Ecotoxicological effect factors for calculating USEtox ecotoxicity characterization factors

Rahul Aggarwal[1*]

Abstract

Background This study aims to tackle the lack of freshwater ecotoxicological effect factors (EFs) crucial for determining freshwater ecotoxicity characterization factors (CFs) using the widely accepted scientific consensus USEtox model for ecotoxicity impact characterization. The objectives are: (1) to offer a collection of experimental EFs to support USEtox ecotoxicity characterization factor computations and (2) to contrast ecotoxicity data produced by various quantitative structure–activity relationship (QSAR) models against experimental data.

Results Experimental ecotoxicity data were gathered from the REACH database and CompTox Version 2.1.1, which includes toxicity information from ToxValDB v9.4. QSAR-driven ecotoxicity data were extracted from ECOSAR v1.11 and T.E.S.T. v5.1.2. The experimental and estimated data underwent a harmonization process to ensure consistency. Subsequently, aquatic ecotoxicological EFs were determined. The merged REACH and CompTox databases list EFs for 11,295 substances, each identified by a unique CAS number. Among these, the USEtox database already catalogs 2426 substances with freshwater ecotoxicological EFs. This study expanded on that by calculating EFs for an additional 8869 substances. Using estimated data, EFs were determined for 6029 chemicals based on ECOSAR data and 6762 chemicals using TEST data.

Conclusions This study calculated EFs for an additional 8869 substances, thereby broadening their inclusion in LCA evaluations. When integrated with the USEtox EFs database, this research encompasses 11,368 chemicals. The high correlation observed between experimental EFs and those in the USEtox database lends significant confidence to the calculations for chemicals not listed in USEtox. Conversely, the low correlation between estimated EFs and those in USEtox suggests limited confidence in calculations based on estimated data. Furthermore, the disparity in correlations between EFs calculated using ECOSAR and TEST indicates that different QSARs can yield varied results. This discrepancy underscores the need for caution when relying on estimated data. Given that EFs are contingent on data availability, it is imperative to periodically update EFs as new data emerges.

Keywords Ecotoxicity, USEtox, Characterization factor, Ecotoxicological effect factor

Background

In the current era of rapid global development and escalating consumption, there is a growing reliance on chemicals essential to nearly all manufacturing products [19, 51]. This has fueled population growth and advancements in sectors, such as energy and food production, yet it also poses significant risks to human and ecological health, including toxic pollution and species extinction [3, 5, 29, 51]. In European waters, organic chemicals are major contributors to biodiversity loss and ecosystem service degradation [24, 55]. Despite recognizing chemical pollution as a critical threat to sustainability, our understanding of its full impact is still emerging [10, 54]. Furthermore, this issue compounds other global challenges, such as climate change and biosphere integrity, with recent assessments

*Correspondence:
Rahul Aggarwal
rahula@chalmers.se
[1] Environmental Systems Analysis, Chalmers University of Technology, Vera Sandbergs Allé 8, 41296 Gothenburg, Sweden

indicating that the pace of chemical production and emission outstrips our capacity to manage [29, 34, 46].

Since the 1800s, the Chemical Abstracts Service (CAS) has documented over 204 million substances, with approximately 350,000 chemicals or combinations listed in national or regional inventories; notably, the European Chemicals Agency (ECHA) has over 26,500 chemical registrations [8, 11, 54]. Increasing production and use of chemicals raises concerns, as lessons from legacy pollutants such as PCBs and mercury highlight the risks of widespread chemical use [28]. Internationally, the Globally Harmonized System (GHS) has been adopted for consistent chemical classification and labeling, categorizing chemicals by properties, health impacts, and environmental hazards [32]. Moreover, the rising concern over Per- and polyfluoroalkyl substances (PFAS) has led to calls for their class-wide regulation. PubChem, one of the world's largest chemical databases, now includes over 7 million PFAS compounds following the OECD revised definition [45].

Frameworks for chemical management and assessment, such as life cycle impact assessment (LCIA), chemical alternatives assessment (CAA), comparative risk screening, and risk assessment, play a crucial role in assessing the toxicological effects of chemical exposures [4, 14, 27]. Life Cycle Assessment (LCA) tools, adhering to standards, such as ISO 14040, are increasingly recognized for quantifying potential toxicity throughout a product's lifecycle [13, 22, 35, 47]. However, comprehensive toxicity quantification relies on having characterization factors (CFs) for every chemical emission [16]. To determine CFs, modeling chemical environmental fate, exposure, and effects is crucial [36]. The USEtox model offers a fate-exposure-effect matrix framework for calculating freshwater ecotoxicity CFs, integrating fate factors (FF), exposure factors (XF), and effect factors (EF) [12]. USEtox uses chronic ecotoxicological data, typically the HC50, to calculate quantitative EFs for chemicals [12, 35].

Although numerous chemicals pose risks to ecosystems, only a small percentage have been thoroughly assessed for LCA, resulting in significant inaccuracies and gaps [37]. This is highlighted by the vast difference between the 145,299 substances listed by the ECHA (ECHA, 2023a) but only 3104 substances (3077 organic and 27 inorganic) detailed in the USEtox database (version 2.01) [12]. With USEtox providing freshwater ecotoxicity EFs for only 2499 substances, a large number of chemicals remain uncharacterized, primarily due to scarce EC50 chronic data.

In recent years, the USEtox consortium and the European Commission have made significant efforts in refining the methodology for calculating EFs and expanding the range of chemicals within the LCA framework by computing EFs for a broader array of chemicals [30, 38]. Such advancements have broadened the chemical scope within the Product Environmental Footprint (PEF), facilitating more comprehensive comparative life cycle analyses in Europe [41–43]. Saouter et al. [39, 40] highlighted the influence of selecting ecotoxicological data sources and toxicological endpoints on the USEtox EFs, underscoring the need for improved substance information within USEtox. Furthermore, Saouter et al. [41] incorporated aquatic toxicity hazard values from REACH, EFSA, and PPDB, yielding HC20 values from REACH/EFSA for 6764 chemicals and from PPDB for 1316 chemicals. However, this newly derived HC20 data has yet to be integrated into USEtox version 3.0, which currently relies on chronic EC50-based data.

The emergence of online experimental databases, particularly REACH and CompTox, offers access to ecotoxicity data for thousands of chemicals to calculate EFs as demonstrated by recent publications. However, in the absence of experimental data, QSAR models are increasingly employed to bridge data gaps [18]. These models correlate chemical structures with compound properties, using chemical descriptors for predictions [20]. Over time, their accuracy and reliability have grown, boosting their application. This trend is particularly significant as regulatory frameworks are progressively leaning towards minimizing animal testing [44]. Yet, millions of animals are still used annually for regulatory testing, a figure projected to rise given the ever-expanding use of chemicals in society [9, 23]. Prominent QSAR-based methods in use include ECOSAR, VEGA, and T.E.S.T. [6, 25, 26]

To effectively incorporate the available ecotoxicological data into CF calculations, the derivation of EFs is crucial. This study endeavors to leverage existing ecotoxicity data for such calculations. The objectives of this study are twofold: first, to develop a rigorous approach for ecotoxicity data harmonization, facilitating the determination of experimental EFs, and second, to compare ecotoxicity data from various QSAR models with experimental findings.

Materials and methods
Aquatic ecotoxicity data collection
Data on aquatic ecotoxicity were gathered from a variety of open-access sources to form an extensive data set. The primary experimental data were obtained from the REACH databases, which has substantial data on aquatic ecotoxicology [33]. This was further enriched by incorporating data from the CompTox version 2.2.1 database, which includes data from ToxValDB v9.4 [1]. A targeted approach was employed to secure experimental data specifically concerning aquatic environments from these

sources. The REACH data were retrieved in August 2020, and the CompTox data were collected in July 2023. Given the regular updates to the CompTox database by the US EPA, the most recent ToxValDB version 9.4 was utilized, rather than the initial data set of ToxValDB version 9.1.1 released in 2022. However, the REACH raw data, acquired from another research project, has not been updated due to time and resource constraints.

The REACH database is structured into sections that detail specific study results. This study concentrated on the "ecotoxicological information" section, more precisely the "aquatic toxicity" subsection. In this each study provides a specific tested concentration value, supplemented with comprehensive details such as the endpoint, testing methodology, environmental conditions, species examined, and the duration of the test. The data from the CompTox database was derived from the U.S. EPA's ToxValDB (version 9.4), a vast repository of experimental toxicity data aggregated from 49 distinct public sources, excluding any data from REACH. Within this collection, 10 sources are particularly focused on ecotoxicity data, including COSMOS, DOE ECORISK, DOE Wildlife Benchmarks, ECHA IUCLID, ECOTOX, EFSA, EnviroTox_v2, HAWC Project, HEAST, and HPVIS.

Estimated aquatic ecotoxicity data were gathered from two quantitative structure–activity relationships (QSARs): ECOSAR™ Version 1.11 through EPI Suite v4.11 [6, 50] and the US EPA Toxicity Estimation Software Tool (TEST) [26, 48], both accessed in August 2023. Estimated ecotoxicity data were first, sourced from ECOSAR™ Version 1.11 via EPI Suite v4.11 [6, 50]. ECOSAR encompasses organic and inorganic chemicals but omits organometallic chemicals and polymers. QSAR model applicability is constrained by the scope of its training set, which can lead to potential uncertainties when applied to substances beyond its intended domain. Consequently, any results generated by a QSAR tool for substances outside its domain were treated with caution due to these significant uncertainties. For this study, the harmonized experimental data set was processed through ECOSAR, estimating endpoints like fish LC50 96 h, daphnia LC50 48 h, and green algae EC50 96 h. More than 100 ECOSAR classes emerged from the analysis. For chemicals that aligned with several classes, both in terms of endpoints and species, a geometric mean was used to aggregate the data, resulting in a singular concentration value per endpoint and species. Second, ecotoxicity data estimates were obtained using the US EPA T.E.S.T. v5.1.2 [26, 48]. The estimated endpoints, derived from the consensus model in TEST, included

fathead minnow LC50 96 h, daphnia magna LC50 48 h, and tetrahymena pyriformis IC50 48 h.

Selection and harmonization of experimental and estimated ecotoxicity data

A data harmonization strategy was formulated in this study to streamline both experimental and estimated ecotoxicological effect data. The primary objective was to harmonize the gathered ecotoxicity data, to ensure uniformity and comparability across crucial aspects like test organisms, endpoints, exposure duration, and ecotoxicity concentration values. Chemicals with uncertain or missing details were omitted. This approach was built upon four steps, as shown in Fig. 1: chemical substance identification, data quality assessment, data harmonization, and consistency verification.

Data harmonization begins with the first step of chemical identification, which ensures that each data point is distinctly linked to its corresponding chemical. Following this initial step, the second step focuses on data quality assessment, emphasizing the importance of data reliability. This involves utilizing the database metadata to filter out lower quality datapoints from the overall data set. The third step of data harmonization ensures uniformity across several key aspects, including the naming and classification of test species, endpoint classification, units of study duration, exposure duration categories, and ecotoxicity concentration units. The process concludes with the final step of consistency checking, which scrutinizes the metadata for each data point to verify its completeness and the uniqueness of the collected data, alongside the removal of duplicates. Detailed descriptions of each of these steps are provided in the Supplementary Information in Table S2.

Ecotoxicity effect factors calculation strategy

This study focuses on the computation of ecotoxicity EFs, which are required to determine freshwater ecotoxicity CFs using USEtox tool. In USEtox, EFs are determined by calculating the linear slope along the concentration–response relationship of the HC50 [kg m^{-3}] up to the point where the PAF (potentially affected fraction of species) is 0.5 [12]. This relationship is quantified using the equation: EF = 0.5/HC50 [PAF m^3 kg^{-1}]. The HC50 represents the concentration at which half of the species in a freshwater ecosystem are exposed above their EC50 value, derived from the geometric mean of chronic EC50s for various freshwater species whereas, the EC50 represents the concentration at which 50% of the exposed species population exhibits an effect. All equations involved in calculating USEtox EFs are detailed in the official USEtox documentation [12]. In this study, the calculation of log HC50$_{EC50eq}$ begins with the extrapolation of
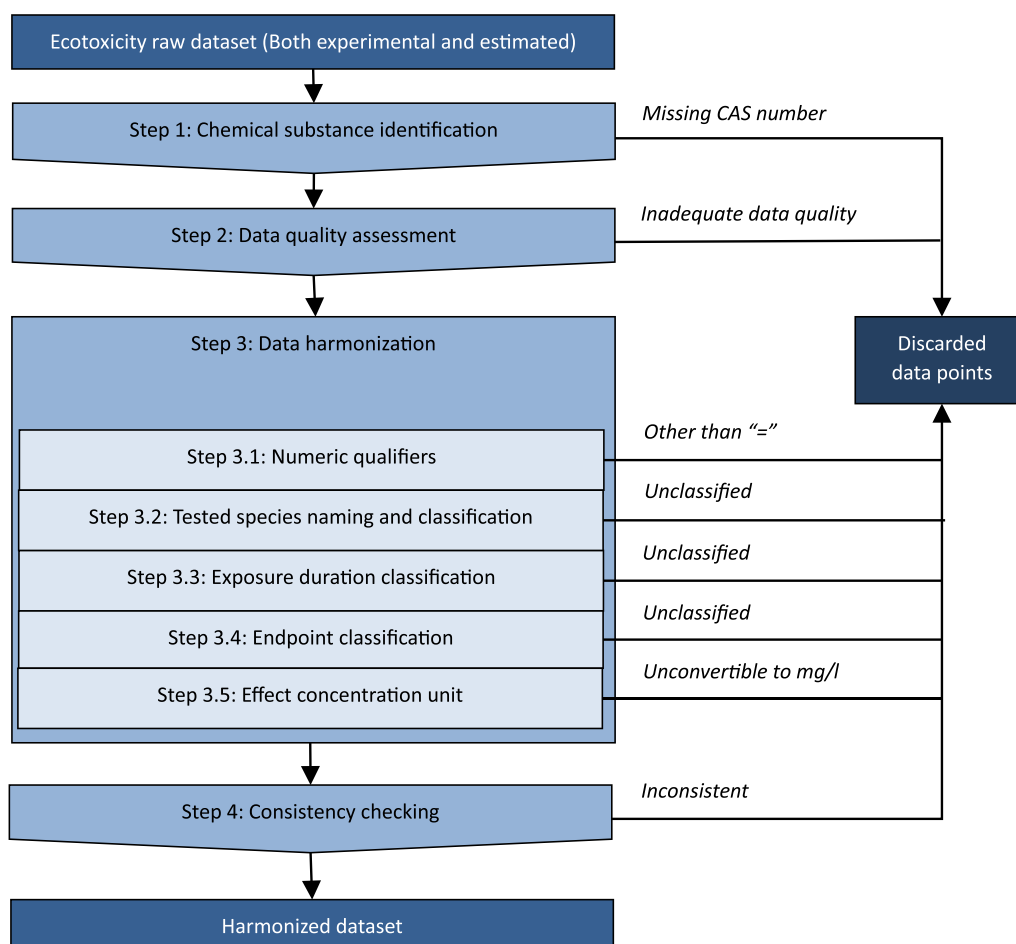
Fig. 1 Decision tree for experimental and estimated ecotoxicity data harmonization strategy

endpoints to chronic EC50 values (if needed), employing species group-specific extrapolation factors [2]. This step is followed by aggregating the extrapolated harmonized data at the species level utilizing the geometric mean for each chemical. Finally, calculating the average of log-values per chemical, resulting in the $logHC50_{EC50eq}$, which is utilized in the USEtox framework to calculate EFs based on $HC50_{EC50eq}$ for individual chemicals.

Building on the collaborative work of the Ecotoxicity Task Force and insights from the SETAC Pellston workshop, a new method for calculating USEtox EFs was proposed by Owsianiak et al. [30]. Traditionally, CFs relied on chronic EC50 values, but the recommendation advocated for an $HC20_{EC10eq}$ based approach. HC20 represents the environmental concentration affecting 20% of species, with the effect on individual species determined using chronic EC10 equivalent ecotoxicity data. This relationship is quantified using the equation: $EF = 0.2/HC20_{EC10eq}$ [PAF m^3 kg^{-1}]. Owsianiak et al. [30] emphasized the importance of utilizing experimentally determined chronic EC10 data for the calculation of HC20,

acknowledging the scarcity of such data. To bridge this gap, extrapolation factors were suggested [41]. Following these updated recommendations, this study also employs this methodology to calculate EFs based on $HC20_{EC10eq}$, thereby aligning with the latest guidelines for calculating EFs.

Comparing ecotoxicity effect factors

The calculated log10 transformed EFs are used for the comparative analysis. Whenever there was at least one chemical shared between the compared data sources, EFs were plotted for analysis. Pairwise correlations were employed to evaluate the linear relationship between these log10 transformed EFs, with the correlation coefficient, r, indicating both strength and direction of the relationship. Further evaluation of the correlation strength involved performing linear regression analyses with unrestricted slopes. The coefficient of determination (R^2) was then used to determine the robustness of the observed relationships.

The comparisons were conducted in several stages. Initially, this study compared the EFs derived from this study's experimental data with the existing USEtox EFs from the USEtox organic substances database (version 2.01). In this study, the EFs calculated were classified in three groups, starting with EFs based on $HC50_{EC50eq}$ using all the toxicity data after extrapolation to EC50 chronic equivalent denoted as EF ($HC50_{EC50eq}$), All data. Secondly, using only EC50 chronic datapoints denoted as EF ($HC50_{EC50eq}$), EC50chronic data only and thirdly using both EC50 chronic and EC50 acute extrapolated to chronic, denoted as EF ($HC50_{EC50eq}$), EC50 data only. Then, this study proceeded to compare the EFs derived from this study estimated data with the USEtox EFs. Following this, this study compared experimental EFs with those calculated using ECOSAR and TEST. Additionally, this study compared the EFs derived from this study experimental data using two different methodologies for calculating EFs. One method is based on the USEtox 2.13 documentation, and the other follows the recommendations by Owsianiak et al. [30], as described in Sect. "Ecotoxicity effect factors calculation strategy". To evaluate the consistency of the relationship, this study utilized the R^2 coefficient obtained from log-transformed regression analysis.

Results and discussion
Aquatic ecotoxicity data collected
The ecotoxicity data used in this study were gathered from two primary sources: experimental data obtained from REACH and CompTox, and estimated data extracted from ECOSAR and TEST. This collected data set underwent a harmonization process to create a harmonized ecotoxicity data set for EF calculations. From the REACH database, this study retrieved 225,517 ecotoxicity datapoints spanning 12,411 chemicals, each associated with an European Commission (EC) Number. In parallel, the CompTox database contributed 517,067 ecotoxicity datapoints encompassing 8640 chemicals, each identified by its CAS number. On the estimated data side, ECOSAR data provided 27,354 toxicity datapoints for 6029 chemicals, while the US EPA T.E.S.T. database accounted for 17,055 datapoints spanning 6762 chemicals. It is important to note that chemical identification in these sources relied on CAS numbers.

Ecotoxicity data selection and harmonization results
The data harmonization strategy, as described in Sect. "Selection and harmonization of experimental and estimated ecotoxicity data", is applied to both experimental and estimated ecotoxicity data. The results of this process at various steps of the ecotoxicity data harmonization strategy, across different data sources, are presented in Table 1 and further detailed in the supplementary information in Table S3. Table 2 provides an overview of the distribution of harmonized ecotoxicity data points and the corresponding number of chemicals across various data sources and endpoint classifications, illustrating the number of datapoints and chemicals at different endpoints. Following harmonization, the REACH data set was refined to include 72,705 data points for 5318 chemicals, while the CompTox data set retained 364,434 datapoints for 7464 chemicals. The aggregation of these two data sets and a subsequent duplicate verification step resulted in the elimination of an additional 2485 data points from the consolidated data set. As a result, this study compiled a harmonized data set of experimental datapoints comprising 434,654 datapoints for 11,295 chemicals, as shown in Tables 1 and 2. Each datapoint in the final combined data set is distinctly

Table 1 Overview of ecotoxicity data harmonization results at different steps of ecotoxicity data harmonization strategy across various data sources

Steps	Step description	REACH	CompTox	ECOSAR	TEST
	Ecotoxicity raw data set	225,517 (12411)	517,067 (8640)	27,354 (6029)	17,055 (6762)
Step 1	Chemical substance identification	177,296 (9035)	515,803 (8639)	27,354 (6029)	17,055 (6762)
Step 2	Data quality assessment	153,751 (8854)	504,355 (8549)	19,719 (6029)	17,055 (6762)
Step 3	Data harmonization	153,751 (8854)	504,355 (8549)	19,719 (6029)	17,055 (6762)
Step 3.1	Numeric qualifiers	110,502 (7333)	477,063 (8217)	19,719 (6029)	17,055 (6762)
Step 3.2	Tested species naming and classification	101,541 (6780)	467,909 (7936)	19,719 (6029)	17,055 (6762)
Step 3.3	Exposure duration classification	91,630 (6728)	434,742 (7704)	19,719 (6029)	17,055 (6762)
Step 3.4	Endpoint classification	90,849 (6719)	412,939 (7464)	19,719 (6029)	17,055 (6762)
Step 3.5	Effect concentration unit	90,849 (6719)	412,939 (7464)	19,719 (6029)	17,055 (6762)
Step 4	Consistency checking	72,705 (5318)	364,434 (7464)	19,719 (6029)	17,055 (6762)
	Harmonized data set	434,654 (11295)		19,719 (6029)	17,055 (6762)

Number of datapoints (number of chemicals)

Table 2 Overview of the distribution of harmonized ecotoxicity datapoints and corresponding number of chemicals across various data sources and endpoint classifications

Data Source	EC10 acute	EC10 chronic	EC50 acute	EC50 chronic	NOEC acute	NOEC chronic	Total
REACH	7373 (751)	10,745 (1986)	22,631 (3661)	9495 (2745)	10,307 (2834)	12,154 (2764)	72,705 (5318)
CompTox	33,830 (2046)	37,099 (2203)	172,505 (5937)	28,618 (2826)	43,101 (3095)	49,281 (2499)	364,434 (7464)
Combined	41,170 (2702)	47,677 (3904)	193,611 (8513)	37,839 (5047)	53,206 (5430)	61,151 (4730)	434,654 (11,295)
ECOSAR	–	–	19,719 (6029)	–	–	–	19,719 (6029)
TEST	–	–	17,055 (6762)	–	–	–	17,055 (6762)

Number of datapoints (number of chemicals)

identified by its CAS, species name, species group, exposure classification, and endpoint classification. The final data sets for ECOSAR and TEST stood at 19,719 datapoints for 6029 chemicals and 17,055 datapoints for 6762 chemicals, respectively, as shown in Tables 1 and 2.

Ecotoxicity effect factors results

The harmonized ecotoxicity data in this study underwent an aggregation process, as outlined in Sect. "Ecotoxicity effect factors calculation strategy" to ensure data integrity and relevance for EF calculations in accordance with USEtox guidelines. Initially, for the calculation of log $HC50_{EC50eq}$ begins with the extrapolation of endpoints to chronic EC50 values employing species group-specific extrapolation factors [2]. This step is followed by aggregating the extrapolated harmonized data at the species level, producing 59,195 ($n=11,295$) aggregated effect concentration data points for the experimental data set, with minor changes in the ECOSAR data set resulted in 16,456 ($n=6029$) aggregated datapoints but no alterations in the TEST data set. Lastly, chronic EC50 data underwent further aggregation, utilizing the average of logarithmic values to calculate log $HC50_{EC50eq}$ [kg m^{-3}]. This data set was then utilized in the USEtox model to compute EFs based on $HC50_{EC50eq}$ for each chemical.

In this study, EFs were derived by combining experimental EC50 chronic values with extrapolated endpoints to EC50 chronic values, resulting in EFs for 11,295 chemicals. While USEtox advocates for the exclusive use of EC50 chronic data for reduced uncertainty, this study also computed EFs solely from EC50 chronic data for 5047 chemicals. Furthermore, the analysis extended to encompass EFs derived from EC50 endpoints, both chronic and extrapolated acute, for 9543 chemicals, broadening the chemical coverage but introducing the uncertainty inherent in relying on a single endpoint. As for the estimated data, EFs were computed for 6029 chemicals using ECOSAR and 6762 chemicals using TEST. In contrast, USEtox provides EFs for 2499 chemicals, with 2426 chemicals overlapping with the experimental data in this study.

Comprehensive statistics, including data points, species groups, log HC50 (log mg/L), and EF [PAF.m^3.kg^{-1}], are presented in Table 3.

Following recommendations proposed by Owsianiak et al. [30], this study also calculated EFs based on $HC20_{EC10eq}$, as outlined in Sect. "Ecotoxicity effect factors calculation strategy" thereby aligning with the latest guidelines for calculating EFs. The calculation of log $HC20_{EC10eq}$ begins with the extrapolation of endpoints to chronic EC10 values employing species group-specific extrapolation factors [2]. This step is followed by aggregating the extrapolated harmonized data at the species level, producing 59,195 ($n=11,295$) aggregated effect concentration data points for the experimental data set. Finally, chronic EC10 data underwent further aggregation, utilizing the procedure given in Owsianiak et al. [30] to calculate log $HC20_{EC10eq}$ [kg m^{-3}]. The equation used is $\log HC20E_{C10eq} = SSD\mu\log_{EC10eq} + z_{0.2}.SSD\sigma\log_{EC10eq}$. This data set was then utilized to compute EFs based on $HC20_{EC10eq}$ for each chemical using the equation: $EF = 0.2/HC20_{EC10eq}$ [PAF m^3 kg^{-1}].

For substances with only one species toxicity data, $HC20_{EC10eq}$ was calculated following the procedure outlined in Saouter et al. [41]. This involves assuming that if there is only one EC10eq, then this value is equivalent to an HC50. Subsequently, an extrapolation equation is applied to convert the HC50 into an estimated HC20. In this study, the extrapolation equation was derived by comparing log-transformed values of $HC50_{EC10eq}$ [mg/l] and $HC20_{EC10eq}$ [mg/l] across chemicals with data from at least three species, encompassing 4563 chemicals. The regression analysis, as shown in Fig. 2, revealed a robust correlation, with an R^2 value of 0.93, indicating a strong consistency between the compared values. The resulting extrapolation equation is log $HC20_{EC10eq} = -0.5 + 1.04*\log HC50_{EC10eq}$, with the default extrapolation factor set as $HC20_{EC10eq} = 0.31* HC50_{EC10eq}$.

It is important to note that not all scientists endorse the USEtox approach, and ongoing improvements are being made to the methodology. These include recommendations by Owsianiak et al. [30] to introduce additional

Table 3 Overview of the summary statistics of different ecotoxicity data sources

Data Source	Statistics	Chemicals	Mean	Std Dev	Sum	Minimum	Maximum
All experimental data, EC50eq	Datapoints	11,295	38.48	235.50	434,654	1	13,727
	Species	11,295	5.24	15.47	–	1	579
	Species groups	11,295	2.31	1.73	–	1	12
	log HC50$_{EC50eq}$ (log mg/L)	11,295	0.56	1.20	–	−7.00	5.90
	EF [PAF.m^3.kg^{-1}]	11,295	5.3E+05	4.7E+07	–	6.0E-04	5.0E+09
Only EC50 chronic experimental data	Datapoints	5047	7.50	29.13	37,839	1	937
	Species	5047	2.14	3.93	–	1	113
	Species groups	5047	1.47	1.04	–	1	11
	log HC50 $_{EC50eq}$ (log mg/L)	5047	0.62	1.41	–	−6.60	9.18
	EF [PAF.m^3.kg^{-1}]	5047	7.6E+05	3.0E+07	–	3.3E-07	2.0E+09
Only EC50 experimental data	Datapoints	9543	24.25	142.33	231,450	1	7285
	Species	9543	4.85	13.39	–	1	449
	Species groups	9543	2.23	1.64	–	1	12
	log HC50 $_{EC50eq}$ (log mg/L)	9543	0.53	1.25	–	−6.45	6.23
	EF [PAF.m^3.kg^{-1}]	9543	2.6E+05	1.5E+07	–	3.0E-04	1.4E+09
All experimental data, EC10eq	Datapoints	11,295	38.48	235.50	434,654	1	13,727
	Species	11,295	5.24	15.47	–	1	579
	Species groups	11,295	2.31	1.73	–	1	12
	log HC20$_{EC10eq}$ (log mg/L)	11,295	−0.50	1.22	–	−8.37	5.37
	EF [PAF.m^3.kg^{-1}]	11,295	4.4E+06	4.4E+08	–	9.0E-04	4.7E+10
ECOSAR data	Datapoints	6029	2.73	0.60	16,456	1	3
	Species groups	6029	2.73	0.60	–	1	3
	log HC50 $_{EC50eq}$ (log mg/L)	6029	0.70	1.69	–	−6.39	5.69
	EF [PAF.m^3.kg^{-1}]	6029	1.1E+06	2.8E+07	–	1.0E-03	1.2E+09
TEST data	Datapoints	6762	2.52	0.68	17,055	1	3
	Species groups	6762	2.52	0.68	–	1	3
	log HC50 $_{EC50eq}$ (log mg/L)	6762	0.63	1.23	–	−5.42	4.45
	EF [PAF.m^3.kg^{-1}]	6762	5.5E+04	2.2E+06	–	1.8E-02	1.3E+08
USEtox data[a]	log HC50 $_{EC50eq}$ (log mg/L)	2426	0.63	1.42	–	−5.21	4.44
	EF [PAF.m^3.kg^{-1}]	2426	1.1E+05	2.4E+06	–	1.8E-02	8.2E+07

[a] In USEtox version 2.13, under the Effect Factor (EF) Input Data section, the number of data points and species groups are not available for all 2426 chemicals; therefore, they are not included in the table
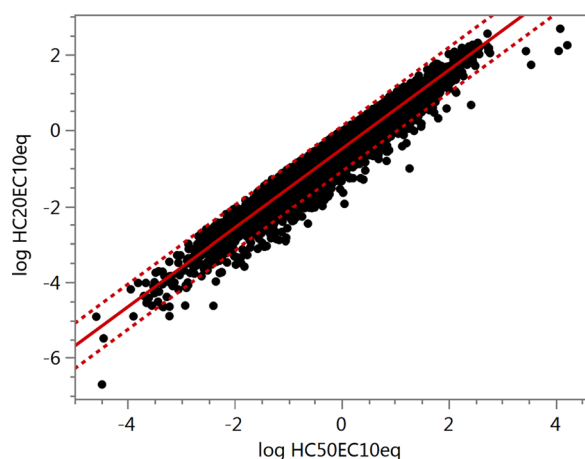
Fig. 2 Regression analysis of log HC20$_{EC10eq}$ [mg/l] versus log HC50$_{EC50eq}$ [mg/l] for chemicals with data from at least 3 species, exhibiting a high correlation (n = 4563, R^2 = 0.93, r = 0.96)

environmental compartments and adjust ecotoxicity effect modeling to reflect more environmentally relevant conditions by shifting focus from EC50 to EC10. USEtox also faces reliability issues with specific chemical groups such as PFAS, prompting Holmquist et al. [17] to develop a PFAS-adapted version of USEtox (version 2.1), which incorporates several enhancements for assessing PFAS ecotoxicity impacts. Adjustments were also made to ecotoxicological EFs considering species richness, following the LC-Impact method [15, 53]. For nanomaterials, USEtox is not suitable because its parameters do not support multimedia fate modeling of nanomaterials [7, 31]. Additionally, the method for including transformation products in CF calculations is not well-defined; Van Zelm et al. [52] suggested adjusting the parent compound CFs in proportion to the CFs of its transformation products. Related to risk assessment, Saouter et al. [39] noted the

importance of aligning the ranking of chemicals by their ecotoxicity impacts in LCA with hazard characterizations used in risk assessment. Saouter et al. [39] compared the ecotoxicity approaches in LCA and chemical risk assessment and observed that these approaches could differ significantly, with discrepancies up to four orders of magnitude for the chemicals compared. LCA toxicity indicators cover various aspects of a chemical's ecotoxicity potential, with USEtox aiming to assess the potential overall toxicity impact of substances on humans and ecosystems rather than focusing solely on hazard identification, which is central to risk assessment [40]. USEtox is a scientific consensus model used to screen, not necessarily to identify the chemical of concern according to risk-assessment approach [40]. This discrepancy underscores the critical need to establish a consensus between risk assessment methodologies and USEtox-based LCA calculations to optimize resource use and enhance the sustainable management of chemicals.

The calculated freshwater EFs exhibit a significant range, spanning between 11 to 16 orders of magnitude, as depicted in Table 3 and elaborated in supplementary information in Table S1. To visually illustrate this variation, Fig. 3 presents box plots categorizing EFs based on the type of data considered and EF calculation method used. These graphical representations effectively showcase the distribution of EFs, highlighting the considerable diversity in values across the different data categories.

Comparison of ecotoxicity effect factors

The initial comparison involved contrasting the EFs based on $HC50_{EC50eq}$ derived from experimental data in this study with the existing USEtox EFs available in the USEtox organic substances database (version 2.01), as illustrated in Fig. 4 and Table 4. This assessment covered EFs for 2426 chemicals already present in the USEtox database, recalculated using experimental data. The results of the regression analysis were indicative of a robust correlation, with an R^2 value of 0.84, demonstrating a strong alignment between recalculated values and those in USEtox.

In line with the USEtox guideline prioritizing EC50 chronic data, this study conducted a separate analysis comparing EFs calculated solely with this data against USEtox EFs. This yielded an R^2 value of 0.64, indicating a moderate correlation. However, when EFs derived from both chronic EC50 and extrapolated acute EC50 data were compared to USEtox EFs, the R^2 value substantially
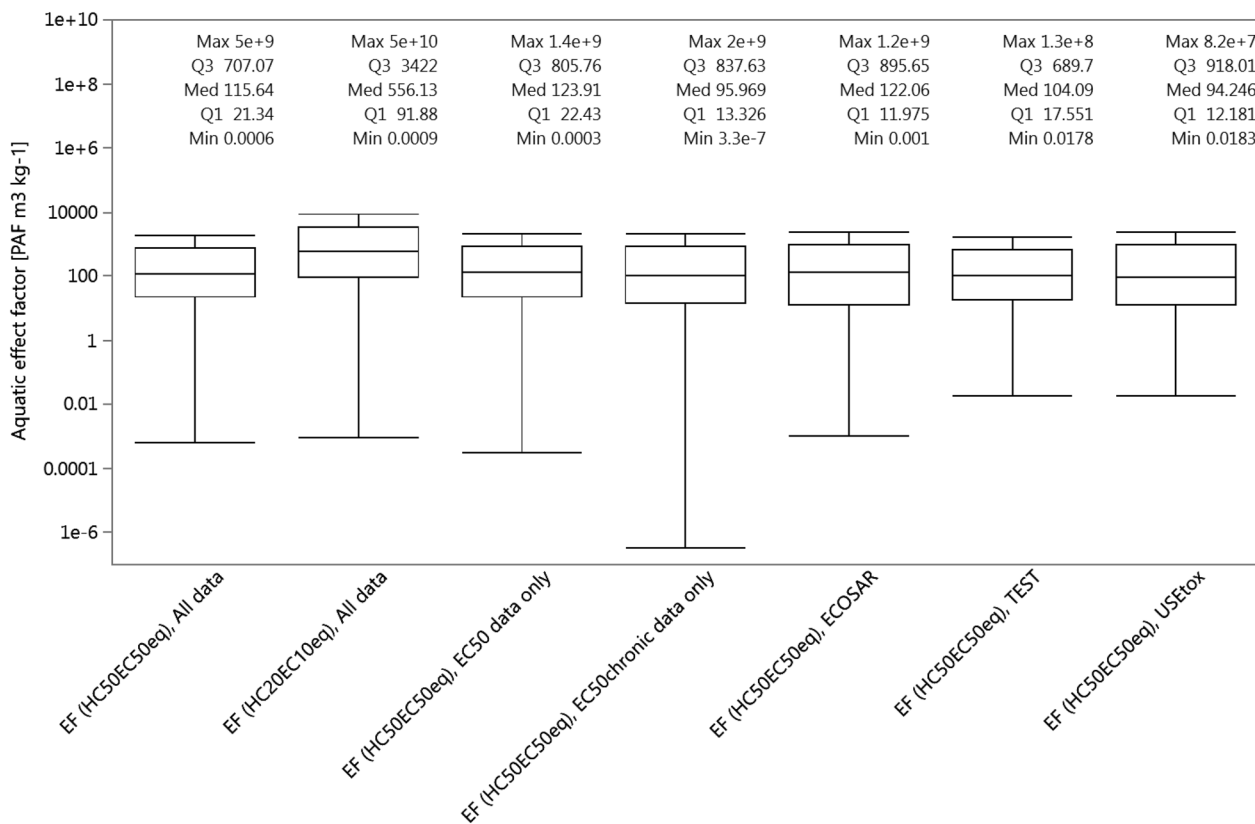
Fig. 3 Box plots of calculated EFs from experimental and estimated data in the freshwater compartment for different types of data

increased to 0.85, signaling a strong correlation. This improvement suggests that incorporating extrapolated data significantly enhances data coverage and species group representation, resulting in normalized values closely aligned with USEtox. The notably high correlation observed between recalculated and USEtox EFs provides confidence that the EFs computed in this study for chemicals absent from the USEtox database adhere to USEtox calculation criteria. This underscores the reliability of the EFs generated for chemicals not originally included in USEtox.

The second comparison involved contrasting the EFs based on $HC50_{EC50eq}$ derived from estimated data in this study with the pre-existing USEtox EFs available in the USEtox organic substances database (version 2.01), as illustrated in Fig. 5. In this assessment, EFs for 2165 chemicals from ECOSAR and 2098 chemicals from TEST, which were already listed in USEtox, were recalculated using their respective estimated data.
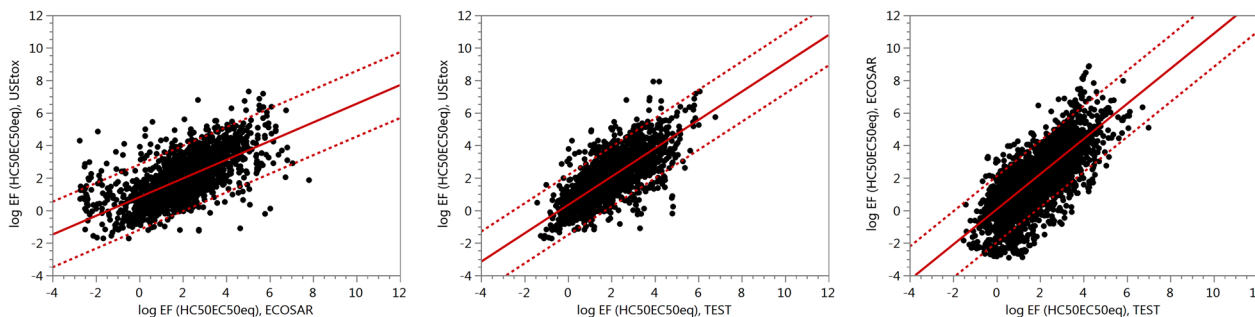
The regression analysis between USEtox and ECOSAR EFs revealed an R^2 value of 0.42, indicating a relatively weak correlation. A similar analysis between USEtox and TEST EFs yielded an R^2 value of 0.55, suggesting a slightly stronger but still weak correlation when compared to ECOSAR. However, when EFs from ECOSAR were compared against TEST EFs, the R^2 value increased to 0.57,
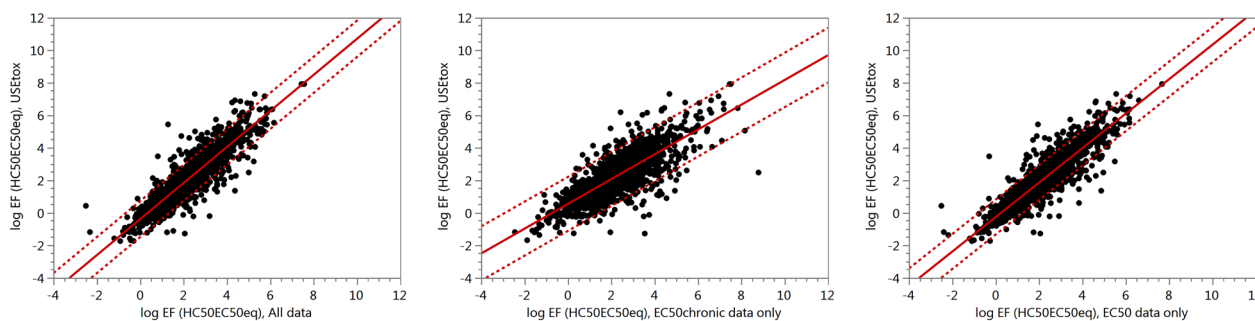
Fig. 4 Regression analysis of pre-calculated log transformed USEtox 2.13 database EFs [PAF m^3 kg^{-1}] in freshwater ecosystem versus log transformed EFs [PAF m^3 kg^{-1}] calculated in this study based on $HC50_{EC50eq}$ with experimental ecotoxicity data: all (left), EC50 chronic (middle), and EC50 (right), with correlation: moderate ($n = 2426$, $R^2 = 0.84$, $r = 0.92$), low ($n = 1418$, $R^2 = 0.64$, $r = 0.80$), and high ($n = 2406$, $R^2 = 0.85$, $r = 0.92$) respectively

Table 4 Overview of the regression analysis of log transformed EFs based on $HC50_{EC50eq}$ calculated in this study versus pre-calculated log transformed USEtox 2.13 database [PAF m^3 kg^{-1}]

Variable	By variable	Correlation	Rsquare	Root mean square error	Covariance	Count	Lower 95%	Upper 95%
log EF, USEtox	log EF, All	0.92	0.84	0.56	1.52	2426	0.91	0.92
log EF, USEtox	log EF, EC50 chronic	0.80	0.64	0.84	1.67	1418	0.78	0.82
log EF, USEtox	log EF, EC50	0.92	0.85	0.55	1.61	2406	0.92	0.93

Fig. 5 Regression analysis between pre-calculated log transformed USEtox 2.13 database EFs, ECOSAR EFs, and TEST EFs [PAF m^3 kg^{-1}] in freshwater ecosystem: USEtox vs ECOSAR (left), USEtox vs TEST (middle), and ECOSAR vs TEST (right), with correlation: weak ($n = 2165$, $R^2 = 0.42$, $r = 0.65$), weak ($n = 2098$, $R^2 = 0.55$, $r = 0.74$), and weak ($n = 4930$, $R^2 = 0.57$, $r = 0.76$) respectively

signifying a relatively stronger, yet still weak, correlation. This observed increase in the R^2 value may be attributed to the inclusion of more data points, which can lead to values that align more closely with each other.

The low correlation between recalculated EFs using estimated data and USEtox EFs implies that the EFs for chemicals with estimated data do not align well with USEtox effect data. This suggests a lower degree of confidence in the EFs calculated for chemicals based on estimated data. Table 5 provides summary statistics of the regression analysis, presenting log-transformed pre-calculated USEtox 2.13 database EFs, ECOSAR EFs, and TEST EFs. Furthermore, the EFs calculated in this study with ECOSAR and TEST exhibit relatively low correlation, indicating that estimations from different QSARs can lead to differing results. It is important to note that QSAR models have a limited applicability domain, determined by their training data, making them more suitable for certain substance groups but not universally applicable. Typically, QSARs provide acute data for specific endpoints, covering a limited range of species groups. This data often requires extrapolation to the necessary endpoints and aggregation to calculate the EF for each chemical, introducing an additional layer of uncertainties. Therefore, when utilizing QSAR-based tools for calculating toxicity data to determine EFs, it is crucial to carefully understand the associated uncertainties and exercise caution. Users should take into account the applicability domain of QSARs for specific chemicals and avoid prioritizing QSAR derived EFs over experimental EFs, except in cases where experimental data are unavailable. This is supported by the findings presented in Table 5, which show a generally low correlation between experimental and estimated data based EFs.

The third comparison focused on contrasting the EFs based on $HC50_{EC50eq}$ derived from the experimental data in this study with the calculated EFs from ECOSAR, as illustrated in Fig. 6 and Table 6. In this investigation, 6029

Table 5 Overview of the regression analysis between log transformed pre-calculated USEtox 2.13 database EFs, ECOSAR EFs, and TEST EFs [PAF m^3 kg^{-1}]

Variable	By variable	Correlation	Rsquare	Root mean square error	Covariance	Count	Lower 95%	Upper 95%
log EF, USEtox	log EF, ECOSAR	0.65	0.42	1.02	1.35	2165	0.63	0.68
log EF, USEtox	log EF, TEST	0.74	0.55	0.94	1.25	2098	0.72	0.76
log EF, ECOSAR	log EF, TEST	0.76	0.57	1.04	1.35	4930	0.75	0.77
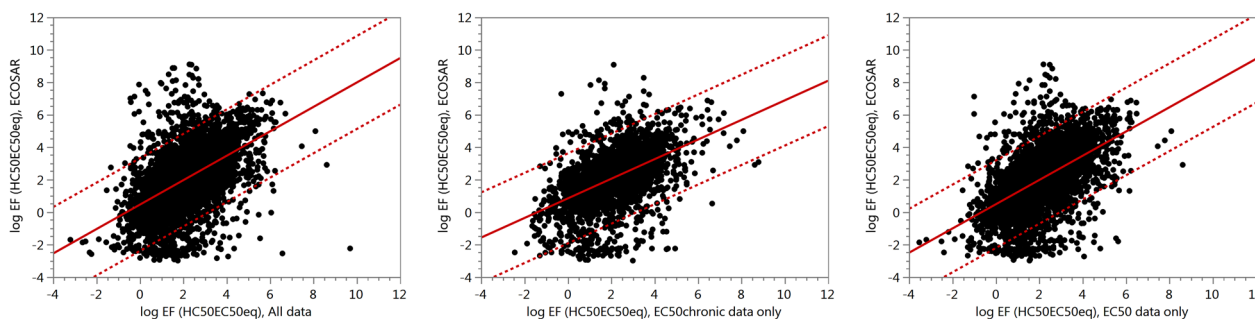
Fig. 6 Regression analysis of ECOSAR EFs [PAF m^3 kg^{-1}] in freshwater ecosystem versus log transformed EFs [PAF m^3 kg^{-1}] calculated in this study with experimental ecotoxicity data: all (left), EC50 chronic (middle), and EC50 (right), with correlation: very weak ($n=6030$, $R^2=0.26$, $r=0.51$), very weak ($n=2964$, $R^2=0.26$, $r=0.51$), and very weak ($n=5409$, $R^2=0.30$, $r=0.55$) respectively

Table 6 Overview of the regression analysis of log transformed EFs versus calculated log transformed ECOSAR EFs [PAF m^3 kg^{-1}]

Variable	By variable	Correlation	Rsquare	Root mean square error	Covariance	Count	Lower 95%	Upper 95%
log EF, ECOSAR	log EF, All	0.51	0.26	1.45	1.00	6029	0.49	0.53
log EF, ECOSAR	log EF, EC50 chronic	0.51	0.26	1.41	1.13	2964	0.48	0.53
log EF, ECOSAR	log EF, EC50	0.55	0.30	1.37	1.07	5408	0.53	0.56

chemicals were found in common between ECOSAR and the experimental data from this study.

The regression analysis between ECOSAR EFs and all the experimental data showed an R^2 value of 0.26, indicating a very weak correlation. Adhering to the USEtox guideline of prioritizing EC50 chronic data, a separate analysis compared EFs calculated using only these data with ECOSAR EFs, resulting in an R^2 value of 0.26, pointing to a very weak correlation. However, when EFs derived from both chronic EC50 and extrapolated acute EC50 were compared against ECOSAR EFs, the R^2 value increased to 0.30, signifying a very weak correlation, but relatively stronger than the previous comparisons.

The low correlation between the experimental EFs and the ECOSAR EFs implies that the estimated values for chemicals do not align well with experimental values in general. This results in a low degree of confidence in the EFs calculated with estimated data. Table 6 provides summary statistics of the regression analysis, presenting log-transformed EFs calculated in this study versus log-transformed ECOSAR EFs.

The fourth comparison centered on contrasting the EFs based on $HC50_{EC50eq}$ derived from the experimental data in this study with the calculated EFs from TEST, as illustrated in Fig. 7 and Table 7. In this research, 6762 chemicals were common between TEST and the experimental data.

The regression analysis between TEST EFs and all the experimental data showed an R^2 value of 0.34, indicating a very weak correlation. Following the USEtox guideline, which prioritizes EC50 chronic data, a separate analysis compared EFs calculated using only this data with TEST EFs. This yielded an R^2 value of 0.31, pointing to a very weak correlation. However, when EFs derived from both chronic EC50 and extrapolated acute EC50 were compared against TEST EFs, the R^2 value increased to 0.38, signifying a very weak correlation, but relatively stronger than the previous comparisons.

The low correlation between the experimental EFs and the TEST EFs implies that the estimated values for chemicals do not align well with experimental values. Consequently, there is a low degree of confidence in the EFs calculated with estimated data. Table 7 provides summary statistics of the regression analysis, presenting log-transformed EFs calculated in this study versus log-transformed TEST EFs.

The fifth comparison focused on evaluating the EFs based on $HC50_{EC50eq}$ derived from the experimental data in this study, considering different endpoint inclusions, as illustrated in Fig. 8. The regression analysis conducted between EC50 chronic data, and all experimental data yielded an R^2 value of 0.82, indicating a strong correlation. Likewise, when comparing combined chronic EC50 and extrapolated acute EC50 against all experimental data, a high R^2 value of 0.94 was observed, signifying a robust correlation.

However, when EFs derived from both chronic EC50 and extrapolated acute EC50 were compared against
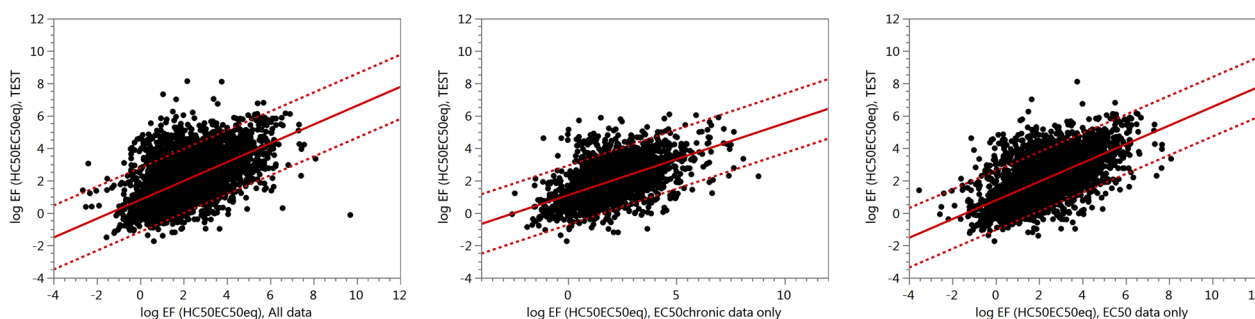
Fig. 7 Regression analysis of TEST EFs [PAF m³ kg⁻¹] in freshwater ecosystem versus log transformed EFs [PAF m³ kg⁻¹] calculated in this study with experimental ecotoxicity data: all (left), EC50 chronic (middle), and EC50 (right), with correlation: very weak ($n = 6762$, $R^2 = 0.34$, $r = 0.58$), very weak ($n = 2994$, $R^2 = 0.31$, $r = 0.56$), and very weak ($n = 5794$, $R^2 = 0.38$, $r = 0.62$) respectively

Table 7 Overview of the regression analysis of experimental log transformed EFs versus calculated log transformed TEST database EFs [PAF m³ kg⁻¹]

Variable	By variable	Correlation	Rsquare	Root mean square error	Covariance	Count	Lower 95%	Upper 95%
log EF, TEST	log EF, All	0.58	0.34	1.01	0.88	6762	0.56	0.60
log EF, TEST	log EF, EC50 chronic	0.56	0.31	0.93	0.88	2994	0.53	0.58
log EF, TEST	log EF, EC50	0.62	0.38	0.93	0.94	5794	0.60	0.64

EC50 chronic data EFs, the R^2 value remained strong at 0.86, suggesting that while the inclusion of different endpoints has an effect, it is not a significant determinant of the correlation strength. This implies a high degree of confidence in the EFs calculated when considering both chronic EC50 and extrapolated acute EC50 data. Table 8 provides summary statistics of the regression analysis, presenting log transformed EFs calculated in this study for all experimental data, all EC50 chronic data, and EC50 data.

The final comparison focused on contrasting the EFs derived from the experimental data in this study based on $HC50_{EC50eq}$ with the calculated EFs based on $HC20_{EC10eq}$, as illustrated in Fig. 9 and Table 9.

The regression analysis with all the experimental data showed an R^2 value of 0.93, indicating a very strong correlation. Adhering to the USEtox guideline of prioritizing EC50 chronic data, a separate analysis compared EFs calculated using only this data with EFs based on $HC20_{EC10eq}$, resulting in an R^2 value of 0.75, pointing to a moderate correlation. However, when EFs derived from both chronic EC50 and extrapolated acute EC50 were compared against EFs based on $HC20_{EC10eq}$, the R^2 value increased to 0.88, signifying a strong correlation.

The strong correlation between the experimental EFs based on $HC50_{EC50eq}$ and the EFs based on $HC20_{EC10eq}$ implies that data can be used to calculate extrapolation factor to convert EF ($HC50_{EC50eq}$)
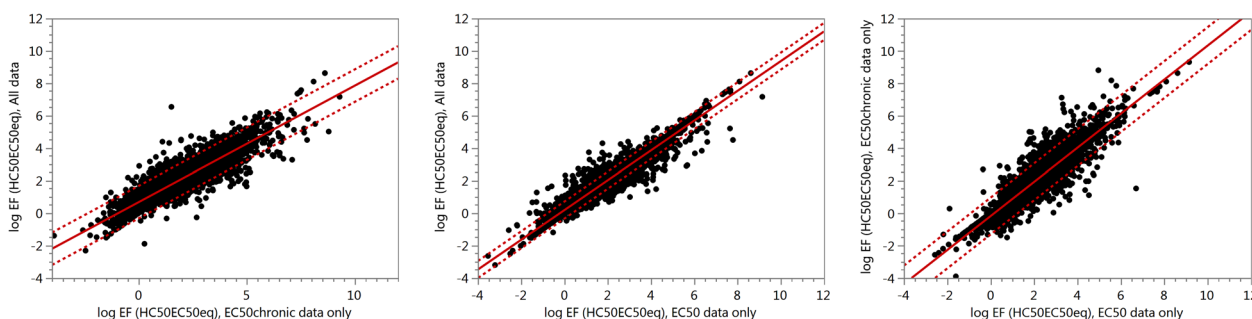
Fig. 8 Regression analysis of EFs [PAF m³ kg⁻¹] in freshwater ecosystem between log transformed EFs [PAF m³ kg⁻¹] calculated in this study with experimental ecotoxicity data: all vs EC50 chronic (left), all vs EC50 (middle), and EC50 chronic vs EC50 (right), with correlation: moderate ($n = 5047$, $R^2 = 0.82$, $r = 0.91$), high ($n = 9543$, $R^2 = 0.94$, $r = 0.97$), and high ($n = 5047$, $R^2 = 0.86$, $r = 0.93$) respectively

Table 8 Overview of the regression analysis between log transformed EFs calculated with all experimental data, all EC50 chronic data, and EC50 data [PAF m3 kg-1]

| Variable	By variable	Correlation	Rsquare	Root mean square error	Covariance	Count	Lower 95%	Upper 95%
log EF, All	log EF, EC50 chronic	0.91	0.82	0.47	1.41	5047	0.90	0.91
log EF, All	log EF, EC50	0.97	0.94	0.28	1.38	9543	0.97	0.97
log EF, EC50 chronic	log EF, EC50	0.93	0.86	0.52	1.58	5047	0.93	0.93
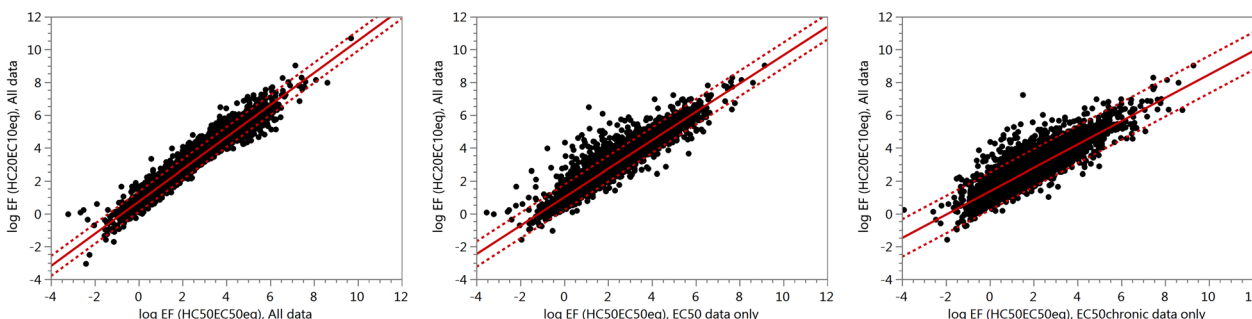
Fig. 9 Regression analysis of log transformed EFs [PAF m³ kg⁻¹] based on $HC20_{EC10eq}$ in freshwater ecosystem versus log transformed EFs [PAF m³ kg⁻¹] based on $HC20_{EC50eq}$ calculated in this study with experimental ecotoxicity data: all (left), EC50 (middle), and EC50 chronic (right), with correlation: high ($n = 11,295$, $R^2 = 0.93$, $r = 0.97$), moderate ($n = 5047$, $R^2 = 0.75$, $r = 0.86$), and high ($n = 9543$, $R^2 = 0.88$, $r = 0.94$) respectively

Table 9 Overview of the regression analysis between log transformed EFs calculated using two different methodologies [PAF m^3 kg^{-1}]

EF based on HC20$_{EC10eq}$	EF based on HC50$_{EC50eq}$	Correlation	Rsquare	Root Mean Square Error	Covariance	Count	Lower 95%	Upper 95%
log EF (HC20$_{EC10eq}$), All	log EF, All	0.97	0.93	0.31	1.41	11,295	0.97	0.97
log EF (HC20$_{EC10eq}$), All	log EF, EC50 chronic	0.86	0.75	0.58	1.41	5047	0.86	0.87
log EF (HC20$_{EC10eq}$), All	log EF, EC50	0.94	0.88	0.39	1.34	9543	0.94	0.94
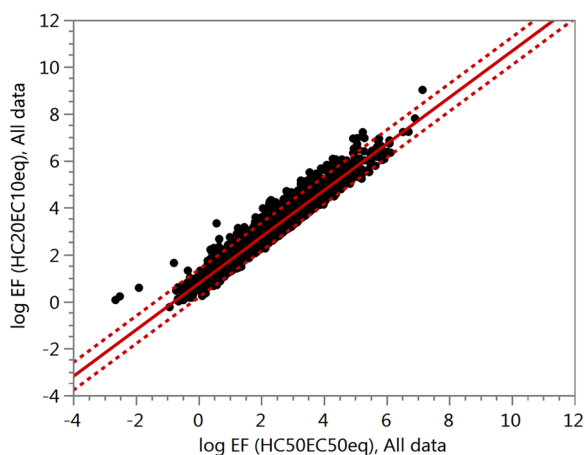
Fig. 10 Regression analysis of log EF (HC20$_{EC10eq}$) versus log EF (HC50$_{EC50eq}$) for chemicals with data from at least 3 species groups, exhibiting a high correlation ($n = 3827$, R$^2 = 0.93$, $r = 0.96$)

to EF (HC20$_{EC10eq}$) for chemicals with available EF (HC50$_{EC50eq}$). In this study, the extrapolation equation was derived by comparing log-transformed values of EF (HC50$_{EC50eq}$) and EF (HC20$_{EC10eq}$) across chemicals with data from at least three species groups, encompassing 3827 chemicals. The regression analysis, as shown in Fig. 10, revealed a robust correlation, with an R^2 value of 0.93, indicating a strong correlation between the compared values. The resulting extrapolation equation is log EF (HC20$_{EC10eq}$) = 0.75 + 0.9897*log EF (HC50$_{EC50eq}$), with the default extrapolation factor as EF (HC20$_{EC10eq}$) = 5.33*EF (HC50$_{EC50eq}$).

In this research, ecotoxicity data from REACH and CompTox version 2.2.1 underwent a systematic harmonization process as discussed in the method section, aligning with USEtox guidelines to ensure data set integrity and relevance for EF calculations. This thorough harmonization procedure led to a 37% reduction in data points. However, the study encountered limitations due to expertise constraints and information gaps inherent in ecotoxicity data with gaps and uncertainties. Challenges arose from the absence of a globally recognized standard for species naming and categorization in ecotoxicity, prompting the use of the US EPA ECOTOX knowledgebase [49]. Exposure classifications, as provided

by Aurisano et al. [2], were employed as the best available estimates for algae tests due to a lack of clear acute/chronic differentiation in typical 72-h exposure. Endpoint classification complexities, especially within lower species sensitivity distributions, made distinguishing between NOEC, LOEC, and EC 1–10 values difficult [21]. Inconsistent units across endpoints resulted in the exclusion of endpoints that were unconvertible. Additionally, the variety of effects studied for a specific endpoint, species, and exposure type introduced further complexity, given varying sensitivities and data limitations. The study aimed to offer a close approximation rather than an accurate representation of the situation, acknowledging the influence of numerous identified and unidentified factors on ecotoxicity test outcomes and derived EFs.

Additionally, analyzing results for specific subgroups of chemicals could enhance our understanding. Nevertheless, a limitation of this study is that chemicals have not been categorized into distinct subgroups such as organic, inorganic, elemental, organometallic, and petroleum products, among others. This limitation stems from the absence of available extrapolation factors needed to convert various endpoints to chronic EC50 and chronic EC10 for different chemical subgroups. This study main aim was to provide EFs that aid in calculating CFs for chemicals, rather than facilitating comparisons between distinct chemical subgroups. However, investigating how different subgroups of chemicals exhibit varying ranges of EFs represents a potential direction for future research. While such a comparative analysis of chemical groups aligns with the future scope of this study, it falls outside the limited scope of the current study.

In terms of applications, USEtox CFs are typically useful for a first-tier assessment in potential toxicity assessment. However, if a substance appears to significantly contribute to the toxicity impact scores, it is recommended to verify and improve the reliability of the chemical-specific input data whenever possible. USEtox classifies ecotoxicity CFs as indicative or recommended based on data coverage across species and trophic levels. Recommended CFs require effect data from at least three different species covering at least three different trophic levels [12]. The scarcity of toxicity data points per chemical for various species at desired endpoints is a notable

limitation in deriving EFs. To address this, Table S1 in the supplementary information of this study details the number of data points, species, and species groups per chemical, highlighting the uncertainties associated with each EF due to limited data availability. Moreover, uncertainty can also arise when the available data require conversion to the desired format through extrapolation methods, such as changing from acute to chronic exposure classes or changing endpoints. The available extrapolation and conversion factors are derived from limited data, contributing further to the uncertainty of the results.

Conclusions

The escalating use of chemicals presents both opportunities and dilemmas. While their benefits are evident, concerns about their potential environmental and health impacts are growing, prompting questions about the sustainability of current practices. This underscores the need to transition from a preliminary lifecycle perspective to a comprehensive LCA that quantifies the environmental concerns of chemicals. Central to this is the availability of EFs for chemical emissions. While risk assessment tools have been employed for years, generating vast data, there is a pressing need to consolidate this information for LCA applications. Drawing from the REACH and CompTox databases, this study identified 11,295 unique chemicals, of which only 2426 are currently cataloged in the USEtox database. To address this, this study calculated EFs for an additional 8869 chemicals, enhancing their representation in LCA studies. In total, this research covers 11,368 chemicals, with 2426 overlapping with USEtox. However, data scarcity posed challenges, leading to employing QSAR models like ECOSAR and TEST for estimating ecotoxicity data.

The strong correlation between the experimental EFs and those in the USEtox database underscores a high degree of confidence in the calculations for chemicals not yet included in USEtox. However, discrepancies arise when comparing USEtox EFs values with derived EC50 chronic values, likely due to variations in foundational ecotoxicity data. This correlation improves when both chronic and extrapolated EC50 data are considered, emphasizing the value of extrapolated data in broadening the data set and species representation. Conversely, the weak correlation between estimated EFs and those in USEtox, as well as between EFs derived from ECOSAR and TEST, suggests that different QSAR models yield varied results. This divergence between estimated and experimental data from REACH and CompTox databases underscores the need for caution when relying on estimated data. As chemical data continues to evolve, EFs should be periodically updated to remain current and accurate, underscoring their dynamic nature in reflecting

the most recent data. There is also a recognized need to enhance consensus between risk assessment and USEtox-based LCA calculations. Achieving this alignment would optimize resources and allow for the reliable use of both methodologies in quantifying the toxicity impacts of chemicals.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12302-024-00947-0.

> Supplementary material 1.

Declarations

Competing interests
There are no competing of interest to declare.

References

1. Adkins, N. (2023). *COMPTOX_Public_ToxValDB v9.4* https://clowder.edap-cluster.com/datasets/61147fefe4b0856fdc65639b#folderId=645a5c0ce4b08a6b39438b10&page=0. Accessed 30 July 2023
2. Aurisano N, Albizzati PF, Hauschild M, Fantke P (2019) Extrapolation factors for characterizing freshwater ecotoxicity effects. Environ Toxicol Chemis 38(11):2568–2582. https://doi.org/10.1002/etc.4564
3. Bălan SA, Andrews DQ, Blum A, Diamond ML, Fernández SR, Harriman E, Lindstrom AB, Reade A, Richter L, Sutton R, Wang Z, Kwiatkowski CF (2023) Optimizing Chemicals management in the united states and Canada through the essential-use approach [Review]. Environ Sci Technol 57(4):1568–1575. https://doi.org/10.1021/acs.est.2c05932
4. Bare JC (2006) Risk assessment and life-cycle impact assessment (LCIA) for human health cancerous and noncancerous emissions: integrated and complementary with consistency within the USEPA. Hum Ecol Risk Assess Int J 12(3):493–509. https://doi.org/10.1080/10807030600561683

5. Barrett H (2000) Six billion and counting: trends and prospects for global population at the beginning of the twenty-first century [Article]. Geography 85(2):107–120

6. Benfenati E, Manganaro A, Gini GC (2013) VEGA-QSAR AI inside a platform for predictive toxicology. PAI@ AI* IA 1107:21–28

7. Blázquez M, Corral B, Buist H, Ligthart T, Henzing B, Rosenbaum R, Cajaraville MP. (2022). On the combination of USEtox® and SimpleBox 4 Nano models for the derivatization of size-dependent characterization factors for engineered nanomaterials. E3S Web of Conferences,

8. CAS. (2023). *CAS Content*. https://www.cas.org/about/cas-content

9. Commission E. (2020). *COM/2020/16 final, REPORT FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT AND THE COUNCIL 2019 report on the statistics on the use of animals for scientific purposes in the Member States of the European Union in 2015–2017*. Publications Office of the European Union. http://op.europa.eu/en/publication-detail/-/publication/04a89 0d4-47ff-11ea-b81b-01aa75ed71a1

10. Diamond ML, de Wit CA, Molander S, Scheringer M, Backhaus T, Lohmann R, Arvidsson R, Bergman T, Hauschild M, Holoubek I, Persson L, Suzuki N, Vighi M, Zetzsch C (2015) Exploring the planetary boundary for chemical pollution. Environ Int 78:8–15. https://doi.org/10.1016/j.envint.2015.02.001

11. ECHA. (2023). *Registered substances - ECHA*. https://echa.europa.eu/information-on-chemicals/registered-substances

12. Fantke, P., Bijster, M., Guignard, C., Hauschild, M., Huijbregts, M., Jolliet, O., Kounina, A., Magaud, V., Margni, M., McKone, T. E., Posthuma, L., Rosenbaum, R., Van de meent, R., & Zelm, R. (2017). *USEtox 2.0 : Documentation (Version 1)*. https://usetox.org/

13. Fantke P, Ernstoff A (2018) LCA of chemicals and chemical products. In: Hauschild MZ, Rosenbaum RK, Olsen SI (eds) Life Cycle Assessment: Theory and Practice. Springer International Publishing, Cham, pp 783–815

14. Fantke P, Huang L, Overcash M, Griffing E, Jolliet O (2020) Life cycle based alternatives assessment (LCAA) for chemical substitution. Green Chem 22(18):6008–6024. https://doi.org/10.1039/d0gc01544j

15. Fantke P, Owsianiak M, Verones F, Huijbregts MAJ, Azevedo L, Chaudhary A, Cosme N, Baan LD, Fantke P, Hauschild M, Henderson AD, Jolliet O, Mutel CL, Owsianiak M, Pfister S, Preiss P, Roy PO, Scherer L, Steinmann Z, Hellweg S (2019) LC-IMPACT version 1.0. a spatially differentiated life cycle impact assessment approach, pp 118–134.

16. Henderson AD, Hauschild MZ, Van De Meent D, Huijbregts MAJ, Larsen HF, Margni M, McKone TE, Payet J, Rosenbaum RK, Jolliet O (2011) USEtox fate and ecotoxicity factors for comparative assessment of toxic emissions in life cycle analysis: sensitivity to key chemical properties. Int J Life Cycle Assess 16(8):701–709. https://doi.org/10.1007/s11367-011-0294-6

17. Holmquist H, Fantke P, Cousins IT, Owsianiak M, Liagkouridis I, Peters GM (2020) An (Eco)toxicity life cycle impact assessment framework for per- and Polyfluoroalkyl substances [Article]. Environ Sci Technol 54(10):6224–6234. https://doi.org/10.1021/acs.est.9b07774

18. Hou P, Jolliet O, Zhu J, Xu M (2020) Estimate ecotoxicity characterization factors for chemicals in life cycle assessment using machine learning models. Environ Inter. https://doi.org/10.1016/j.envint.2019.105393

19. ICCA. (2011). ICCA Review 2010. *ICCA Review 2010*, 1–32. https://icca.cefic.org/ICCADocs/ICCA-review-2009-2010.pdf. Accessed 1 Dec 2023

20. Isarankura-Na-Ayudhya C, Naenna T, Nantasenamat C, Prachayasittikul V (2009) A practical overview of quantitative structure-activity relationship. EXCLI J 8:874–888

21. Iwasaki Y, Kotani K, Kashiwada S, Masunaga S (2015) Does the Choice of NOEC or EC10 affect the hazardous concentration for 5% of the species? Environ Sci Technol 49(15):9326–9330. https://doi.org/10.1021/acs.est.5b02069

22. Jacobs MM, Malloy TF, Tickner JA, Edwards S (2016) Alternatives assessment frameworks: research needs for the informed substitution of hazardous chemicals. Environ Health Perspect 124(3):265–280. https://doi.org/10.1289/ehp.1409581

23. Llanos EJ, Leal W, Luu DH, Jost J, Stadler PF, Restrepo G (2019) Exploration of the chemical space and its three historical regimes. Proc Natl Acad Sci 116(26):12660–12665. https://doi.org/10.1073/pnas.1816039116

24. Malaj E, von der Ohe PC, Grote M, Kühne R, Mondy CP, Usseglio-Polatera P, Brack W, Schäfer RB (2014) Organic chemicals jeopardize the health of freshwater ecosystems on the continental scale. Proc Natl Acad of Sci 111(26):9549–9554. https://doi.org/10.1073/pnas.1321082111

25. Martin T. (2020). User's guide for TEST (Toxicity Estimation Software Tool) version 5.1-a java application to estimate toxicities and physical properties from molecular structure. *US Enviorn Prot. Agency*.

26. Mayo-Bean K, Moran-Bruce K, Nabholz J, Meylan W, Howard P, Cassidy L. (2012). Operation manual for the ecological structure-activity relationship model (ECOSAR) Class Program. *Risk Assessment Division*, *US EPA*.

27. McCarty LS, Borgert CJ, Posthuma L (2018) The regulatory challenge of chemicals in the environment: toxicity testing, risk assessment, and decision-making models. Regul Toxicol Pharmacol 99:289–295. https://doi.org/10.1016/j.yrtph.2018.10.001

28. Munthe J, Brorström-Lundén E, Rahmberg M, Posthuma L, Altenburger R, Brack W, Bunke D, Engelen G, Gawlik BM, van Gils J, Herráez DL, Rydberg T, Slobodnik J, van Wezel A (2017) An expanded conceptual framework for solution-focused management of chemical pollution in European waters. Environ Sci Eur 29(1):13. https://doi.org/10.1186/s12302-017-0112-2

29. Ogunseitan OA (2023) Chemicals management approach to sustainable development of materials [Review]. MRS Bull 48(4):368–374. https://doi.org/10.1557/s43577-023-00518-3

30. Owsianiak M, Hauschild MZ, Posthuma L, Saouter E, Vijver MG, Backhaus T, Douziech M, Schlekat T, Fantke P (2023) Ecotoxicity characterization of chemicals: global recommendations and implementation in USEtox. Chemosphere. https://doi.org/10.1016/j.chemosphere.2022.136807

31. Praetorius A, Tufenkji N, Goss K-U, Scheringer M, Von der Kammer F, Elimelech M (2014) The road to nowhere: equilibrium partition coefficients for nanoparticles. Environ Sci Nano 1(4):317–323

32. Pratt IS (2002) Global harmonisation of classification and labelling of hazardous chemicals. Toxicol Lett 128(1):5–15. https://doi.org/10.1016/S0378-4274(01)00529-X

33. REACH. (2020). *REACH Dossier Database*. https://echa.europa.eu/fr/search-for-chemicals. Accessed Aug 2020

34. Rockström J, Steffen W, Noone K, Persson A, Chapin Iii FS, Lambin E, Lenton TM, Scheffer M, Folke C, Schellnhuber HJ, Nykvist B, de Wit CA, Hughes T, van der Leeuw S, Rodhe H, Sörlin S, Snyder PK, Costanza R, Svedin U, Foley J (2009) Planetary boundaries: exploring the safe operating space for humanity. Ecol Soc. https://doi.org/10.5751/ES-03180-140232

35. Rosenbaum RK (2015) Ecotoxicity. In: Hauschild MZ, Huijbregts MAJ (eds) Life Cycle Impact Assessment. Springer Netherlands, Dordrecht, pp 139–162

36. Rosenbaum RK, Bachmann TM, Gold LS, Huijbregts MAJ, Jolliet O, Juraske R, Koehler A, Larsen HF, MacLeod M, Margni M, McKone TE, Payet J, Schuhmacher M, van de Meent D, Hauschild MZ (2008) USEtox—the UNEP-SETAC toxicity model: recommended characterisation factors for human toxicity and freshwater ecotoxicity in life cycle impact assessment. Inter J Life Cycle Assess 13(7):532–546. https://doi.org/10.1007/s11367-008-0038-4

37. Rosenbaum RK, Hauschild MZ, Boulay AM, Fantke P, Laurent A, Núñez M, Vieira M (2017) Life cycle impact assessment. In: Michael ZH, Ralph KR, Stig Irving O (eds) Life Cycle Assessment. Springer International Publishing, Cham, pp 167–270

38. Sala S, Biganzoli F, Mengual ES, Saouter E (2022) Toxicity impacts in the environmental footprint method: calculation principles. Inter J Life Cycle Assess 27(4):587–602

39. Saouter E, Aschberger K, Fantke P, Hauschild MZ, Bopp SK, Kienzler A, Paini A, Pant R, Secchi M, Sala S (2017) Improving substance information in USEtox®, part 1: discussion on data and approaches for estimating freshwater ecotoxicity effect factors. Environ Toxicol Chem 36(12):3450–3462

40. Saouter E, Aschberger K, Fantke P, Hauschild MZ, Kienzler A, Paini A, Pant R, Radovnikovic A, Secchi M, Sala S (2017) Improving substance information in USEtox®, part 2: data for estimating fate and ecosystem exposure factors. Environ Toxicol Chem 36(12):3463–3470. https://doi.org/10.1002/etc.3903

41. Saouter E, Biganzoli F, Ceriani L, Versteeg D, Crenna E, Zampori L, Sala S, Pant R (2020) Environmental footprint: update of life cycle impact assessment methods-ecotoxicity freshwater, human toxicity cancer, and non-cancer. Publications Office of the European Union, Luxembourg

42. 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. Int Environ Assess Manage 15(5):783–795. https://doi.org/10.1002/ieam.4168

43. Saouter E, Wolff D, Biganzoli F, Versteeg D (2019) Comparing options for deriving chemical ecotoxicity hazard values for the European union environmental footprint, part II. Integr Environ Assess Manag 15(5):796–807. https://doi.org/10.1002/ieam.4169

44. Schaafsma G, Kroese ED, Tielemans ELJP, Van de Sandt JJM, Van Leeuwen CJ (2009) REACH, non-testing approaches and the urgent need for a change in mind set. Regul Toxicol Pharmacol 53(1):70–80. https://doi.org/10.1016/j.yrtph.2008.11.003

45. Schymanski EL, Zhang J, Thiessen PA, Chirsir P, Kondic T, Bolton EE (2023) Per- and polyfluoroalkyl substances (PFAS) in pubchem: 7 million and growing. Environ Sci Technol 57(44):16918–16928. https://doi.org/10.1021/acs.est.3c04855

46. Steffen W, Richardson K, Rockström J, Cornell SE, Fetzer I, Bennett EM, Biggs R, Carpenter SR, De Vries W, De Wit CA, Folke C, Gerten D, Heinke J, Mace GM, Persson LM, Ramanathan V, Reyers B, Sörlin S (2015) Planetary boundaries: guiding human development on a changing planet. Science. https://doi.org/10.1126/science.1259855

47. Tickner JA, Simon RV, Jacobs M, Pollard LD, van Bergen SK (2021) The nexus between alternatives assessment and green chemistry: supporting the development and adoption of safer chemicals. Green Chem Lett Rev 14(1):23–44. https://doi.org/10.1080/17518253.2020.1856427

48. U.S.EPA. (2020). User's Guide for T.E.S.T. (version 5.1) (toxicity estimation software tool): a program to estimate toxicity from molecular structure [Data and Tools]. https://www.epa.gov/chemical-research/toxicity-estimation-software-tool-test. Accessed 1 Dec 2023

49. U.S.EPA. (2023a). *ECOTOX | Explore*. September https://cfpub.epa.gov/ecotox/explore.cfm?sub=Species. Accessed 1 Sept 2023

50. U.S.EPA. (2023b). *Estimation programs interface suite™ for microsoft® windows, v 4.11*. March https://www.epa.gov/tsca-screening-tools/epi-suitetm-estimation-program-interface. Accessed 9 Mar 2023

51. UNEP. (2019). (United Nations Environment Programme (UNEP), global chemicals outlook II: from legacies to innovative solutions, 2Nd Edn., Issue. https://www.unep.org/resources/report/global-chemicals-outlook-ii-legacies-innovative-solutions. Accessed 12 Dec 2023

52. Van Zelm R, Huijbregts MA, van de Meent D (2010) Transformation products in the life cycle impact assessment of chemicals. Environ Sci Technol 44:1004

53. Verones F, Hellweg S, Azevedo L, Chaudhary A, Cosme N, Fantke P, Goedkoop M, Hauschild M, Laurent A, Mutel C. (2019). LC-Impact Version 0.5: a spatially differentiated life cycle impact assessment approach. *Adv Access publ*.

54. Wang Z, Walker GW, Muir DCG, Nagatani-Yoshida K (2020) Toward a global understanding of chemical pollution: a first comprehensive analysis of national and regional chemical inventories. Environ Sci Technol 54(5):2575–2584. https://doi.org/10.1021/acs.est.9b06379

55. Wolfram J, Stehle S, Bub S, Petschick LL, Schulz R (2021) Water quality and ecological risks in European surface waters—monitoring improves while water quality decreases. Environ Int 152:106479. https://doi.org/10.1016/j.envint.2021.106479

Publisher's Note