Uncertainty_analysis

Oskar Nyberg

2023-04-17

Contents

Uncertainty analysis of toxicological characterization factors at the HC20 response level

Oskar Nyberg[1,†], Reinout Heijungs[2], Patrik Henriksson[3,4,5]

[1]Department of Ecology, Environment and Plant Sciences, Svante Arrhenius väg 20 A, Stockholm University
[2] Leiden University
[3] Stockholm Resilience Center, XXX, Stockholm University
[4] WorldFish Center
[5] Beijer Institute

† Corresponding author:oskar.nyberg@su.se

Abstract

We calculate characterization factors for chemicals with potential negative environmental impact by constructing a database with all openly available ecotoxicological effect data for a set of 16797 substances

Introduction

Increasing chemical use is a major concern for operating within the safe space of planetary boundaries, and agriculture is a major driving force behind this increase, through pesticide use, veterinarian drugs, and disinfectants (?Gordon et al. 2017; ?Persson et al. 2022). Chemical use in agriculture is also expected to continue to increase in the coming decade, as a result of larger production volumes and more intensive production systems (Schreinemachers and Tipraqsa 2012). However, data on potentially toxic chemicals used for different farming systems, toxicity evaluations for different compounds, and variability in toxicity potentials are incomplete and fragmented, hampering accurate global comparisons of ecotoxicity impacts of different food products. Accurate toxicological characterization of chemicals are essential for ensuring the safety of human health and the environment, where proper toxicological characterization involves identifying and evaluating the potential adverse effects of chemicals, determining the level of exposure that is safe for humans and the environment, and assessing the risk posed by exposure to chemicals (Krewski et al. 2010). Environmental Risk Assessment (ERA) has been the primary tool for evaluating toxicological risks at fields and farms, but it is limited to individual processes in production chains (e.g. grow-out). This can be limiting when trying to understand the overall impacts of food products, including e.g. all ingredients used to produce animal compound feeds. Subsequently, life cycle assessments (LCAs) have increasingly been used as a complementing framework to benchmark toxicological impacts of agrifood products. LCA is an ISO-standardized environmental framework that seeks to aggregate the emissions and products that are needed throughout a value chain, and characterize these towards one or more environmental impacts. The results are also scaled to a predefined unit of reference (functional unit), which allows for comparisons to be made between products (e.g. in terms of kg. food), functions (e.g. in terms of kcal), or services (e.g. washing one plate). It is most commonly encountered as the methodology behind carbon footprints, but is also used to quantify freshwater consumption, land occupation, eutrophication, biodiversity loss, or toxicity impacts (Hauschild and Huijbregts 2015). Toxicity impacts, in turn, are commonly evaluated in terms of human toxicity, either including or excluding cancer cases, and ecotoxicity impacts on freshwater, marine, or terrestrial ecosystems (Hauschild and Huijbregts 2015).

LCA is cruder than ERA when it comes to evaluating toxicological effects, as it disregards temporal scales and critical concentrations, assumes inputs equal to emissions, simplifies assumptions related to release location and exposure, and disregards 'cocktail effects' between chemicals (Fantke et al., 2015). In the meantime, it can capture toxicological impacts throughout whole value chains, from mercury emissions from coal-fired power plants to therapeutants used in aquaculture ponds. LCAs can thereby provide useful insights into where in value chains the largest toxicity reductions can be achieved, such as that most freshwater ecotoxicological impacts related with prawn farming in Bangladesh are related to the production certain agricultural materials used for feeds, and not the prawn grow-out (Henriksson et al. 2015). To assess ecotoxicological impacts throughout the value chain of a product, ecotoxicological characterization factors for chemicals are used within as indicators for potential impact the life cycle impact assessment (LCIA) phase of an LCA. Toxicological effect data for chemicals provide the base for characterization factor calculations within life cycle assessments. These data, however, are notoriously heterogeneous since they are reported across thousands of tested species at variable concentrations, measured effects and at various empirical or modeled endpoints, and as such provide a fair amount of uncertainty to characterization factor calculations. To investigate the magnitude of uncertainties for the ecotoxicological concentration-response slope factor (CRF) at the 20% response level (HC20) we have queried the OECD QSAR Toolbox for ecotoxicological effect data for the 16797 chemicals included in the HESTIA inventory and gathered all non-proprietary records available. Additional data was sourced from EnviroTox, a curated ecotoxicological database for 4267 substances (Connors et al. 2019).

HESTIA is a free open-access platform that provides a data repository for life cycle inventory data using a harmonized schema and glossary of terms, and calculations tools for various emissions and impact assessments. The ambition of HESTIA is to make environmental benchmarks of agrifood commodities more accessible and transparent, by providing a free harmonized framework. Given that the agrifood sector is a major user of potentially toxic chemicals, mainly in terms of pesticides and therapeutants, we deem it important to provide as a complete and accurate set of ecotoxicity potentials as possible.

The evaluation of toxicological impacts in LCA is limited by the number of chemical compounds characterized

by impact assessment methodologies. The characterization factor is the value used to translate the amounts of chemicals used to its potential toxicity impact. There are several different impact assessment methodologies to derive these characterization factors, including USES-LCA v1&2 (Huijbregts et al. 2000; Van Zelm, Huijbregts, and De Meent 2009), IMPACT 2002 (Pennington et al. 2005), and UNEP-SETAC's USEtox (Fantke 2017), but ultimately they all rely on fundamental data on toxicity and physicochemical properties. Among these impact assessment methodologies, the USEtox model is most widely used at present, and also the one promoted by the European Platform on Life Cycle Assessment (ILCD, 2010). The USEtox model is currently on version 2.1, and readily presents 2499 freshwater ecotoxicological characterization factors for different chemicals, derived from physicochemical properties and empirical data at the 50% response level (EC50), presented as the 50th percentile on the SSD curve (HC50) (ref). For chemicals not readily characterized, the USETox model also allows users to derive their own characterization factors through an Excel spreadsheet, but this is, in turn, dependent on access to chemical properties.

Based on a series of recent expert workshops, updated recommendations on which toxicological input data should form the SSD slopes for characterization of substances has been published (Mikolaj Owsianiak et al. 2019). The authors argued for the use of effect data at concentrations within a range similar to ambient environmental concentrations, implying the use of effect data from no observed effect concentration (NOEC), lowest observed effect concentration (LOEC), or effect concentrations at 0, 10 and 50% response level with the 20th percentile as a working point on the SSD curve (HC20). This implies harmonizing all different effect data into one coherent effect concentration or equivalent thereof, recommended as effect concentration at the 10% response level equivalent data (EC10eq) and using as the input toxicity parameter for the USETox model (Aurisano et al. 2019). In october 2022, Sala et al. (2022) publish the EU environmental footprint (EF v3.1) database database with ecotoxicological characterization factors for 6,038 chemicals based on data from the REACH toxicological database using the HC20EC10eq impact score (Saouter et al. 2019; Sala et al. 2022). While this is a huge improvement, it is still far from the complete list of potentially toxic chemicals encountered in agrifood production. For example, the US-EPA presents a list of almost 17,000 organic and inorganic chemicals with potential toxicity (Ref?).**Not sure we should discuss any particular database**

The EF v3.1 database is substantially (140%) larger than the former USEtox v2.1 database of toxicological characterization factors, easily accessible online and presents substances' physicochemical and toxicological data as well as characterization factors for each substance respectively. The database is constructed on toxicological data from three repositories; the OpenFoodTox database, the Pesticide Property Database, the Registration, Evaluation and Restriction och Chemicals (REACH) Database as well as the USEtox v2.1 database when data was not available elsewhere but the majority of toxicological data is gathered from the REACH database. Unfortunately, a substantial part of the toxicological records within REACH-DB are proprietary substance registrations which are protected by confidentiality agreements and have been made unavailable to the public. This is unfortunate, as one of the 19 key recommendations of the Ecotoxicity Task Force and the Pellston workshop that was held in 2018 in Valencia, Spain is to "use data that has a traceable origin" (Mikołaj Owsianiak et al. 2022). Since the present HESTIA database was under construction at the release of EF v3.1, data sources overlap only to a minor extent, and original data sources are accessible within the HESTIA database, the continued assembly of the HESTIA database was motivated.

The present objective is to derive freshwater ecotoxicity potential characterization factors, including uncertainty estimates at the 20th percentile working point, to better inform environmental impact calculations for agrifood LCAs. With a large database built on openly available ecotoxicological data, we are able to investigate and report statistical uncertainty for the generated characterization factors, as is prompted by (Mikołaj Owsianiak et al. 2022). Additionally, we estimate ecotoxicological data based on quantitative structure-activity relationships (QSAR) for evaluating the applicability of such data in covering more substances or complement toxicological data based on animal testing. The ultimate goal of these outcomes is to support environmental impact calculations on the HESTIA.earth online platform.

Methods

HESTIA Database construction.

The starting point for constructing a database with ecotoxicological effect data to calculate characterization factors suited for the online life cycle assessment (LCA) application HESTIA (http://HESTIA.earth) is the substance inventory: `PesticideAI.csv`. This is a list of 16797 CAS registry numbers (CASRN) and chemical names for potentially harmful substances based on the United States Environmental Protection Agency's Substance Registry Services (USEPA SRS) inventory.

CASRN and chemical names were queried to match SMILES configurations acquiring from the NCBI PubChem database using the R package Webchem (Szöcs et al. 2020). Based off of the CASRN-SMILES matches three actions are taken: 1) substance use patterns information is gathered from several repositories for each substance respectively to categorize chemicals into groups: Antibiotic, Antiviral, Other inorganic chemicals, Other organic chemicals, PPCP, Pesticide, Pharmaceutical, and Unknown, 2) query the OECD QSAR Toolbox (Dimitrov et al. 2016) for physicochemical properties for substances and subsequently read and wrangle data for a USEtox-friendly format, and 3) query the OECD QSAR Toolbox for aquatic ecotoxicological records from the following toxicological endpoints: EC10, EL10, IC10, LC10, LOEC (grouped as "EC10"); EC50, EL50, IC50, LC50 (grouped as "EC50"); EC0, LC0, NOAEC, NOEC, NOER, NOEL (grouped as "NOEC"), according to the methodology of (Aurisano et al. 2019). For details on the processing of data, please see supporting information (`code/Pesticide_annotations.Rmd`, `Physchem_read_wrangle_function.R`, `data/raw_data_read_and_wrangle.R`, and`HESTIA_HC20_DB.Rmd`).

Merging the EnviroTox database:

While using the HESTIA toxicological database as a base, the EnviroTox toxicological database (Connors et al., 2019) is a curated aquatic toxicological database containing a large set of toxicological data from a broad range of potentially toxic substances. Prior to merging the two databases, curation of the EnviroTox database was performed accordingly:
Toxicological endpoints were selected and grouped as for HESTIA database (Table STX), Acute and Chronic test duration definitions were performed according to taxonomic group and duration of experiment. The EnviroTox database has put a lot of effort into a very refined acute/chronic exposure selection methodology. For the current purpose, however, we chose to categorize acute/chronic definitions in a harmonized (e.g. less refined) way according to the methodology described above. Taxonomic information from the EnviroTox database was slightly revised to match the HESTIA dataset (e.g minor spelling corrections). With harmonized acute/chronic definitions, identical taxonomic descriptions and endpoint conversions, extrapolation factors for EC10eq conversions were added to effect data. Since several database queries overlap between the HESTIA and EnviroTox databases, prioritization of the HESTIA dataset was chosen. Adding on to the HESTIA database are records from EnviroTox where 1) substances are unique to EnviroTox, and 2) Species, but not substances are unique to EnviroTox.

From raw data to the nls output

Based on the assumption of lognormal distribution of ecotoxicological effect data, we assess the investigate the uncertainties of the $\mathrm{CRF_{HC20}}$ by fitting each substance to a nonlinear least squares model to investigate the minimum μ and σ at the $logHC20_{EC10eq}$ respectively. First, we define the raw data. For a specific substance x, we have $EC10^{eq}$ data for species i, done in a number of experiments, labeled as $k = 1, \ldots, n_i$. The data are thus indicated as $EC10^{eq}_{i,k}$, Throughout the analysis, we will work with the logarithm of the data, and for conciseness the data are indicated as $L_{(i,k)}$. Thus

$$L_{i,k} = log(EC10^{eq}_{i,k}) \tag{1}$$

The per-species average over all samples are found as

$$M_i = \frac{1}{n_i} \sum_{k=1}^{n_i} L_{i,k} \tag{2}$$

and the corresponding standard deviation as

$$S_i = \sqrt{\frac{1}{n_i - 1} \sum_{k=1}^{n_i} (L_{i,k} - M_i)^2} \tag{3}$$

The average values M_i have a standard error E_i that is given by

$$E_i = \frac{s_i}{\sqrt{n_i}} = \sqrt{\frac{1}{n_i(n_i - 1)} \sum_{k=1}^{n_i} (L_{i,k} - M_i)^2} \tag{4}$$

Next, we make the assumption that species sensitivity follows a lognormal distribution. This is in agreement with a mainstream practice of ecologists (ref Posthuma et al?). Because we have transformed the data with a logarithm, we now have a normal distribution for the M_i-values. This distribution is characterized by two parameters, the mean (μ) and the standard deviation (σ), which are traditionally estimated as follows:

$$\hat{\mu} = \frac{1}{m} \sum_{i=1}^{m} M_i \tag{5}$$

$$\hat{\sigma} = \sqrt{\frac{1}{(m-1)} \sum_{i=1}^{m} (M_i - M)^2} \tag{6}$$

Also traditionally, the HC20 is found as the 20-percentile value of the distribution:

$$HC20 = \hat{\mu} + z_{0.2}\hat{\sigma} \tag{7}$$

where $z_{0.2}$ is approximately -0.84.
The question is now what we can say about the uncertainty in the value of $logHC20_{EC10eq}$. This basically depends on two elements: the uncertainty in $\hat{\mu}$ and the uncertainty in $\hat{\sigma}$. Both of these depend on the degree of fit with the normal distribution, and with the intra-species variation. Our approach will be to fit a normal distribution to the vector of mean values M_i, weighted with the reciproca 1 of the variance $(\frac{1}{E_i})^2$. Suppose we fit a function $F(x; \mu, \sigma)$ to the cumulative distribution of M_i-values, each of which is associated with a standard error E_i. We can find the optimal values of μ and σ through a least-squares minimization of the residual:

$$min \left(\sum_{i=1}^{m} \frac{1}{E_i^2} (y_i - F(m_i; \mu, \sigma))^2 \right) \tag{8}$$

Here, y_i denotes the order of appearance in the cumulative form. More precisely,

$$y_i = \frac{rank(M_i - 0.5)}{m} \tag{9}$$

For F, we take the cumulative normal distribution, given by

$$F(m_i; \mu, \sigma) = \frac{1}{2} + \frac{1}{2} erf \left(\frac{x - \mu}{\sigma\sqrt{2}} \right) \tag{10}$$

To model estimates for $\hat{\mu}$ and $\hat{\sigma}$ we fit a nonlinear least-squares (nls) regression model to the toxicological dose-response data using R-programming language. The script defines the cumulative normal distribution function

Table 1: Summary of HESTIA Environmental Toxicology dataset by taxonomic group and endpoint

Taxonomy.Group	EC10_Chronic	EC50_Chronic	NOEC_Chronic	EC10_Acute	EC50_Acute	NOEC_Acute
Algae	2659	7021	3187	NA	NA	NA
Amphibian	577	325	1402	540	2170	564
Annellidae	43	32	110	76	704	44
Crustacean	3169	4579	5068	755	12360	2059
Fish	3845	2115	5900	2564	31135	5282
Insect	344	433	966	284	6872	379
Mollusca	343	241	691	199	2640	279
Others	684	2263	950	53	251	86
Plant	1383	1837	1369	217	248	152
Rotifera	322	532	557	NA	NA	NA
Total	13369	19378	20200	4688	56380	8845

(Eq. 10) and a self-starting function for the cumulative normal distribution to solve for the least squares residual of μ and σ, programming in R allow the use of the `nls()`function for solving cumulative normal distribution functions. Starting points the model are defined as $\mu_{start} = \frac{1}{m}\sum_{i=1}^{m} M_i$ and $\sigma_{start} = \frac{1}{m}\sum_{i=1}^{m} S_i$ respectively. In cases where only one toxicological record per species is available, σ_{start} is not obtainable thus an arbitrary value of $\sigma_{start} = 1$ is used. Then, the function selects the unique CAS numbers and checks that there are at least 5 distinct species and 3 distinct taxonomic groups for each CAS number. If there is insufficient data, the function skips to the next CAS number. If there is sufficient data, the function calculates the mean and standard deviation of the log-transformed EC10eq values for each species within each taxonomic group. The mean values are then ranked and the self-starting function is used to fit a cumulative normal distribution to the ranked means. For each CAS number, the function outputs a list that contains the CAS number, the $logHC20_{EC10eq}$, the concentration-response slope factor (CRF), the estimated parameters of the cumulative normal distribution (mu and sigma), the 2.5 and 97.5 quantiles of the parameter estimates, the number of species and taxonomic groups used in the analysis, the number of records (i.e., EC10 values) for the CAS number, the status of the analysis (i.e., "not enough data" or "converged"), and the raw output from the nls regression (nls_results). If specified, the function also creates a plot of the dose-response curve for each CAS number.

Results

We need to start with a summary of the HESTIA Toxicological database.
The curated data set contain toxicological data with 122860 records across 2210 species (Table 1), adapted for freshwater aquatic ecotoxicological potential (e.g. ecotoxicological effect factor) calculations at HC20EC10eq (hazard concentration where 20% mortality is expected at the EC10eq concentration) for a set of 3705 chemicals. Additionally, physicochemical properties required for freshwater aquatic toxicity potential characterization in USEtox was gathered for 13715 chemicals.

Describe the grouping summary results

From the available chemicals that we apply the nls model to, we are able to fit 1197 chemicals that have enough data to be able fit the model. Data availability is a major issue here, since 3798 (23.963964% of all records) have insufficient data (see Table 4).

To allow for comparative assessment of uncertainties at the $logHC20_{EC10eq}$ response level across the entire data set, we define uncertainty ratios(U_r) (or perhaps σ_r) given by 0.5 * 2.5 to 97.5 quantiles range divided by the central value ($logHC20_{EC10eq}$) by Eq 11

$$U_r = \frac{0.5(Q2.5 - Q97.5)}{logHC20_{EC10eq}} \tag{11}$$

Table 2: Summary of the records from EnviroTox dataset joined to HESTIA by taxonomic group and endpoint

Taxonomy.Group	EC10_Chronic	EC50_Acute	EC50_Chronic	NOEC_Acute	NOEC_Chronic	EC10_Acute
Algae	213	1451	1	2	678	NA
Amphibian	NA	228	NA	NA	3	NA
Annellidae	NA	67	NA	NA	1	NA
Crustacean	51	3128	NA	NA	594	NA
Fish	30	8688	6	41	1013	NA
Insect	8	384	NA	NA	8	1
Mollusca	NA	294	1	NA	16	NA
Others	13	576	NA	NA	382	NA
Rotifera	NA	211	NA	NA	44	NA
NA	NA	225	NA	NA	NA	NA
Total	315	15252	8	43	2739	1

Table 3: Summary Toxicological use annotations identified for chemicals

Group	n
Antibiotic	31
Antiviral	6
Other inorganic chemicals	435
Other organic chemicals	9191
PPCP	219
Pesticide	2802
Pharmaceutical	1031
Unknown	1593

Table 4: Summary overview of the nonlinear least square model fit for the HESTIA ecotoxicological database

status	n_records	min_species_per_substance	max_species_per_s
OK	1197	5	
Warning: too many missing values	2	7	
not enough data	3756	1	
singular gradient matrix at initial parameter estimates	18	5	
warning: step factor reduced below 'minFactor'	22	5	

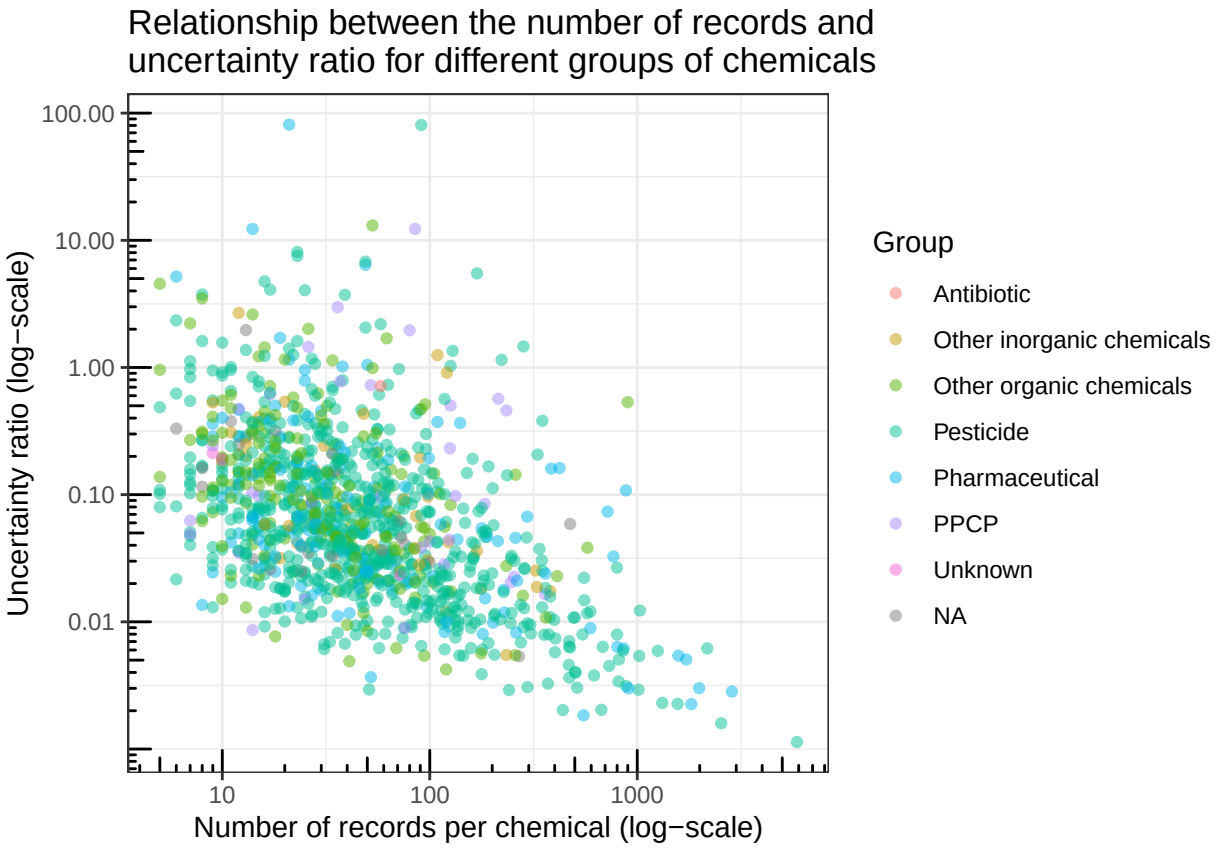
\begin{table}

\caption{Uncertainty ratios distribution across the dataset with minimum, 25%-quantile (Q1), median, 75%-quantile(Q3) and maximum}

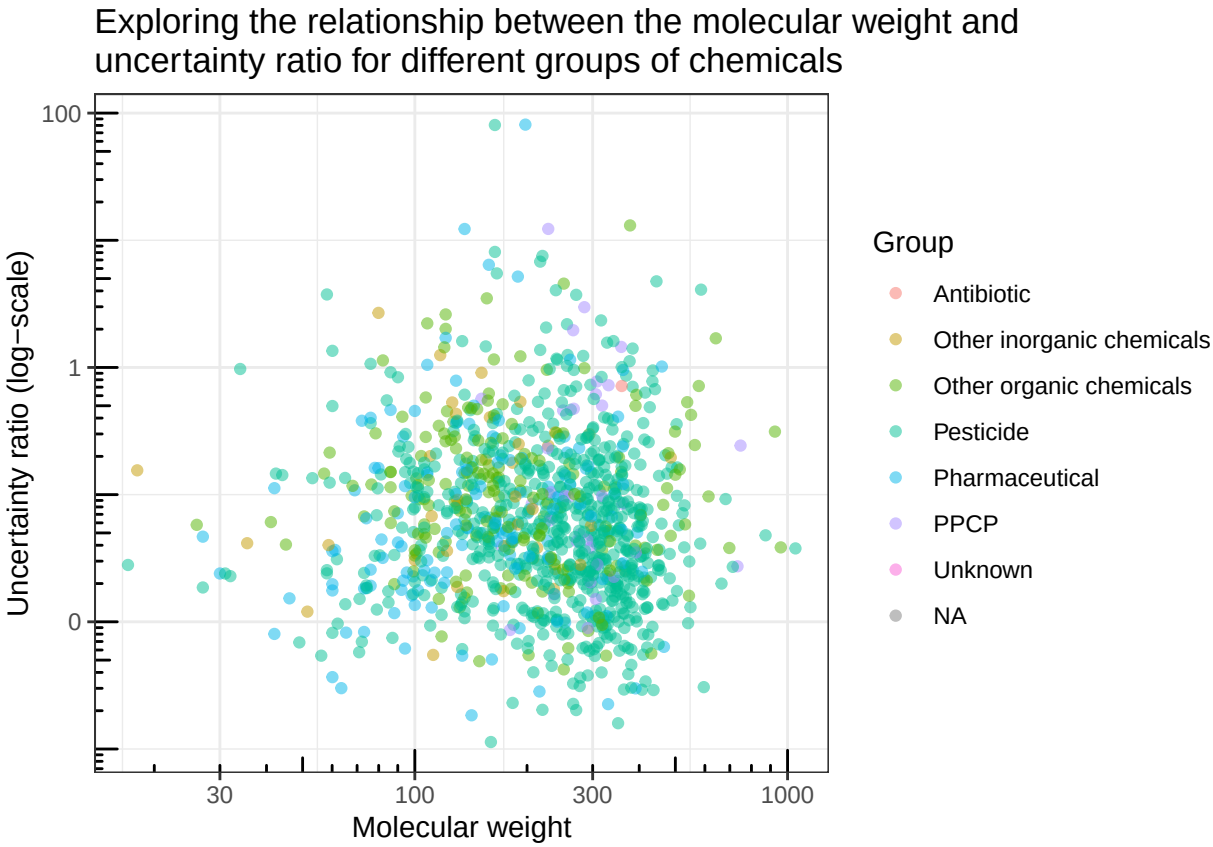
min	Q1	median	Q3	max
0.0011345	0.0235643	0.0574276	0.1609564	81.17155

\end{table}

When exploring the relationship between the number of records and uncertainty ratio for different groups of chemicals,



Uncertainty ratio variability as a result of molecular weight


Exploring the relationship between the molecular weight and uncertainty ratio for different groups of chemicals
Uncertainty ratio (log−scale)
100
1
0
30
100
300
1000
Molecular weight
Group
Antibiotic
Other inorganic chemicals
Other organic chemicals
Pesticide
Pharmaceutical
PPCP
Unknown
NA

References

Aurisano, Nicolò, Paola Federica Albizzati, Michael Hauschild, and Peter Fantke. 2019. "Extrapolation Factors for Characterizing Freshwater Ecotoxicity Effects." *Environmental Toxicology and Chemistry* 38 (11): 2568–82. https://doi.org/10.1002/etc.4564.

Connors, Kristin A., Amy Beasley, Mace G. Barron, Scott E. Belanger, Mark Bonnell, Jessica L. Brill, Dick de Zwart, et al. 2019. "Creation of a Curated Aquatic Toxicology Database: EnviroTox." *Environmental Toxicology and Chemistry* 38 (5): 1062–73. https://doi.org/10.1002/etc.4382.

Dimitrov, SD, R Diderich, T Sobanski, TS Pavlov, GV Chankov, AS Chapkanov, YH Karakolev, et al. 2016. "QSAR Toolbox–Workflow and Major Functionalities." *SAR and QSAR in Environmental Research* 27 (3): 203–19.

Fantke, Bijster, P. 2017. "USEtox® 2.0, Documentation Version 1." https://doi.org/10.11581/DTU:00000011.

Hauschild, Michael Z, and Mark AJ Huijbregts. 2015. *Introducing Life Cycle Impact Assessment*. Springer.

Henriksson, Patrik JG, Andreu Rico, Wenbo Zhang, Sk Ahmad-Al-Nahid, Richard Newton, Lam T Phan, Zongfeng Zhang, et al. 2015. "Comparison of Asian Aquaculture Products by Use of Statistically Supported Life Cycle Assessment." *Environmental Science & Technology* 49 (24): 14176–83.

Huijbregts, Mark AJ, U Thissen, JB Guinée, T Jager, D Kalf, D Van de Meent, AMJ Ragas, A Wegener Sleeswijk, and L Reijnders. 2000. "Priority Assessment of Toxic Substances in Life Cycle Assessment. Part i: Calculation of Toxicity Potentials for 181 Substances with the Nested Multi-Media Fate, Exposure and Effects Model USES–LCA." *Chemosphere* 41 (4): 541–73.

Krewski, Daniel, Daniel Acosta Jr, Melvin Andersen, Henry Anderson, John C Bailar III, Kim Boekelheide, Robert Brent, et al. 2010. "Toxicity Testing in the 21st Century: A Vision and a Strategy." *Journal of Toxicology and Environmental Health, Part B* 13 (2-4): 51–138.

Owsianiak, Mikolaj, Peter Fantke, Leo Posthuma, Erwan Saouter, Martina Vijver, Thomas Backhaus, Christian Schlekat, and Michael Z. Hauschild. 2019. "Ecotoxicity." In *Global Guidance on Environmental Life Cycle Impact Assessment*, edited by Rolf Frischknecht and Olivier Jolliet, 138–72. Paris, France: UNEP/SETAC Life Cycle Initiative.

Owsianiak, Mikołaj, Michael Z. Hauschild, Leo Posthuma, Erwan Saouter, Martina G. Vijver, Thomas Backhaus, Mélanie Douziech, Tamar Schlekat, and Peter Fantke. 2022. "Ecotoxicity Characterization of Chemicals: Global Recommendations and Implementation in USEtox." *Chemosphere*, no. 310 (October). https://doi.org/10.1016/j.chemosphere.2022.136807.

Pennington, David W, Manuele Margni, Christoph Ammann, and Olivier Jolliet. 2005. "Multimedia Fate and Human Intake Modeling: Spatial Versus Nonspatial Insights for Chemical Emissions in Western Europe." *Environmental Science & Technology* 39 (4): 1119–28.

Sala, Serenella, Fabrizio Biganzoli, Esther Sanye Mengual, and Erwan Saouter. 2022. "Toxicity Impacts in the Environmental Footprint Method: Calculation Principles." *The International Journal of Life Cycle Assessment* 27 (4): 587–602. https://doi.org/10.1007/s11367-022-02033-0.

Saouter, Erwan, Fabrizio Biganzoli, Rana Pant, Serenella Sala, and Donald Versteeg. 2019. "Using REACH for the EU Environmental Footprint: Building a Usable Ecotoxicity Database, Part I." *Integrated Environmental Assessment and Management* 15 (5): 783–95. https://doi.org/10.1002/ieam.4168.

Schreinemachers, Pepijn, and Prasnee Tipraqsa. 2012. "Agricultural Pesticides and Land Use Intensification in High, Middle and Low Income Countries." *Food Policy* 37 (6): 616–26.

Szöcs, Eduard, Tamás Stirling, Eric R. Scott, Andreas Scharmüller, and Ralf B. Schäfer. 2020. "webchem: An R Package to Retrieve Chemical Information from the Web." *Journal of Statistical Software* 93 (13): 1–17. https://doi.org/10.18637/jss.v093.i13.

Van Zelm, Rosalie, Mark AJ Huijbregts, and Dik van De Meent. 2009. "USES-LCA 2.0—a Global Nested Multi-Media Fate, Exposure, and Effects Model." Springer.