals differs vastly per unit emitted mass, as expected. The denomination PAF of species was kept in the unit to differentiate it from other damage metrics used in LCA, such as potentially disappeared fraction of species. The values of the EFs are provided in Supporting Information, S5.

The uncertainty analysis criterion was set at GSD total ≤2.23 based on experiences with regulation-accepted SSDs to characterize sufficiently robust SSDs based on EC10-equivalents and the $HC20_{EC10}$ values derived from those. As illustrated and described in relation to Figure 2, a robust SSD is defined as an SSD that is largely insensitive to changes regarding the addition or removal of test data, whereby nonrobustness occurs especially in case of limited initial test data. Removal or addition of test data may in such cases cause substantial but random effects on slope and position of the SSD. The uncertainty analysis findings for 9862 chemicals revealed that $HC20_{EC10}$ values derived from only sufficient and originally measured data exhibited the highest robustness ($HC20_{EC10,extrap}$), with 57% of chemicals meeting this criterion ($n = 401$ out of 701), followed by the $HC20_{EC10}$ values complemented by QSAR (29% of chemicals, $n = 83$ out of 285). The $HC20_{EC10}$ values derived from the default slope ($HC20_{EC10,slope}$) proved the least robust, with only 1% of chemicals meeting the criterion ($n = 132$ out of 8876). Regardless of the source of data used to compute the $HC20_{EC10}$ values, a higher number of initial ecotoxicity data generally improves SSD robustness (insensitivity to adding or removing test data), thus improving the SSD output, that is, $HC20_{EC10}$ and the EF (Figure 4).

Comparison between past and current modeling in life cycle impact assessment. To evaluate whether the latest GLAM recommendation to base the EFs on $HC20_{EC10}$ instead of $HC50_{EC50}$ implied a consistent shift in the EFs, and similar hazard ranking of the chemicals, the $HC50_{EC50}$ values were compared to the associated $HC20_{EC10}$ values for the 629 chemicals with both values (Figure 5). The results show a close to 1 slope and a constant factor of 0.07 between $HC20_{EC10}$ and $HC50_{EC50}$ with, as expected, systematically lower $HC20_{EC10}$ values compared to the $HC50_{EC50}$. This finding implies that

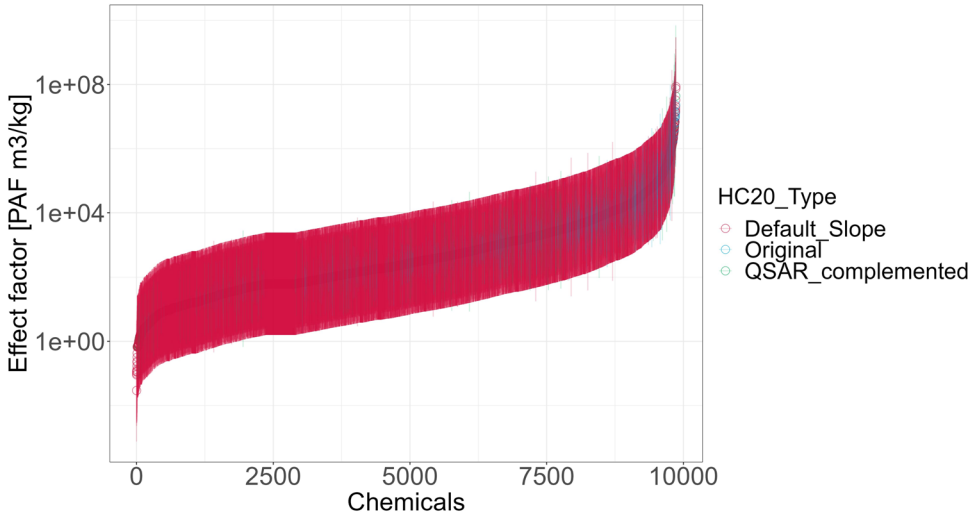

FIGURE 4: Effect factors (EFs; 20% hazardous concentration of the 10% effect concentration [$HC20_{EC10}$]) expressed as potentially affected fraction of species derived for all the chemicals in DatabaseCaseStudy (=9862). The uncertainty around each EF is presented as well depending on the approach followed to derive the $HC20_{EC10}$. *Default_Slope* stands for the default slope approach, *Original* is the approach proposed in the second global life cycle impact assessment method phase, and *QSAR_Complemented* are the $HC20_{EC10}$ values based on measured as well as quantitative structure–activity relationship-based EFs. PAF = potentially affected fraction.

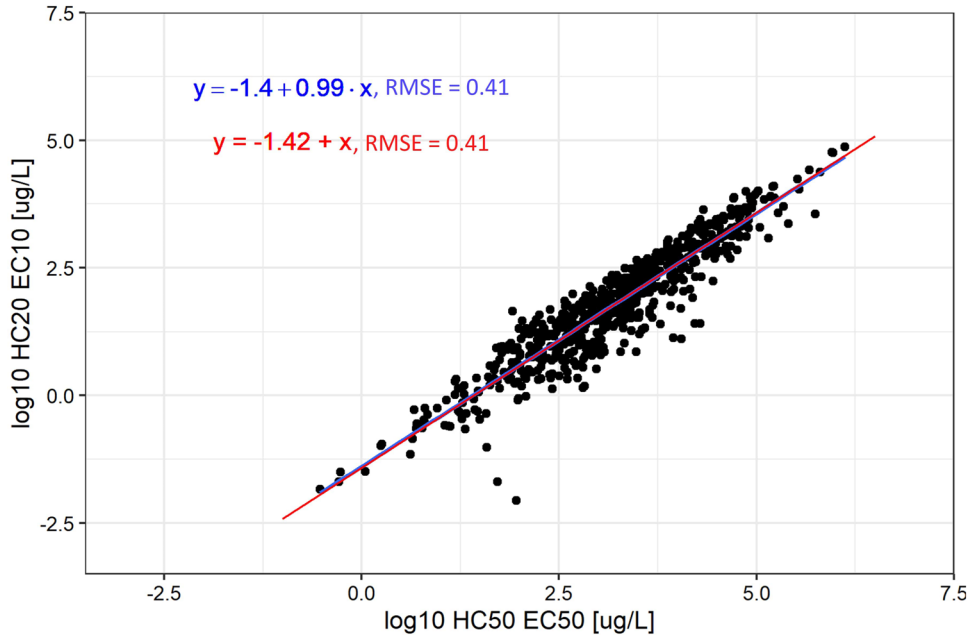

FIGURE 5: The 20% hazardous concentration of the 10% effect concentration ($HC20_{EC10}$) estimated from measured and extrapolated EC10 compared to $HC50_{EC50}$ estimated from measured and chronic EC50 and extrapolated acute EC50. Two linear regression equations are shown: one with a fixed slope of 1 (red) and another with best fit (blue). RMSE = root mean squared error.

application of the extrapolation methods, needed to derive $HC20_{EC10}$, did not cause substantial disturbance in the ranking of the ecotoxicity of chemicals: A higher $HC50_{EC50}$ (without extrapolations) covaries strongly with a higher $HC20_{EC10}$.

DISCUSSION

The current section discusses the applicability of our proposed approach to handle data-poor chemicals, addresses the limitations of the method, and lists future research needs.

Applicability of the approach for data-poor chemicals

Applied ecotoxicology is confronted with two issues: first, that a relative large amount of measured EC50s requires extrapolation to ambient exposure levels to obtain environmentally relevant outcomes; second, that a choice for lower-level effect data as alternative input for existing assessments (because they represent ambient concentrations) trades off with a far lower availability of such data. This limited availability of measured chronic EC10 makes the estimation of $HC20_{EC10,measured}$ on only measured EC10 data difficult and implies neglect of a large number of chemicals from existing assessment frameworks. This contrasts with the knowledge that many chemicals potentially pose ecotoxic risks and the societal requirement to evaluate the ecotoxic risks of a wide diversity of products, embodied in concepts such as environmental footprinting and the recently adopted Corporate Sustainability Reporting Directive (European Commission, 2022) or the European Union product environmental footprint (European Commission, 2021). The present study therefore aimed to propose an approach to derive $HC20_{EC10}$ values for data-poor chemicals relying on intra- and interspecies extrapolation equations.

Given the relatively good overlap of the $HC20_{EC10,measured}$ and the one based on extrapolated EC10 (bias close to 1), and moreover the observed close covariation of the past and current consensus methods for deriving EFs (Figure 5), we concluded that intraspecies extrapolated EC10 values can be used to derive $HC20_{EC10}$ values whenever (at least) five measured effect data from three species groups are not available. We used extrapolation methods listed in Equations (1) to (5) but assume that the outcome applies to other extrapolation equations that satisfy criteria for good extrapolation functions.

The in silico approaches used in the method proposed in *Approach to characterize data-poor chemicals* for data-poor chemicals have been chosen according to the findings in *Performance of the in silico approaches*. Overall, for chemicals with fewer than 20 measured effect data, adding QSAR-based effect data led to $HC20_{EC10}$ values close to the reference values with median differences between $HC20_{EC10,extrap}$ and $HC20_{EC10,QSAR}$ between 0.88 and 1.06 over the range of 1 to 20 measured data points. The usefulness of QSAR-based effect data to derive hazardous concentrations was also shown in other studies (Belanger et al., 2016; Lu et al., 2020). On the other hand, the median R_{repr} for $HC20_{EC10,ICE}$ ranged between 0.98 and 1.62 and for $HC20_{EC10,QSAR,ICE}$ between 0.85 and 1.65 over the range of measured data points from 1 to 20.

The restricted applicability domain of QSARs (<50% for the chemicals investigated in the present study) currently limits their use for more chemicals currently in trade. This limited applicability for QSARs is due, on the one hand, to the availability of SMILES notations and, on the other hand, to the applicability domain of the QSARs. For the former, SMILES notations could not be retrieved for 43 chemicals out of the

670. The latter relates directly to the "chemical space" covered by the training data set. This could be counterbalanced by QSARs with larger applicability domains or alternatively by choosing among the wide range of available QSARs the one tailored specifically for the considered chemical. Despite this, QSARs are likely to be not available for all existing chemicals, so it was necessary to include another in silico approach in the developed approach to ensure a broad chemical coverage.

The $HC20_{EC10}$ derived from measured and ICE-based effect data ($HC20_{EC10,ICE}$) had a median difference between 0.98 and 1.62 compared the to $HC20_{EC10,extrap}$ and was comparable when combining measured, ICE-based, and QSAR-based effect data ($HC20_{EC10,QSAR\ ICE}$). Bejarano et al. (2017) report a threefold difference between SSD-based HC5 estimates based on ICE-complemented SSDs and HC5 estimates relying only on measured effect data in 58% of the cases. The difference derived for $HC20_{QSAR\ ICE}$, so only QSAR and ICE-based effect data, was approximately 3. Following a similar approach, Douziech et al. (2020) reported a difference between 1 and 2 for HC50 based on QSAR and ICE and between 3 and 4 for an SSD-based estimate of HC5s based on QSAR and ICE effect data, thus comparable to the values in the present study. While more species can be included using ICE than what is currently done in assessments based solely on measured effect data, the observed difference for $HC20_{EC10,extrap}$ is a potential sign of species selection bias as discussed in *Limitations and future research*. On the contrary, the default slope method showed an average mean difference of 1.3 for data-poor chemicals (fewer than 20 measured data points). Based on this, the average-slope method was chosen over the combination of QSAR and ICE for application in our proposed approach, instead of the alternative of neglecting a chemical in impact assessment if its SSD was missing.

Limitations and future research

The proposed approach was derived on the basis of a large data set, although the available data were, in part, not yet ideal. For example, the implications of assigning ECHA-dossier data to the most tested species were not investigated in detail. One could alternatively have assumed a different species for each test entry. Given the link of ECHA data with legal requirements, the assumption that tests represent the most widely tested species seems, however, valid. Further, an ideal data set would represent all relevant test data globally collated. However, the data set used for the present study was slightly smaller than the data used by Posthuma et al. (2019) due to the removal of proprietary data.

Refinements of the average-slope method used as part of the approach to characterize data-poor chemicals might be necessary in the future, by either restricting the application of the average-slope approach or updating the average-slope value itself. In fact, the outcomes suggest that the slope of 0.7 is underestimated because it was determined based on data sets with restrictive biodiversity and overrepresentation of most commonly tested species such as *D. magna*. Oginah et al. (2023), for example, obtained a slope of 0.9 on their data set on data-rich chemicals that, once confirmed in a larger data set, might be considered to avoid this bias. This would require sufficiently rich data sets for chemicals with either narcotic or any of the specific modes of action so as to investigate robust slopes (due to data richness) in relation to subgroups and/or all tested species in separate or overall SSDs. Such data are, however, not yet available (Oginah et al., 2023).

An aspect that therefore deserves particular attention in the future is the question around the potential bias introduced by the focus of current ecotoxicological tests on a limited number of species. *Species selection bias* is defined as the bias that results from systematic, or random, trends that appear to have occurred in the collection of test data for particular species, in comparison to the diversity of species representing aquatic life forms. This may be a random process, whereby researchers happen to test some species and species groups, as well as chemicals and chemical groups, more than others (Gustavsson et al., 2017; Kristiansson et al., 2021). But it may also be the systematic result of a policy-requested minimum set of tested species, such as the algae, daphnia, fish (ADF) minimum triplet for European Union chemical safety assessments under the REACH regulation. Such species selection bias may be helpful to characterize and rank chemical safety in a standard way when evaluating potential hazards of chemicals to be judged for market entry but may counter the representativity of SSDs to cover the biodiversity of field species assemblages, with an effect on slope for the compounds currently studied. Such species selection bias has been discussed previously in relation with ICE estimates (Golsteijn et al., 2012) and is also reflected by the QSARs available especially for the ADF test species triplet. This species selection bias and its potential implications on the differences between $HC20_{EC10}$ were not further investigated in the present study. The species selection bias was, however, a likely explanation for the higher $HC20_{EC10}$ estimated from the fixed-slope extrapolation compared to raw data–driven SSDs.

The present study focuses on the ecotoxicological impacts on freshwater ecosystems—the most commonly addressed type of ecosystem in both ecotoxicity testing and therefore practices such as LCA. As next steps, we recommend addressing additional ecosystems, such as terrestrial, for which the problem of lacking measured effect data is probably even larger than for freshwater ecosystems (Li et al., 2023).

Despite including several approaches to estimate ecotoxicity data for chemicals and comparing them in a systematic way, we could not reflect the entire breadth of in silico approaches available. The approaches for deriving SSDs from data is an ever-evolving field of research. For example, while QSAR-based SSDs were evaluated, approaches like that in Hoondert et al. (2019), where the per-chemical median and standard deviation of the log-normal SSDs themselves were directly estimated from those of data-rich chemicals, were not. Expanding on these per-chemical methods can be done not only with the methods of the present paper but also with more overarching methods in which larger parts of the available ecotoxicity data are used at the level of EC10 of

individual species as well as directly at the SSD level of, for example, $HC50_{EC50}$ (Hou, Jolliet, et al., 2020; Hou, Zhao, et al., 2020). Multichemical data (or the whole database) can be used to characterize the ecotoxicity of a data-poor chemical. Building on earlier research at the Dutch National Institute for Public Health and the Environment (Notenboom et al., 1995), we can refer to this as *quantitative species sensitivity relationships* (QSSRs), which can currently be based on describing such patterns. Promising results are currently being obtained, exploring such QSSRs. Among others, such methods might help address chemicals with no measured effect concentrations (no data) at all, which were not addressed in the present study.

CONCLUSIONS AND RECOMMENDATIONS

The approach proposed in the present study derived EC10-based SSDs and estimated $HC20_{EC10}$ values and their uncertainty from them for 9862 chemicals, including data-poor chemicals. The approach proposed is compliant with the current consensus approach for the ecotoxicological impact assessment in LCA. The explicit quantification of uncertainty allows LCA assessors to set criteria for their specific assessment condition, allowing, for example, for higher uncertainty in exploratory and lower uncertainty for decision-related settings. This way, insight is provided into the ecotoxic potential of more chemicals emitted from product systems at field-relevant exposure concentrations, and it is possible to account for their potential impact in assessment methods relying on SSDs. This was possible following three hierarchical steps: (1) intraspecies extrapolating effect data to estimate EC10 from other test endpoints, (2) using QSAR approaches to reach five effect data from three species groups, or—if those are lacking—(3) assuming a set slope of 0.7 to derive $HC20_{EC10}$. A comparison of the estimated $HC20_{EC10}$ with EFs being defined from SSD-EC50 showed that these covary closely and so support the adoption of these recommendations. In fact, extensive studies on contemporary ambient exposures have shown that chemical concentrations in surface waters do not reach the $HC50_{EC50}$ level, whereas the concentrations expressed as $HC20_{EC10}$ are in a more realistic range (see Rorije et al., 2022). We conclude that the approach to derive SSDs and EFs developed in the present study, using measured chronic EC10 when available or relying on intraspecies extrapolations, QSARs, and an average-slope approach when not, covers the main effects in a relevant way and can be considered scientifically robust. The choice and use of in silico methods are namely supported by the empirical evidence in the present study based on the most recent database of measured effect data available so far and align with published literature.

Supporting Information—The Supporting Information is available on the Wiley Online Library at https://doi.org/10.1002/etc.5929.

Conflict of Interest—The authors declare no conflict of interest.

Author Contribution Statement—**Mélanie Douziech**: Conceptualization; Methodology; Software; Validation; Writing–original draft. **Susan Anyango Oginah**: Software; Data curation; Writing–review & editing. **Laura Golsteijn, Michael Zwicky Hauschild, Olivier Jolliet, Mikołaj Owsianiak**: Conceptualization; Writing–review & editing. **Leo Posthuma, Peter Fantke**: Conceptualization; Writing–original draft; Methodology.

Data Availability Statement—The R scripts used to carry out the analysis presented in this article are not available online. The calculated EFs for 9862 chemicals are available in the Supporting Information.

Acknowledgments—Open access funding provided by Agroscope.

REFERENCES

Bejarano, A. C., Raimondo, S., & Barron, M. G. (2017). Framework for optimizing selection of interspecies correlation estimation models to address species diversity and toxicity gaps in an aquatic database. *Environmental Science & Technology*, *51*(14), 8158–8165. https://doi.org/10.1021/acs.est.7b01493

Belanger, S. E., Brill, J. L., Rawlings, J. M., & Price, B. B. (2016). Development of acute toxicity quantitative structure activity relationships (QSAR) and their use in linear alkylbenzene sulfonate species sensitivity distributions. *Chemosphere*, *155*, 18–27. https://doi.org/10.1016/j.chemosphere.2016.04.029

Bjørn, A., Diamond, M., Birkved, M., & Hauschild, M. Z. (2014). Chemical footprint method for improved communication of freshwater ecotoxicity impacts in the context of ecological limits. *Environmental Science & Technology*, *48*(22), 13253–13262. https://doi.org/10.1021/es503797d

Dong, Y., Rosenbaum, R. K., & Hauschild, M. Z. (2016). Assessment of metal toxicity in marine ecosystems: Comparative toxicity potentials for nine cationic metals in coastal seawater. *Environmental Science & Technology*, *50*(1), 269–278. https://doi.org/10.1021/acs.est.5b01625

Douziech, M., Ragas, A. M., van Zelm, R., Oldenkamp, R., Hendriks, A. J., King, H., Oktivaningrum, R., & Huijbregts, M. A. (2020). Reliable and representative in silico predictions of freshwater ecotoxicological hazardous concentrations. *Environment International*, *134*, Article 105334.

Dyer, S. D., Versteeg, D. J., Belanger, S. E., Chaney, J. G., Raimondo, S., & Barron, M. G. (2008). Comparison of species sensitivity distributions derived from interspecies correlation models to distributions used to derive water quality criteria. *Environmental Science & Technology*, *42*(8), 3076–3083. https://doi.org/10.1021/es702302e

European Commission. (2021). *European climate law*. https://ec.europa.eu/clima/eu-action/european-green-deal/european-climate-law_en

European Commission. (2022). *Directive (EU) 2022/2464 of the European Parliament and of the Council of 14 December 2022 amending Regulation (EU) No 537/2014, Directive 2004/109/EC, Directive 2006/43/EC and Directive 2013/34/EU, as regards corporate sustainability reporting*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32022L2464

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., & Hauschild, M. (2018). Toward harmonizing ecotoxicity characterization in life cycle impact assessment: Harmonizing ecotoxicity characterization in LCIA. *Environmental Toxicology and Chemistry*, *37*(12), 2955–2971. https://doi.org/10.1002/etc.4261

Golsteijn, L., van Zelm, R., Veltman, K., Musters, G., Hendriks, A. J., & Huijbregts, M. A. (2012). Including ecotoxic impacts on warm-blooded predators in life cycle impact assessment. *Integrated Environmental Assessment and Management*, *8*(2), 372–378. https://doi.org/10.1002/ieam.269

Gramatica, P. (2007). Principles of QSAR models validation: Internal and external. *QSAR & Combinatorial Science*, *26*(5), 694–701. https://doi.org/10.1002/qsar.200610151

Gustavsson, M. B., Hellohf, A., & Backhaus, T. (2017). Evaluating the environmental hazard of industrial chemicals from data collected during the REACH registration process. *Science of the Total Environment*, *586*, 658–665. https://doi.org/10.1016/j.scitotenv.2017.02.039

Henderson, A. D., Hauschild, M. Z., Van De Meent, D., Huijbregts, M. A. J., Larsen, H. F., Margni, M., McKone, T. E., Payet, J., Rosenbaum, R. K., & Jolliet, O. (2011). USEtox fate and ecotoxicity factors for comparative assessment of toxic emissions in life cycle analysis: Sensitivity to key chemical properties. *The International Journal of Life Cycle Assessment*, *16*(8), 701–709. https://doi.org/10.1007/s11367-011-0294-6

Hoondert, R. P. J., Oldenkamp, R., de Zwart, D., van de Meent, D., & Posthuma, L. (2019). QSAR-based estimation of species sensitivity distribution parameters: An exploratory investigation. *Environmental Toxicology and Chemistry*, *38*(12), 2764–2770. https://doi.org/10.1002/etc.4601

Hou, P., Jolliet, O., Zhu, J., & Xu, M. (2020). Estimate ecotoxicity characterization factors for chemicals in life cycle assessment using machine learning models. *Environment International*, *135*, Article 105393. https://doi.org/10.1016/j.envint.2019.105393

Hou, P., Zhao, B., Jolliet, O., Zhu, J., Wang, P., & Xu, M. (2020). Rapid prediction of chemical ecotoxicity through genetic algorithm optimized neural network models. *ACS Sustainable Chemistry & Engineering*, *8*(32), 12168–12176. https://doi.org/10.1021/acssuschemeng.0c03660

Istituto di Ricerche Farmacologiche Mario Negri. (2018). *VEGA HUB—Virtual models for property: Evaluation of chemicals within a global architecture*. https://www.vegahub.eu/

Jolliet, O., Rosenbaum, R. K., Chapman, P. M., McKone, T. E., Margni, M., Scheringer, M., van Straalen, N. M., & Wania, F. (2006). Establishing a framework for life cycle toxicity assessment: Findings of the Lausanne Review Workshop. *The International Journal of Life Cycle Assessment*, *11*(3), 209–212.

Kosnik, M. B., Hauschild, M. Z., & Fantke, P. (2022). Toward assessing absolute environmental sustainability of chemical pollution. *Environmental Science & Technology*, *56*(8), 4776–4787. https://doi.org/10.1021/acs.est.1c06098

Kristiansson, E., Coria, J., Gunnarsson, L., & Gustavsson, M. (2021). Does the scientific knowledge reflect the chemical diversity of environmental pollution?—A twenty-year perspective. *Environmental Science & Policy*, *126*, 90–98. https://doi.org/10.1016/j.envsci.2021.09.007

Li, T., Cui, L., Xu, Z., Liu, H., Cui, X., & Fantke, P. (2023). Micro- and nanoplastics in soil: Linking sources to damage on soil ecosystem services in life cycle assessment. *Science of the Total Environment*, *904*, Article 166925. https://doi.org/10.1016/j.scitotenv.2023.166925

Lu, B.-Q., Liu, S.-S., Wang, Z.-J., & Xu, Y.-Q. (2020). Conlecs: A novel procedure for deriving the concentration limits of chemicals outside the criteria of human drinking water using existing criteria and species sensitivity distribution based on quantitative structure–activity relationship prediction. *Journal of Hazardous Materials*, *384*, Article 121380. https://doi.org/10.1016/j.jhazmat.2019.121380

Martin, T. M. (2020). User's guide for T.E.S.T. (Toxicity Estimation Software Tool). https://www.epa.gov/sites/default/files/2016-05/documents/600r16058.pdf

Müller, N., De Zwart, D., Hauschild, M., Kijko, G., & Fantke, P. (2017). Exploring REACH as a potential data source for characterizing ecotoxicity in life cycle assessment. *Environmental Toxicology and Chemistry*, *36*(2), 492–500. https://doi.org/10.1002/etc.3542

Netzeva, T., Pavan, M., & Worth, A. (2007). *Review of data sources, QSARs, and integrated testing strategies for aquatic toxicity* (EUR 22943 EN). Joint Research Center.

Notenboom, J., Vaal, M. A., & Hoekstra, J. A. (1995). Using comparative ecotoxicology to develop quantitative species sensitivity relationships (QSSR). *Environmental Science and Pollution Research*, *2*(4), 242–243. https://doi.org/10.1007/BF02986776

Nugegoda, D., & Kibria, G. (2013). Water quality guidelines for the protection of aquatic ecosystems. In J.-F. Férard & C. Blaise (Eds.), *Encyclopedia of aquatic ecotoxicology* (pp. 1177–1196). Springer. https://doi.org/10.1007/978-94-007-5704-2_105

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. *Environmental Science & Technology*, *57*(39), 14526–14538. https://doi.org/10.1021/acs.est.3c04968

Organisation for Economic Co-operation and Development. (2007). *Guidance document on the validation of (quantitative) structure–activity relationship [(Q)SAR] models* (Series on Testing and Assessment 69; ENV/JM/MONO[2007]2). https://www.oecd.org/env/guidance-document-on-the-validation-of-quantitative-structure-activity-relationship-q-sar-models-9789264085442-en.htm

Owsianiak, M., Fantke, P., Posthuma, L., Saouter, E., Vijver, M. G., Backhaus, T., Schlekat, T., & Hauschild, M. (2019). Ecotoxicity. In J. Lynch (Ed.), *Global guidance for life cycle impact assessment indicators* (Vol. *2*, pp. 139–172). United Nations Environment Programme.

Owsianiak, M., Hauschild, M. Z., Posthuma, L., Saouter, E., Vijver, M. G., Backhaus, T., Douziech, M., Schlekat, T., & Fantke, P. (2023). Ecotoxicity characterization of chemicals: Global recommendations and implementation in USEtox. *Chemosphere*, *310*, Article 136807. https://doi.org/10.1016/j.chemosphere.2022.136807

Posthuma, L., & De Zwart, D. (2012). Predicted mixture toxic pressure relates to observed fraction of benthic macrofauna species impacted by contaminant mixtures. *Environmental Toxicology and Chemistry*, *31*(9), 2175–2188. https://doi.org/10.1002/etc.1923

Posthuma, L., Suter, G. W., & Traas, T. P. (2002). *Species sensitivity distributions in ecotoxicology*. Lewis.

Posthuma, L., van Gils, J., Zijp, M. C., van de Meent, D., & de Zwart, D. (2019). Species sensitivity distributions for use in environmental protection, assessment, and management of aquatic ecosystems for 12,386 chemicals. *Environmental Toxicology and Chemistry*, *38*(4), 905–917. https://doi.org/10.1002/etc.4373

Posthuma, L., Zijp, M. C., De Zwart, D., Van De Meent, D., Globevnik, L., Koprivsek, M., Focks, A., Van Gils, J., & Birk, S. (2020). Chemical pollution imposes limitations to the ecological status of European surface waters. *Scientific Reports*, *10*(1), Article 14825. https://doi.org/10.1038/s41598-020-71537-2

Raimondo, S., Lilavois, C., & Barron, M. G. (2009). *Interspecies correlation estimation manual v1.0*. US Environmental Protection Agency.

Rorije, E., Wassenaar, P. N. H., Slootweg, J., Van Leeuwen, L., Van Broekhuizen, F. A., & Posthuma, L. (2022). Characterization of ecotoxicological risks from unintentional mixture exposures calculated from European freshwater monitoring data: Forwarding prospective chemical risk management. *Science of the Total Environment*, *822*, Article 153385. https://doi.org/10.1016/j.scitotenv.2022.153385

Rosenbaum, R. K., Bachmann, T. M., Gold, L. S., Huijbregts, M. A., Jolliet, O., Juraske, R., Koehler, A., Larsen, H. F., MacLeod, M., & Margni, M. (2008). USEtox—The UNEP-SETAC toxicity model: Recommended characterisation factors for human toxicity and freshwater ecotoxicity in life cycle impact assessment. *The International Journal of Life Cycle Assessment*, *13*(7), Article 532.

Saouter, E., Biganzoli, F., Pant, R., Sala, S., & Versteeg, D. (2019). Using REACH for the EU environmental footprint: Building a usable ecotoxicity database, Part I. *Integrated Environmental Assessment and Management*, *15*(5), 783–795. https://doi.org/10.1002/ieam.4168

Szöcs, E., Allaway, R., Muench, D., Ranke, J., Scharmüller, A., Scott, E. R., Stanstrup, J., Cavalcante, J. V. F., Getzinger, G., Bass, E., & Stirling, T. (2023). *webchem: Chemical information from the Web* (Version 1.3.0). https://cran.r-project.org/web/packages/webchem/index.html

Von Borries, K., Holmquist, H., Kosnik, M., Beckwith, K. V., Jolliet, O., Goodman, J. M., & Fantke, P. (2023). Potential for machine learning to address data gaps in human toxicity and ecotoxicity characterization. *Environmental Science & Technology*, *57*(46), 18259–18270. https://doi.org/10.1021/acs.est.3c05300

Westh, T. B., Hauschild, M. Z., Birkved, M., Jørgensen, M. S., Rosenbaum, R. K., & Fantke, P. (2015). The USEtox story: A survey of model developer visions and user requirements. *The International Journal of Life Cycle Assessment*, *20*(2), 299–310. https://doi.org/10.1007/s11367-014-0829-8

Zijp, M. C., Posthuma, L., & van de Meent, D. (2014). Definition and applications of a versatile chemical pollution footprint methodology. *Environmental Science & Technology*, *48*(18), 10588–10597. https://doi.org/10.1021/es500629f

