Hazard/Risk Assessment

Species Sensitivity Distributions for Use in Environmental Protection, Assessment, and Management of Aquatic Ecosystems for 12 386 Chemicals

Leo Posthuma,[a,b,]* Jos van Gils,[c] Michiel C. Zijp,[a] Dik van de Meent,[a,d,e] and Dick de Zwart[d,e]

[a]National Institute for Public Health and the Environment (RIVM), Centre for Sustainability, Environment and Health, Bilthoven, The Netherlands
[b]Department of Environmental Science, Institute for Water and Wetland Research, Faculty of Science, Radboud University, Nijmegen, The Netherlands
[c]Deltares, Delft, The Netherlands
[d]Mermayde, Groet, The Netherlands
[e]ARES, Odijk, The Netherlands

Abstract: The present study considers the collection and use of ecotoxicity data for risk assessment with species sensitivity distributions (SSDs) of chemical pollution in surface water, which are used to quantify the likelihood that critical effect levels are exceeded. This fits the European Water Framework Directive, which suggests using models to assess the likelihood that chemicals affect water quality for management prioritization. We derived SSDs based on chronic and acute ecotoxicity test data for 12 386 compounds. The log-normal SSDs are characterized by the median and the standard deviation of log-transformed ecotoxicity data and by a quality score. A case study illustrates the utility of SSDs for water quality assessment and management prioritization. We quantified the chronic and acute mixture toxic pressure of mixture exposures for >22 000 water bodies in Europe for 1760 chemicals for which we had both exposure and hazard data. The results show the likelihood of mixture exposures exceeding a negligible effect level and increasing species loss. The SSDs in the present study represent a versatile and comprehensive approach to prevent, assess, and manage chemical pollution problems. *Environ Toxicol Chem* 2019;38:905–917. © 2019 SETAC

Keywords: Aquatic ecotoxicity data; Chronic no-observed-effect concentration; Acute median effective concentration; Species sensitivity distribution; Environmental risk assessment; Life cycle assessment

INTRODUCTION

Human activities cause the emission of more than 100 000 chemical substances, with expected increases in compound diversity and emitted masses (United Nations Environment Programme 2013; European Chemicals Agency 2016; Bernhardt et al. 2017). This results in diverse ambient concentrations (e.g., the EMPODAT Database, Norman Association 2013), body residues (US Environmental Protection Agency 2009), ecological risks (Malaj et al. 2014), and eventual ecological and human health impacts (e.g., Vörösmarty et al. 2010; Hoekstra and Wiedmann 2014; Schäfer et al. 2016). Chemical pollution is a main driver of deterioration of freshwater biodiversity (Vörösmarty et al. 2010). Complementary policy approaches (chemical safety assessment, environmental quality assessment

and management, and product environmental footprints) are used to prevent and limit impacts of such pollution. These require ecotoxicity data and a method to convert these data into estimates of benchmark concentrations for no or negligible impacts (further referred to as "sufficiently protected") or in expected impact magnitudes of pollution (expressed as, e.g., species loss). Species sensitivity distributions (SSDs) support making both these conversions, for separate compounds and mixtures (De Zwart and Posthuma 2005).

An SSD reflects the observation that interspecies differences in sensitivity to a chemical resemble a bell-shaped distribution (on a log scale). An SSD is derived by fitting a selected statistical model (e.g., log normal) to compound-specific ecotoxicity data. Lack of ecotoxicity data is often mentioned as a reason that we have SSDs for only a few chemicals for current policy applications. Criteria for SSD data selection thereby vary among the policy approaches and jurisdictions, for example, for the minimum number of data points (taxa) needed and for minimum study quality characteristics (Posthuma et al. 2002). Despite that, SSDs are widely used for decision support (Supplemental Data,

Section 1). This likely relates to an observed association between SSD-predicted and observed biodiversity impacts (Posthuma and De Zwart 2012), to relative ease of use, and to representing a higher-tier approach compared with using benchmark concentrations.

Several hundreds of thousands of ecotoxicity test results are available globally, but currently few are used to derive SSDs. Often, strict criteria for SSD data selection (Klimisch et al. 1997; Moermond et al. 2016) and a minimum diversity of taxonomic groups and species (European Chemicals Agency 2008) are prescribed in SSD derivation. That has currently resulted in a low number of compounds with sufficient data to derive an SSD as well as in SSDs that are "unstable" because of low data numbers. To enable decision-support applications of SSDs for as many compounds and uses as possible, by reconsidering the afore-mentioned criteria, we collated species sensitivity data from existing sources. We created a base set of ecotoxicity data for test situations that may occur in nature. We were able to derive SSDs for a large number of compounds and added a quality score to each SSD. We illustrate how SSDs and their quality scores are used in assessment planning and interpretation.

The aims of the present study were to describe the following: 1) data collection and curation for as many chemicals as feasible; 2) the derivation of chemical-specific SSDs, each with a quality score; 3) the application and evaluation of methods to bridge data gaps; 4) the utility and limitations of using SSDs in practical assessments (case study); and 5) the set of SSDs for further use (Supplemental Data, Table S2).

The scope for using the SSDs is global. Supplemental Data, Table S2, presents 2 operationally defined SSD models for the studied compounds, based on chronic no-effect or negligible-effect data (e.g., no-observed-effect concentration [NOEC], 10% effect concentration [EC10], etc.) and acute median effective concentration (e.g., EC50) data, respectively. The former SSDs relate to current global practices in chemical safety assessment and environmental quality assessment and man-agement, operating via protective benchmark concentrations (such as the predicted-no-effect concentration [PNEC] and environmental quality standards [European Commission 2003, 2011] or similar benchmarks for geographies outside the European Union). Exposures below protective benchmarks are considered to imply no or negligible impacts, and ecosystems are considered sufficiently protected at exposures below the benchmark. The latter SSDs relate to current global practices in life cycle impact assessment and other environmental assess-ments in which likely impacts of chemical pollution are quantified. That is commonly done in a comparative way, between products and environmental samples (see Supplemen-tal Data, Section 1). Increasing the number of compound-specific SSDs is relevant for all policy uses.

MATERIALS AND METHODS

Ecotoxicity data

Ecotoxicity data were collated from many sources, curated and operationally characterized for the 2 (chronic and acute)

TABLE 1: Criteria for operationally characterizing ecotoxicity data as "acute" (ECETOC 1993); longer exposure durations were classified as "chronic"

Species group	Acute test duration
Algae	12 h
Bacteria	12 h
Unicellular animals	12–24 h
Crustaceans	24–48 h
Fish	4–7 d
Mollusks, worms, etc.	2–7 d

SSDs aimed at the following. 1) A validated set of existing and well-referenced aquatic ecotoxicity database is described in De Zwart (2002). All available toxicity data were designated to represent chronic or acute toxicity criteria (Table 1): records with the endpoints NOEC, lowest-observed-effect concentra-tion, maximum acceptable toxicant concentration, EC0, EC5, EC10, and EC20 are marked as "chronic NOEC" when they have an appropriate taxon-dependent test duration (see Table 1) and population-relevant effect criterion (e.g., repro-duction, growth, population growth, and development, next to mortality and immobility); and records with a sublethal (EC) or lethal endpoint ranging from 30 to 70% are marked as "acute EC50" when they have an appropriate taxon-dependent test duration (see Table 1) and effect criterion (e.g., mortality and immobility). This data set comprised of 30 806 records (3445 substances, 1556 different taxa, 2513 chronic NOEC values, 28 293 acute EC50 values). As described by De Zwart (2002), this data set was comparatively checked for plausible toxicity estimates. Implausible outcomes were traced to the original reference for data misinterpretations, often attributable to errors in unit transformations, typing errors, and/or tests conducted under less optimal conditions. Erroneous entries were corrected when possible, and data were removed when original sources could not be checked (in this and later steps). 2) Further referenced data were added from an analysis and categorization of listed compounds of established and emerging concern under various national and international legislations (see Stichting Toegepast Onderzoek Waterbeheer 2016.). Curation was as in data set 1. In addition, acute NOEC values and chronic EC50 values were identified using the test duration criteria. Data were obtained from a variety of sources: the US Environmental Protection Agency's ECOTOX database (1995), comprising 58 714 records (1853 substances, 942 taxa, 15 019 acute EC50 values, 19 875 acute NOEC values, 21 676 chronic NOEC values, 2144 chronic EC50 values); a total of 952 test results from fish embryo toxicity tests on 214 substances with 4 different fish species adopted from the Procter & Gamble Company (Oris et al. 2012); Das et al. (2013) and Sanderson and Thomsen (2009) provided 334 records of measured acute EC50 toxicity concerning algae, daphnids, and fish for 162 different pharmaceuticals' active ingredients; a series of reports generated to define preliminary environmental quality criteria for compounds suspected to cause impact provided additional information mainly on pesticides and pharmaceuticals (Osté et al. 2010; Harezlak and Keijzers 2011;

Smit and Keijzers 2015), consisting of 1059 records on acute and chronic ecotoxicity for 37 substances involving 215 different taxa; a series of draft assessment reports from the European Food Safety Authority (2015) and the Pesticide Properties Database (University of Hertfordshire 2007); the Swiss Centre for Applied Ecotoxicology provided a large number of dossiers on the ecotoxicity of pharmaceuticals and pesticides (EAWAG 2016); and the WikiPharma database (MistraPharma 2015) was used to complement ecotoxicity data for listed pharmaceuticals. 3) Data in the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) registry (24 April 2015; European Commission 2006) after code harmonization added 207 943 records on acute and chronic toxicity on 8787 different substances, mainly for algae, daphnids, and fish; REACH data (European Commission 2006) are not properly documented regarding test conditions, test duration, and exposed taxa and are not transparently traceable to peer-reviewed origin; they were therefore analyzed both separately as well as after combining with the other data. 4) Read-across acute ecotoxicity data (baseline, acute median lethal concentration) for algae, daphnids, and fish were derived for 5201 substances. Toxicity estimates were evaluated as the lowest value derived by 2 different estimation methods: by ECOSAR prediction (Mayo-Bean et al. 2017) and by methods utilized by the Helmholtz-Zentrum für Umweltforschung (UFZ). The UFZ method for acute fish toxicity consists of an automated read-across approach (Schüürmann et al. 2011). This model estimates baseline toxicity from log octanol–water partition coefficients and corrects it by a toxicity enhancement derived from experimental data for similar compounds. Compound similarity is deducted by comparison of atom-centered fragments (Kühne et al. 2009). For daphnids, a refined version of this approach has been applied (Kühne et al. 2013). For algae, a simple model similar to that for fish was used, but the acute toxicity was directly derived from experimental data of similar chemicals taken from an internal data set. Again, these data were analyzed both separately as well as after combining with the other data.

SSD derivation

A variety of log-normal SSDs was derived. Available data were used to derive a compound-specific SSD for all species tested. The location and scale parameters of log-normal SSDs (mu and sigma) were first derived with optimum data, and the resulting SSDs were assigned a high quality score ("1111"). Because decreasing test data numbers trades off into limited numbers of compounds and potentially lower SSD "stability," additional SSDs were derived with various (inter-SSD) extrapolations to bridge data gaps, followed by an evaluation of similarity to the high-quality ones.

In line with the data origin and the use contexts (both chemical safety and environmental quality assessment), data were subdivided according to source quality (literature referenced vs nonreferenced and measured vs read-across) and to effect endpoints (chronic and acute) with characterization of data that were strictly measured ecotoxicity endpoints and obtained by extrapolation.

With the selection of strictly measured data, if more than 2 different taxa are tested for acute EC50 and chronic NOEC, the summarization process consisted of estimating the 2 moments of a log-normal SSD (Posthuma et al. 2002): 1) mu, the population median of toxicity values with equal weight per taxon, by first calculating the geometric average toxicity within taxa and subsequently calculating the geometric average toxicity over taxa of the geometric average toxicities per taxon, and 2) sigma, the population standard deviation of ^{10}log-transformed toxicity data, without considering taxon weight. If ecotoxicity data were available for less fewer 3 taxa, the process was restricted to mu—the average of the population of ^{10}log-transformed toxicity values with equal weight per taxon. For substances with too little data, an intermediate median SSD slope (sigma) was adopted with a value of 0.7 (the average slope over all data is 0.71).

The "strictly measured" data selection leaves many of the collated data unused. When sufficient acute EC50 or chronic NOEC values are not available but other toxicity endpoints are, those are extrapolated to acute EC50 and chronic NOEC values using empirically derived extrapolation values (De Zwart 2002; Duboudin et al. 2004). Note that the correct interpretation of this extrapolation is a parallel shift in an SSD, which is far more robust than per-species acute–chronic ratio extrapolations (De Zwart 2002). Further test endpoints on acute EC50 and chronic NOEC were first obtained after extrapolation from the most prominently available single other test endpoints in the collated data set. This was similar to the procedure for the strictly reported ecotoxicity data. Extrapolation factors are summarized in Table 2. If more than 4 substances shared the same primary mode of action in this extrapolation procedure, the SSD slope (sigma) was averaged over all substances with a similar primary mode of action. If more than 5 different taxa are tested for either acute EC50 or chronic NOEC, the input data were not further extrapolated. If test data are available in this extrapolation process for fewer than 3 taxa, the test endpoints (acute EC50 and chronic NOEC) were summarized by extrapolation from all available data, irrespective of the reported test endpoint. This

TABLE 2: Species sensitivity distribution extrapolation scheme by factor from/to

From/to	Order of extrapolation attempts to acute EC50[a]	Acute EC50 extrapolation factor[b]	Order of extrapolation attempts to chronic NOEC[a]	Chronic NOEC extrapolation factor[b]
Acute EC50	0	Multiply by 1	3	Multiply by 1/10
Acute NOEC	1	Multiply by 3	2	Multiply by 1/3
Chronic EC50	2	Multiply by 3	1	Multiply by 1/3
Chronic NOEC	3	Multiply by 10	0	Multiply by 1

[a]Numbers relate to quality scores in Table 4.
[b]Numbers express inter–species sensitivity distribution (SSD) extrapolations by parallel SSD shift.
EC50 = median effect concentration; NOEC = no-observed-effect concentration.

was done according to the scheme and the order presented in Table 2, as derived from De Zwart (2002) and approximately reconfirmed in the present study. If in the extrapolation process the maximum coverage of ecotoxicity data is available for fewer than 3 taxa, the summarization process is based on strictly measured data and restricted to mu—the average of the population of ^{10}log-transformed toxicity values with equal weight per taxon.

SSD types and quality scores

The database contains far more acute EC50 data than chronic NOEC data, making SSDs from the former data more robust for the majority of compounds. There are also other reasons that some SSDs are likely better estimates of true but unknown assemblage-wide sensitivity differences than others. Therefore, we added 4-digit quality scores as shown in Table 3, with the modality "1111" representing the highest quality. Derivation of the slope may or may not be possible (digit 1). Representation of different taxonomic groups was ranked (digit 2). Data origin was classified (digit 3). When SSDs were derived via data-bridging techniques, the method was scored (digit 4). The quality score information was added for planning and interpreting an assessment. This has a specific utility when large numbers of samples and compounds are assessed and where an assessment commonly involves prioritization of management efforts to "most impacted sites" and "most contributing compounds." That is, prioritization is straightforward if all SSDs used in an assessment are of high quality. If some SSDs are of low quality, this generally indicates a need for collecting additional hazard data. An exception may occur for assessments with limited resources, with clearly high-ranking cases and some low-ranking cases derived from (partially) lower-quality SSDs. Uncertainty analysis may reveal whether additional hazard data can "move"

TABLE 3: Four-digit quality scoring of species sensitivity distribution (SSD), based on underlying data types, quality, and numbers

Digit	Quality aspect	Modality	Meaning
1	SSD fullness	1	Data on full SSD available (mu and sigma)
1		2	Not sufficient data to calculate SSD slope
2	Biodiversity coverage	1	Number of taxa evaluated >10
2		2	Number of taxa evaluated >5
2		3	Number of taxa evaluated >2
2		4	Number of taxa evaluated <3
3	Data origin quality	1	Strictly measured
3		2	Extrapolated
3		3	Read-across
4	Extrapolation quality	1	Not extrapolated
4		2	Single-step extrapolation (Table 2)
4		3	Double-step extrapolation (Table 2)
4		4	Triple-step extrapolation (Table 2)
4		5	All available toxicity data extrapolation (Table 2)
4		6	Read-across

cases with a low rank and a low SSD quality "up" to the group prioritized for management attention.

Evaluation of SSDs

Various regressions of "other" SSDs on high-quality ones were performed to evaluate the quality of the "other" SSDs (for mu, the SSD midpoint). Similar outcomes suggest that data sets can be merged to derive SSDs from more data per compound. The regressions involved REACH data (European Commission 2006), read-across data, and extrapolated acute EC50 or chronic NOEC data (cf. Table 2). These were regressed against the geometrical average of high-quality SSDs, defined as strictly measured acute EC50 or chronic NOEC toxicity data over tested taxa derived with the curated and validated data for the substances. Also, the geometric average of strictly measured chronic NOEC toxicity data over tested taxa was regressed against the geometric average of strictly measured acute EC50 toxicity data.

Example case study: Water quality assessment and management prioritizations

Scope. The case study was set up to examine the largest possible fraction of chemicals in commerce in Europe, focusing on water quality and pollution impacts.

Exposure assessment. Predicted exposure concentrations (PECs) were derived from European Union production data with an integrated European Union-wide emission–fate–hydrological model (J. Van Gils et al., Deltares, Delft, The Netherlands, unpublished manuscript); details on PEC derivation and accuracy are in Supplemental Data, Section 4. Combined with the SSD results, required exposure and hazard data were available for a subset of 1760 compounds, selected for adequate physicochemical and ecotoxicological data representation. Predicted exposure concentrations are freely dissolved concentrations and were derived for 22 278 modeled European Union subcatchments (median spatial resolution 214 km^2) for a 365-d period (with weather data for the year 2013). This yielded 1.4×10^{10} PECs.

Impact assessments. Toxic pressures of the individual chemicals and their mixtures were derived from the PECs per subcatchment per day using the SSDs of the present study. Compared with the mixture model of De Zwart and Posthuma (2005), we applied a simplified approach to derive mixture toxic pressures, by operationally assuming that all chemicals act concentration additively,

$$\sum_{i=1}^{1760} HU_i, \text{ where } HU_i \text{ is } \frac{PEC_i}{SSD \text{ midpoint mu}_i}$$

with HU being hazard unit. This is similar to the frequent practice of water quality assessment for mixtures by linearly summed risk quotients over compounds, whereby numerical values estimated with the simplified and the original (more complex) model are similar. The approach was modeled with the Microsoft Excel

function NORMDIST (^{10}log[ΣHU],0,Average SSD slope $= 0.7$,1). Chronic and acute mixture toxic pressures were quantified per subcatchment per day.

Assessment and interpretation examples. The vast number of daily mixture toxic pressure data for individual subcatchments (8.1×10^6) were summarized using various approaches and statistics. Relative impact rankings across water bodies were visualized as Geographic Information Systems (GIS) maps, based on the years' 95th percentile concentration (P95) multisubstance potentially affected fraction (msPAF) values. Relative rankings of the contributions of chemicals to those impacts on a regional scale were derived in 2 steps. In the first step, a relative toxicity score for each compound in a subset of samples (Europe or a specific example river basin) was determined as the product of the mean compound toxic pressure in the set and the ratio between the non-zero values for that compound and the non-zero values for the mixtures. Those represent the magnitude and the relative contribution and frequency of increased compound pressures to total mixture pressures. All scores were then relatively ranked using the lowest-ranking compound as the baseline (defined as "1"). Case study example data are shown for all European basins or, for some assessments, for a typical northwestern (Rhine), a southeastern (Danube), and a set of southern basins (the Spanish basins of Ebro, Guadalquivir, Xuquer, and Llobregat combined).

RESULTS

Ecotoxicity data

The collated and curated data set for deriving SSDs consists of 256 409 records (details on the data set are in Supplemental Data, Section 2). Data origins (strictly measured and referenced up to read-across) were tagged, to allow derivation and comparisons of various types of SSDs for a compound.

SSD types and quality scores

A single compound may have various SSDs: from chronic or acute data, from referenced data, from unreferenced REACH data (European Commission 2006) or read-across data, and from strictly measured or extrapolated data. The SSDs could be derived for 12 386 compounds, where 12 214 originate from acute EC50 data and 7540 from chronic NOEC data. Their characteristics are summarized in Supplemental Data, Table S2, where the acute and chronic SSD data are selected to represent the lowest available quality score (best-quality SSD), while combining the referenced and unreferenced REACH data (European Commission 2006). The quality scores vary from 1111 to 2436 or 2411 for acute and chronic SSDs, respectively. Supplemental Data, Table S2, also contains a summary overview of the numbers of compounds per SSD type.

Evaluation of SSDs

The first evaluation considered regressions of acute SSDs derived from various methods ("other") on high-quality

nonextrapolated SSD midpoints. The regressions were all significant (Figure 1): SSDs based on acute EC50 REACH data (European Commission 2006; A), overlapping number of compounds 927, $y = 0.8173x + 0.7025$, $R^2 = 0.65$, $p < 0.001$; SSDs based on acute EC50 read-across data (B), overlapping number of compounds 1116, $y = 0.5914x + 1.2515$, $R^2 = 0.49$, $p < 0.001$; SSDs based on acute EC50 data extrapolated from other test endpoints (C), overlapping number of compounds 3827, $y = 0.9611x + 0.1362$, $R^2 = 0.95$, $p < 0.001$. The read-across SSDs (related to baseline toxicity) showed more variability and appeared to underestimate the mu for the most toxic substances up to a factor of 10 compared with high-quality SSDs. The reliability of acute EC50 SSDs based on the other data analysis methods decreases in the sequence extrapolated SSDs $\approx$ REACH-based SSDs > read-across SSDs.

The second evaluation similarly considered 2 regressions, now for SSDs from chronic NOEC data (Figure 2). For chronic NOEC REACH data (European Commission 2006), the relationship with the chronic high-quality nonextrapolated SSD midpoint data was less strong than for the EC50 data but still highly significant (A; overlapping number of compounds 251, $y = 0.7448x + 0.8502$, $R^2 = 0.60$, $p < 0.001$). For the most toxic substances, the REACH data (European Commission 2006) tend to underestimate toxicity compared with the well-referenced data set by a factor of up to approximately 30. A highly significant relationship was found between extrapolated chronic data and the chronic high-quality nonextrapolated SSD midpoint data (B; overlapping number of compounds 1131, $y = 0.7941x + 0.5902$, $R^2 = 0.74$, $p < 0.001$). The extrapolated data also tend to underestimate toxicity compared with the high-quality data by a factor of up to approximately 10. The reliability of chronic-NOEC SSDs based on the other data analysis was relatively similar for both other methods.

The third evaluation considered the comparison of SSDs based on NOECs versus EC50s. This showed a significant association between the mu values for the 250 substances for
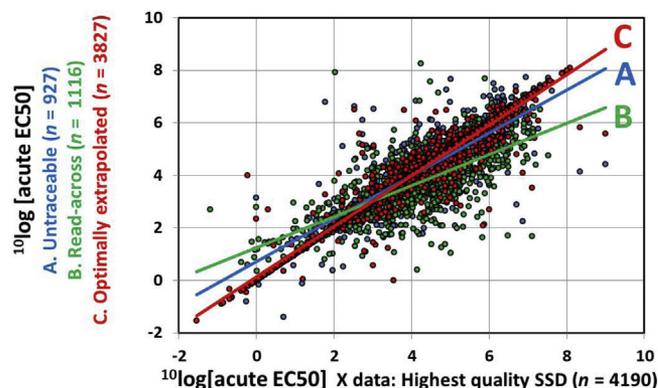

FIGURE 1: Comparison of acute median effect concentration (EC50) species sensitivity distribution (SSD) midpoint data for 3 SSD derivation approaches regressed on highest-quality SSD midpoint data. The x-axis shows highest-quality SSDs (traceable data, all taxa, geometrically averaged, measured data, acute EC50). The y-axis shows "other" and "derived" SSDs (A = untraceable data [REACH], blue; B = read-across, green; C = optimally extrapolated, red).
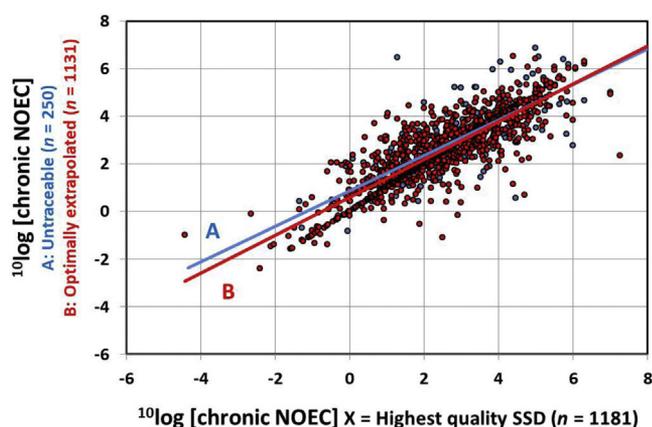
FIGURE 2: Comparison of chronic no-observed-effect concentration (NOEC) species sensitivity distribution (SSD) midpoint data for 2 SSD derivation approaches regressed on highest-quality midpoint data. The x-axis shows highest-quality SSDs (traceable data, all taxa, geometrically averaged, measured data, chronic NOEC). The y-axis shows other SSDs (A = untraceable data [REACH], B = optimally extrapolated, red).

which both aspects are strictly measured and quantified (Figure 3; $n = 250$, $y = 0.7304x - 0.4414$, $R^2 = 0.60$, $p < 0.001$). On average, the chronic NOEC is less than a factor of 10 lower than the acute EC50, which represents an observed factorial shift of the SSD. A limited number of outliers have a strong influence on this regression. A restriction of the regression to the 5th to 95th percentile data intervals yielded a factorial difference of approximately 6.6 and a slope of approximately 0.85 ($n = 225$, $y = 0.8509x - 0.8237$, $R^2 = 0.76$, $p < 0.001$). These outcomes suggest that, on average, an SSD NOEC can be derived from an SSD EC50 by extrapolation (as shown in Table 2) but also that the variability around the average should be taken into account when such an extrapolation would be used in practice.
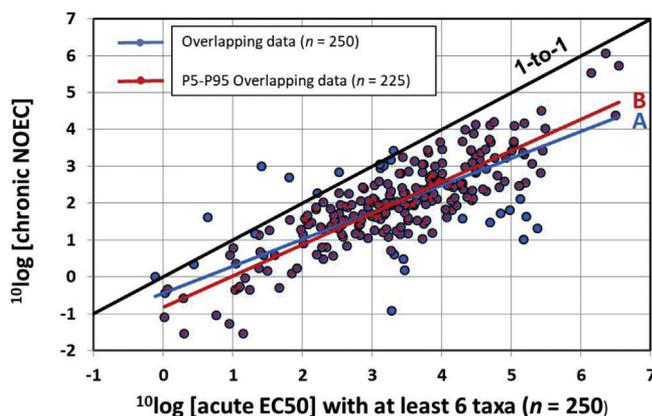

FIGURE 3: Comparison of species sensitivity distribution midpoints for acute median effect concentration (x) and chronic no-observed-effect concentration (y). A = All data points, blue dots and blue line; B = all data points, restricted to the P5-P95 data ranges, red dots and red line. EC50 = median effect concentration; NOEC = no-observed-effect concentration; P5–P95 = 5th to 95th percentile concentration.

Example case study: Water quality assessment and management prioritizations

Scope and illustrative purpose. Presented results illustrate how the method allows for impact rankings of sites and of relative importance of substances within basins or water bodies. The impact assessments are based on the large series of exposure concentrations (detailed in Supplemental Data, Section 4). It should be noted that alternative data summary choices—related to the assessment problem—yield different results. Results are therefore explicitly not to be interpreted as a list of Europe-wide priority sites or priority chemicals in a Water Framework Directive (European Commission 2000) context.

Impact assessment and prioritizations. Various impact assessment and prioritization outcomes are illustrated, based on the large set of PECs for 1760 substances that all have relatively high SSD quality scores, from 1111 to 1324 or 1325 for chronic and acute, respectively. Mixture toxic pressures based on chronic or acute SSDs were derived and mapped, to illustrate the spatial variation of these impact-related metrics (Figures 4 and 5, respectively). because GIS maps cannot plot both space and time, the example figures are based on the 95th percentiles of predicted mixture toxic pressures over a year (P95-year = 18 d). The mixture toxic pressure of a water body is higher for 5% of days in the modeled year. The interpretation proceeds as follows. 1) Per-species interpretation: Colors represent the variation of the probability that a randomly selected species from the set of tested species would be exposed beyond the species' chronic no-effect level or the acute EC50 for at most 18 d per year. These outcomes represent the "probability of effects on a species" (PES) of the water pollution at the level of "initiation" or "substantial harm," respectively (Suter et al. 2002). A PES value characterizes the potential of the polluted water to cause harm, which is the basis for the term "toxic pressure." 2) Biodiversity interpretation: A quantitatively identical expression of the outcome is the potentially affected fraction of species (PAF), but the interpretation narrative differs. The PAF expression relates to the regulation-defined endpoint of protecting against biodiversity reduction in species assemblages. A PAF value relates to the concept of protective benchmarks, if those are derived from an SSD of chronic NOECs, such as the hazardous concentration for 5% of the species (HC5). Higher toxic pressures imply higher fractions of species affected for the studied acute or chronic endpoint. 3) Regulatory interpretation: The 2 maps relate to 2 current policy approaches: chemical safety and environmental quality assessment policies and ecological impact assessment. The 2 approaches operate via the protective benchmark no-effect concept (using PNEC and environmental quality standards for REACH [European Commission 2006] and the Water Framework Directive [European Commission 2000], respectively), below which ecosystems are considered "sufficiently protected" for expected or observed exposures. For single compounds or mixtures, sufficient protection relates to PAF-NOEC$_{max}$ = msPAF-NOEC$_{max}$ = 0.05, which is in regulatory terms considered to protect 95% of the species against adverse effects. In
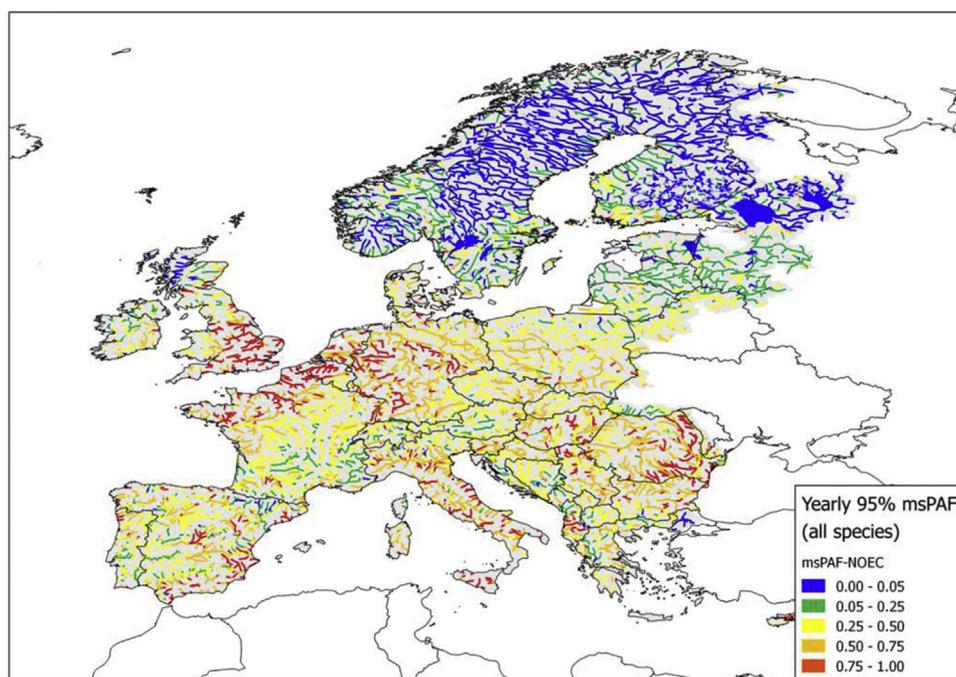
FIGURE 4: Results mapped in relation to the regulatory concept of "sufficient protection" of aquatic ecosystems (initial effects, distress), used in chemical safety assessment (e.g., Registration, Evaluation, Authorisation and Restriction of Chemicals, predicted-no-effect concentration) and water quality assessment of chemicals according to the Water Framework Directive (environmental quality standard). Blue: species assemblages in the water bodies are "sufficiently protected" for 95% of the days (multisubstance potentially affected fraction [msPAF]–no-observed-effect concentration [NOEC] <0.05, 95th percentile concentration [P95] of a year = 18 d; see text). Technical output specification: numbers are P95-year of msPAF–chronic NOEC per site ($n = 22\,728$ water bodies).
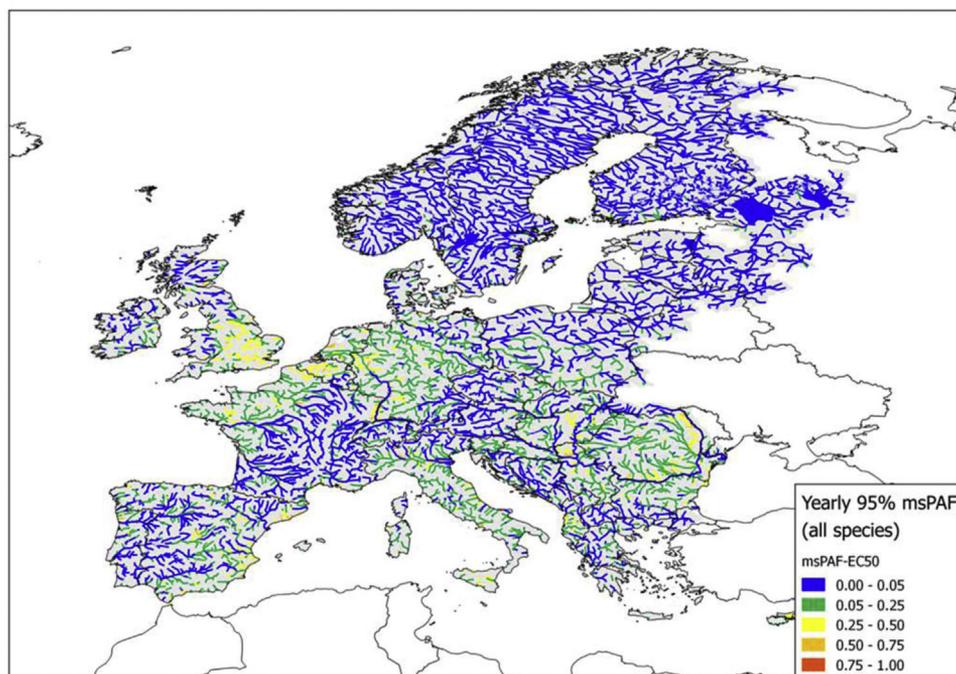

FIGURE 5: Results mapped in relation to the regulatory concept of ecological impact magnitudes (species loss), as utilized to classify ecological impacts under the Water Framework Directive. Higher multisubstance potentially affected fraction (msPAF)–acute median effective concentration (EC50) values empirically relate to increased species loss. Technical output specification: numbers are 95th percentile concentration of a year msPAF–acute EC50 per site ($n = 22\,728$ water bodies).

Figure 4, this equivalency was used as a class boundary between sufficient and insufficient protection (blue–green boundary), with other colors represent increasing distress (exposure higher than the no-effect level). In Figure 5, the color scale relates to increased biodiversity effects, found in empirical studies, which can be aligned with ecological impact classification used in the Water Framework Directive (European Commission 2000) to define moderate, poor, and bad water quality.

Impact distributions across water bodies were investigated with the same data. The water quality would be classified as "insufficiently protected" for approximately 65% of all European water bodies (P95-year msPAF-NOEC data; Figure 6). Specific basins have higher fractions of water bodies with insufficient protection, with average values of approximately 93, 88, and 79% for the example basins of the Rhine, the Danube, and the Spanish basins, respectively. The observed Pareto-type (skewed) distributions imply that relatively few sites are characterized by relatively high chronic toxic pressures.

Relative impact contributions of chemicals were investigated, with the same data (example for acute SSD EC50 ranking) and toxic pressure data aggregated over an area and over time. We derived a top-15 ranking of substances (Table 4). The top 15 explained nearly 99.5% of the mixture exposure effects, with <0.5% explained by the remaining 1745 compounds. This Pareto-type distribution implies that approximately 1% of the compounds cause 99% of the exposure impacts, leading to a Pareto-type "99–1" rule for species loss for the P95-year assessment.

Some of the top 15 chemicals were not identified as potentially relevant following current compound prioritizations according to the Water Framework Directive (European Commission 2000) or the NORMAN network (kindly provided by Valeria Dulio, executive secretary of NORMAN Association 2016). Modeling expected impacts may identify chemicals that likely affect ecosystems but that are not identified by monitoring (because of lack of attention, lack of analysis methods, or detection limits that are higher than effect benchmarks).

Final interpretation of an assessment. In a comprehensive assessment, these types of results would be further checked before being used for management prioritization. First, ranking outcomes should be checked on their fit to the assessment problem. For example, prioritizations will differ when chemicals with peak exposures (such as pesticides) are involved, and the assessor may then evaluate outcomes from P99-year-based ranking. The outcomes of this are illustrated in Supplemental Data, Section 5. Second, SSD-quality scores should be evaluated for their potential influence on the interpretation (quality scores were high for the top 15 chemicals). Third, outcomes will normally be interpreted with other lines of evidence. In doing so, collection of data on the top 15 chemicals showed that 2 compounds, terbufos and phorate, are no longer approved in the European Union. Because all of our assessment runs were based on dossier data which we did not a priori screen on forbidden compounds so that the high ranks are not relevant for current prioritizations; however, the past nonapproval decision is supported by the high ranks found for these compounds. Further compound details are in Supplemental Data, Section 6, showing that all of the top 15 compounds are characterized by high production mass, ubiquitous use, and high hazard classifications.
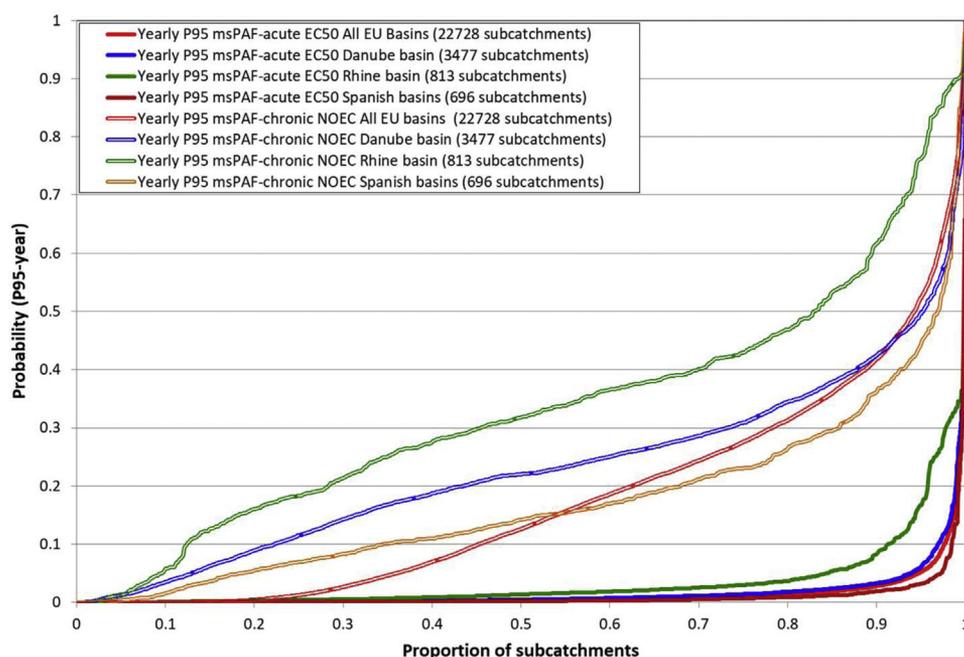

FIGURE 6: Cumulative distributions of mixture toxic pressures of the individual subcatchments of Europe (*n* = 22 728) and of 3 example basin areas (Rhine, *n* = 813; Danube, *n* = 3477; Spanish basins, *n* = 696). *x*-axis = rank order of subcatchments; *y*-axis = probability that exposures exceed the no-effect or 50%-effect level, under the assumption of concentration additivity of the substances in the mixtures. EC50 = median effective concentration; msPAF = multisubstance potentially affected fraction; NOEC = no-observed-effect concentration; P95 = 95th percentile concentration.

TABLE 4: Example of chemical ranking[a]

Substance	CAS	WFD priority	Norman priority	All 22 EU basins	Danube basin	Rhine basin	Spanish basins	Count	Rank
Bisphenol-A	80-05-7	√	√	90 316	85 935	277 239	110 079	4	1
N-1,3-Dimethylbutyl-N'-phenyl-p-phenylenediamine	793-24-8			2573	2451	20 344	3528	4	2
Chlorpyrifos	2921-88-2	√		1629	2755		78 285	3	3
Anthracene	120-12-7	√		1502	6675	3528	247	4	4
Octamethylcyclotetrasiloxane	556-67-2		√	1483	6325	1320	92	4	5
N-(4-Aminophenyl)aniline	101-54-2			1381	268	15 144	393	4	6
Cumene hydroperoxide	80-15-9			1123	4332	2242	634	4	7
Difenylamine	122-39-4		√	589	308	8667	1028	4	8
1-Dodecanol	112-53-8			48	558			2	9
Pyraclostrobin	175013-18-0			41	151	114		3	10
Cyhexatin	13121-70-5			25	8		2821	3	11
p-Phenylenediamine	106-50-3			19	17	97		3	12
Dimoxystrobin	149961-52-4			11	46	37		3	13
Terbufos	13071-79-9			6	75		96	3	14
Phorate	298-02-2			1			110	2	15
Count		3	3	15	14	10	11		
Sum relative score of 15 substances				100 748	109 904	328 733	197 314		
Sum relative score of remaining 1760 substances				208	571	0	0		
Percent of sum relative score in top 15 substances				99.8%	99.5%	100.0%	100.0%		

[a]An illustration of the top 15 ranked toxicants, their species sensitivity distribution quality scores, and their relative impact potential to species loss. Outcomes based on 95th percentile concentration (P95)–multisubstance potentially affected fraction–median effect concentration, for Europe and for 3 example case study basins (relative ranks defined by phorate for whole-Europe data defined as baseline, "1"). Priority marks: compound listed as priority compounds for management attention following WFD or NORMAN methods. Note that the choice for P95 excludes peak exposures of, for example, pesticides (see Supplemental Data, Section 6).
CAS = Chemical Abstracts Service; EU = European Union; WFD = Water Framework Directive.

DISCUSSION

Key innovations

The present study addresses the problem that different environmental policy frameworks have developed very different practices in handling ecotoxicity data and assessment models, with major trade-offs for practical assessments. Various guidance documents prescribe strict criteria for data selection and SSD derivation. The downside of such criteria is that the risks of most compounds and their mixtures cannot be evaluated, prioritized, and managed. Current water quality assessment under the Water Framework Directive (European Commission 2000) considers, for example, only 0.2% of the compounds in commerce (environmental quality standards for approximately 300 compounds [European Environment Agency 2018] vs >146 000 registered compounds on the REACH website). To address that trade-off, we developed SSDs and associated mixture approaches for a large number of chemicals, to enable more comprehensive and realistic assessments.

Innovative aspects of the present study are 1) the methods and sources for the collection and curation of ecotoxicity data, 2) the SSD derivation and quality scoring method, 3) the extra information that can be gained from inter-SSD comparisons and extrapolations, 4) the utility of SSD-based assessment outcomes for ranking sites and compounds (illustrated in the case study), and eventually 5) the opportunity to use a consistent set of ecotoxicity data and SSDs for various practical policies. For the practical uses, assessors should be aware of the limitations of SSD-based assessments because SSDs do not consider food-chain exposures or indirect effects (e.g., on predators via toxicity

to prey). Moreover, they should be aware of the fact that the simple expression of the toxic pressure for a water body has the complex interpretation that the species that are exposed will show widely different impacts, related to the species sensitivity differences that are the basis of SSDs.

Expanded number of compounds, SSDs, quality scores, and SSD applications

We operationally derived separate acute and chronic SSDs for many compounds. Those may—in principle—be used for all policy purposes. Checking SSD-quality scores is thereby always important. Low-quality scores (e.g., caused by few ecotoxicity data or by extrapolation) may be consequential. For example, the calculated acute-median sensitivity (mu-acute) may be smaller than the calculated chronic value (mu-chronic) because of the haphazard effect of small data sets. This specific effect occurred for approximately 1.3% of the SSDs (160 data lines in Supplemental Data, Table S2). Low SSD-quality scores can only be improved by collecting more test data. When quality scores are considered sufficient, the way to use the SSDs for practical assessments is basically as follows.

First, we acknowledge that current guidelines exist on deriving protective benchmark concentrations, and the results of the present study may therefore not be adopted for specific policies. However, if there are no data that fit the current guidance for a contaminant of potential concern, a provisional benchmark concentration can be derived, for example, via the chronic-acute SSD-level relationship shown in Figure 3, to

914

Environmental Toxicology and Chemistry, 2019;38:905–917—L. Posthuma et al.

provide provisional insights into potential impacts for data-poor chemicals. Second, the utility for water quality assessment and management prioritization was shown in the case study, based on relative rankings of impact levels across water bodies and compounds within the studied basins. Third, the SSDs can be used for establishing the environmental (ecotox-icity) footprints of products as well as production-chain ecotoxicity hot spots, to enable production and selection of environmentally benign products. The European Union recently derived SSD-based effect factors for this purpose from 3 regulatory data sources for thousands of chemicals (Saouter et al. 2018). Consistent environmental protection and pollution management can be based on the SSDs presented in the present study, for a wide array of compounds. Critical use must be supported by the quality scores. If needed, specific SSD types with specific quality scores may be earmarked for specific purposes by the user, for example, for repetitive assessments.

Prioritization opportunities

All assessment outcomes of SSDs relate to ranking as the basis for management prioritization and efficacy. For an array of compounds, the rank order of protective benchmarks (e.g., HC5s, PNECs, environmental quality standards) reflects the relative potency of different compounds to cause harm. For an array of sites, the order of (mixture) toxic pressures similarly reflects impact differences across sites. For an array of products, outcomes identify benign products and produc-tion-chain hotspots. Outcomes of acute-data SSDs have even been used in disaster assessment and management by United Nations Disaster Assessment and Coordination teams (Sup-plemental Data, Section 1). The case study only illustrates the ranking of polluted water bodies. Based on experiences with case study data (such as in Table 4), we recommend that rankings for all applications are not interpreted as absolute and as fixed order of cases because natural and human-made variabilities of exposure occur. That is, pesticides may have a high rank order only during the growing season, with further influences of weather (e.g., rainfall events) affecting emissions of some chemicals (via runoff) and dilution of all chemicals. Instead of assuming an absolute idea of site or substance ranking, we propose the use of classes. For example, 1) an "always high"-ranking class of sites or compounds, where the probability of impact is always high; 2) an "always low"-ranking class, where the low probability of impact allows one to "exclude the innocent"; and 3) an "intermediate" class, where the probability of impact depends on the situation. These classes would discriminate 3 clear management perspectives: 1) action needed, 2) no action needed, and 3) possible further lines of evidence needed. The concept of using classes is further supported by results from field monitoring (Vallotton and Price 2016). Based on this phenomenon, assessors can also consider the opposite of prioritizing the higher-impact sites or compounds for management, by considering the lower tail of the distribution (Figure 6). With a large number of ranked cases and limited management resources, there may be an

opportunity to "exclude the innocent," even when some SSD-quality scores underlying the lower-tail ranking are low.

Case study: Utility example on ranking toxic pressures for sites and substances

The case study illustrates how site and substance ranking provides management prioritization insights, refining global water quality classification practices. The latter are currently based on comparing single-chemical exposure concentrations to a protective benchmark. This results in a binary outcome of water quality assessment: there is "(in)sufficient protection," which is then interpreted and communicated as "polluted" or "unpolluted." When applied to mixtures, the application of such an approach to European surface waters showed that all water bodies of a country can be interpreted as polluted (see, e.g., European Environment Agency 2012). This incorrectly seems to suggest equal management needs for all studied water bodies, neglecting that low and high benchmark exceedances imply lower and higher impacts and lower or higher motives for pollution reduction, respectively. The case study shows that current practices can be refined, to highlight where impacts are likely highest, according to compound(s)/groups. These rank-ings help prioritization of management and select measures that reduce impacts (improve water quality) most. Novel case study insights are as follows. First, European water quality is currently insufficiently protected (Figure 4), corroborating the study of Malaj et al. (2014). Second, this implies an associated degree of likely species loss (Figure 5), based on empirical evidence for the association between msPAF-acute EC50 and species loss (e.g., Posthuma et al. 2016). Third, this is attributable to relatively few compounds at the European and basin scales (Table 4), also found by others (e.g., Vallotton and Price 2016). Fourth, the high-ranking compounds share specific characteristics: high mass used in Europe and ubiquitous use and high hazard character-istics for aquatic ecosystems (see Supplemental Data, Section 6). This means that chemical safety assessment knowledge (as collected for, e.g., REACH [European Commission 2006]) provides meaningful indications of potential impacts in aquatic ecosystems. All of these insights were obtained with compounds with high SSD-quality scores. The sequence of analysis steps was designed in line with the holistic principles of the Water Framework Directive (European Commission 2000). It implies a stepwise and meaningful "system–site–substance–solution" focus in the assessment of impacts and planning of management of European surface waters.

Assessment problem definition, model choice, and interpretation

The case study resulted in outcomes for a specific set of conditions (e.g., year P95 data). In practical assessments, the most informative outcomes for management prioritization should be generated by tailoring the assessment to the specific conditions. For example, if the emissions of all chemicals in a region are rather constant (e.g., household chemicals), the assessor may

investigate especially spatial exposure and impact ranking using, for example, the year-P50 or P95 toxic pressures, whereas for exposures in an agricultural landscape the assessor may focus on peak exposures (e.g., year P99 data). Supplemental Data, Section 5, illustrates the increased toxic pressures when using, for example, the P99-year data (4-d peak exposures are covered). The number of sufficiently protected sites (chronic mixture toxic pressure <0.05) is lower, and predicted species loss is higher. It is important to note that the applied European Union–wide model operates with a spatial resolution of approximately 200 km^2, and consequently, probabilities of impacts at point sources (e.g., downstream wastewater-treatment plants) are not shown in detail. In the vicinity of point sources, impacts may be much higher than shown in the present study via water body–level exposure data.

Combining lines of evidence and planning of monitoring

An assessment of the likelihood of impacts under the Water Framework Directive (European Commission 2000) consists of various lines of evidence. Model results can be combined with information from an "assessment of pressures" (human activities), available monitoring data, and other information. Information may consider sources as diverse as nontarget screening of the presence of chemicals (Hollender et al. 2017) to effect-based methods that assess impacts of complex mixtures in water samples (Altenburger et al. 2015). Planning and management of river basins combine these lines of evidence, and the use of SSDs in this process allows the assessor to obtain highly specific information on the likelihood that chemical pollution causes harm.

The case study results also contain a suggestion on monitoring and management for 10 additional compounds compared with current practices (Table 4).

Communicating results and evaluating trends

Currently, one of the conundrums of pollution assessment and management is the communication of results. In the case study, 1.5×10^{10} exposure concentrations had to be summarized for management planning, and this number multiplies if an assessor wants to evaluate trends of past management or of optional abatement strategies. To address this problem, mixture toxic pressures can be summarized as chemical footprints for a region (Bjørn et al. 2014; Zijp et al. 2014). The changes caused by past management or future abatement scenarios can then be summarized and communicated via chemical footprints, to summarize spatial or temporal trends in mixture impacts for large regions. Thus, SSDs can be used as an effective, though lower-tier (screening), approach for water quality assessment and management in the context of a wide diversity of policy fields, with opportunities to explore "big patterns" (footprints) as well as "details" (specific sites and chemicals within sites). The use of SSDs provides an intermediary tool that lies between generic assessments of chemical safety and more specific impact assessments based on more complex (ecological) modeling.

CONCLUSIONS

The following conclusions were reached. 1) Species sensitivity distributions are used in environmental protection, assessment, and management practices, currently for a few to a few thousand compounds only. 2) Species sensitivity distributions are provided for 12 386 substances, with a quality score to assist in planning and interpretation of assessments. 3) The utility of the SSDs was illustrated for water quality assessment at the European scale considering 1760 compounds and their mixtures. 4) The role of chemical pollution in aquatic ecosystems could be specified, regarding both the regulatory concept of sufficient protection (REACH [European Commission 2006] and Water Framework Directive [European Commission 2000]) as well as species loss (Water Framework Directive [European Commission 2000], ecological status impact classification). 5) The use of models is suggested in the European Water Framework Directive (European Commission 2000), and SSDs are fit for that use because they help to express expected impact magnitudes of pollution. 6) The use of SSDs for water quality assessment follows the holistic principles of the Water Framework Directive (European Commission 2000) and supports a "system–site–substance–solution" sequence in the assessment of impacts and planning of management. 7) The use of SSDs complements the current per-chemical benchmark approach, which substantially improves the diagnosis and communication of chemical pollution in surface waters.

Data Accessibility—Data, associated metadata, and calculation tools are available from the corresponding author (leo.posthuma@rivm.nl) and, the Excel Supplementary Data Table file is also deposited in FigShare.

REFERENCES

Altenburger R, Ait-Aissa S, Antczak P, Backhaus T, Barceló D, Seiler TB, Brion F, Busch W, Chipman K, de Alda ML, De Aragão Umbuzeiro G, Escher BI, Falciani F, Faust M, Focks A, Hilscherova K, Hollender J, Hollert H, Jäger F, Jahnke A, Kortenkamp A, Krauss M, Lemkine GF, Munthe J, Neumann S, Schymanski EL, Scrimshaw M, Segner H, Slobodnik J, Smedes F, Kughathas S, Teodorovic I, Tindall AJ, Tollefsen KE, Walz KH, Williams TD, Van den Brink PJ, van Gils J, Vrana B, Zhang X, Brack W. 2015. Future water quality monitoring—Adapting tools to deal with mixtures of pollutants in water resource management. *Sci Total Environ* 512–513:540–551.

Bernhardt ES, Rosi EJ, Gessner MO. 2017. Synthetic chemicals as agents of global change. *Front Ecol Environ* 15:84–90.

Bjørn A, Diamond M, Birkved M, Hauschild MZ. 2014. Chemical footprint method for improved communication of freshwater ecotoxicity impacts in the context of ecological limits. *Environ Sci Technol* 48:13253–13262.

Das RN, Sanderson H, Mwambo AE, Roy K. 2013. Preliminary studies on model development for rodent toxicity and its interspecies correlation with aquatic toxicities of pharmaceuticals. *Bull Environ Contam Toxicol* 90:375–381.

De Zwart D. 2002. Observed regularities in species sensitivity distributions. In Posthuma L, Suter GW, Traas TP, eds, *Species Sensitivity Distributions in Ecotoxicology*. CRC, Boca Raton, FL, USA, pp 133–154.

De Zwart D, Posthuma L. 2005. Complex mixture toxicity for single and multiple species: Proposed methodologies. *Environ Toxicol Chem* 24:2665–2676.

Duboudin C, Ciffroy P, Magaud H. 2004. Acute-to-chronic species sensitivity distribution extrapolation. *Environ Toxicol Chem* 23:1774–1785.

ECETOC. 1993. Aquatic toxicity data evaluation. Brussels, Belgium: European Centre for Ecotoxicology and Toxicology of Chemicals. Ecetoc Technical Report No. 56. Brussels, Belgium.

EAWAG. 2016. Proposals for acute and chronic quality standards. Lausanne, Switzerland. [2016 September 7]. Available from: http://www.ecotoxcentre.ch/expert-service/quality-standards/proposals-for-acute-and-chronic-quality-standards/.

European Chemicals Agency. 2008. Guidance on information requirements and chemical safety assessment. Helsinki, Finland.

European Chemicals Agency. 2016. Information on chemicals. Helsinki, Finland. [cited 2018, March 6]. Available from: http://echa.europa.eu/information-on-chemicals

European Commission. 2000. Directive 2000/60/EC of the European Parliament and of the Council establishing a framework for the community action in the field of water policy. *Official J Eur Union* L327:01–73.

European Commission. 2003. Technical guidance document on risk assessment, part II. Eur 20418 EN/2. European Chemicals Bureau, Joint Research Centre, Ispra, Italy. [cited 2018 March 6]. Available from: https://echa.Europa.Eu/documents/10162/16960216/tgdpart2_2ed_en.pdf

European Commission. 2006. Regulation (EC) 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the registration, evaluation, authorisation and restriction of chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) 793/93 and Commission Regulation (EC) No. 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. *Official J Eur Union* L396:374–375.

European Commission. 2011. Common implementation strategy for the Water Framework Directive (2000/60/ec). Guidance document 27. Technical Guidance for Deriving Environmental Qualilty Standards. Brussels, Belgium.

European Environment Agency. 2012. European waters—Assessment of status and pressures. 8/2012. Copenhagen, Denmark.

European Environment Agency. 2018. European waters—Assessment of status and pressures 2018. 7/2018. Copenhagen, Denmark.

European Food Safety Authority. 2015. Rapporteur member state assessment reports submitted for the EU peer review of active substances used in plant production products (published before March 2015). Parma, Italy. [cited 2017 December 14]. Available from: http://dar.efsa.europa.eu/dar-web/provision

Harezlak V, Keijzers R. 2011. Afleiding van het ecotoxicologisch deel voor 29 ad hoc mtr's voor 2010. 1203121-001. Deltares, Delft, The Netherlands (in Dutch).

Hoekstra AY, Wiedmann TO. 2014. Humanity's unsustainable environmental footprint. *Science* 344:1114–1117.

Hollender J, Schymanski EL, Singer HP, Ferguson PL. 2017. Nontarget screening with high resolution mass spectrometry in the environment: Ready to go? *Environ Sci Technol* 51:11505–11512.

Klimisch HJ, Andreae M, Tillmann U. 1997. A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data. *Regul Toxicol Pharmacol* 25:1–5.

Kühne R, Ebert R-U., Schüürmann G. 2009. Chemical domain of Qsar models from atom-centered fragments. *J Chem Inf Model* 49:2660–2669.

Kühne R, Ebert R-U, von der Ohe PC, Ulrich N, Brack W, Schüürmann G. 2013. Read-across prediction of the acute toxicity of organic compounds toward the water flea *Daphnia magna*. *Mol Inform* 32:108–120.

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 Sci USA* 111:9549–9554.

Mayo-Bean K, Moran-Bruce K, Meylan W, Ranslow P, Lock M, Nabholz V, Von Runnen J, Cassidy L, Tunkel J. 2017. Methodology document for the ecological structure–activity relationship model (ECOSAR) class program, Ver 2. Office of Pollution Prevention and Toxics, US Environmental Protection Agency, Washington DC.

MistraPharma. 2015. Welcome to the wikipharma databse. The Swedish Foundation for Strategic Environmental Research, Stockholm, Sweden. [cited 2016 November 18]. Available from: http://www.wikipharma.org/welcome.asp

Moermond CTA, Kase R, Korkaric M, Ågerstrand M. 2016. CRED: Criteria for reporting and evaluating ecotoxicity data. *Environ Toxicol Chem* 35:1297–1309.

Norman Association. 2013. NORMAN–EMPODAT database. Network of reference laboratories, research centres and related organisations for monitoring of emerging environmental substances. Environmental Institute, Kos, Slovak Republic. [cited 2018 October 10]. Available from: http://www.normandata.eu/empodat/.

Norman Association. 2016. Norman priority list. Environmental Institute, Kos, Slovak Republic. [cited 2018 October 10]. Available from: https://www.norman-network.com/sites/default/files/files/suspectListExchange/NORMAN_PriorityList_2016.csv

Oris JT, Belanger SE, Bailer J. 2012. Baseline characteristica and statistical implications for the OECD 210 fish early-life stage chronic toxicity test. *Environ Toxicol Chem* 31:370–376.

Osté L, Keijzers R, Mastalerz V. 2010. Afleiding van het ecotoxicologisch deel voor 28 ad hoc mtr's voor 2009. 1203121-001. Deltares, Delft, The Netherlands (in Dutch).

Posthuma L, De Zwart D. 2012. Predicted mixture toxic pressure relates to observed fraction of benthic macrofauna species impacted by contaminant mixtures. *Environ Toxicol Chem* 31:2175–2188.

Posthuma L, De Zwart D, Keijzers R, Postma J. 2016. Ecologische sleutelfactor toxiciteit. Deel 2: Kalibratie: Toxische druk en ecologische effecten op macrofauna. Dutch Foundation for Applied Water Research (STOWA), Amersfoort, The Netherlands (in Dutch).

Posthuma L, Suter GWI, Traas TP. 2002. *Species Sensitivity Distributions in Ecotoxicology*. CRC, Boca Raton, FL, USA.

Sanderson H, Thomsen M. 2009. Comparative analysis of pharmaceuticals versus industrial chemicals acute aquatic toxicity classification according to the United Nations classification system for chemicals. Assessment of the (Q)SAR predictability of pharmaceuticals acute aquatic toxicity and their predominant acute toxic mode-of-action. *Toxicol Lett* 187:84–93.

Saouter E, Biganzoli F, Ceriani L, Versteeg D, Crenna E, Zampori L, Sala S, Pant R. 2018. Environmental footprint: Update of life cycle impact assessment methods—Ecotoxicity freshwater, human toxicity cancer, and non-cancer. European Union, Luxembourg.

Schäfer RB, Kühn B, Malaj E, König A, Gergs R. 2016. Contribution of organic toxicants to multiple stress in river ecosystems. *Freshw Biol* 61:2116–2128.

Schüürmann G, Ebert R-U, Kühne R. 2011. Quantitative read-across for predicting the acute fish toxicity of organic compounds. *Environ Sci Technol* 45:4616–4622.

Smit CE, Keijzers R. 2015. Indicatieve waterkwaliteitsnormen voor bestrijdingsmiddelen—normvoorstellen voor 19 stoffen. Briefrapport 2015-0124. National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands (in Dutch).

Stichting Toegepast Onderzoek Waterbeheer. 2016. Ecological key factor toxicity. Main report, partial reports and calculation tools. STOWA, Amersfoort, The Netherlands. [cited 2018 September 21]. Available from: https://www.stowa.nl/publicaties/ecologische-sleutelfactor-toxiciteit-hoofdrapport-deelrapporten-en-rekentools

Suter GW, Traas TP, Posthuma L. 2002. Issues and practices in the derivation and use of species sensitivity distributions. In Posthuma L, Suter GWI, Traas TP, eds, *Species Sensitivity Distributions in Ecotoxicology*. CRC, Boca Raton, FL, USA, pp 437–474.

United Nations Environment Programme. 2013. Global chemicals outlook. Towards sound management of chemicals. Geneva, Switzerland.

University of Hertfordshire. 2007. PPDB: Pesticide properties database. Agriculture and Environmental Research Unit, Hatfield, Hertfordshire, UK. [cited 2016 September 22]. Available from: http://sitem.herts.ac.uk/aeru/ppdb/en/search.htm

US Environmental Protection Agency. 2009. The national study of chemical residues in lake fish tissue. Washington, DC.

US Environmental Protection Agency. 1995. ECOTOX database. Wahsington, DC. [cited 2014 October 21]. Available from: https://cfpub.epa.gov/ecotox/data_download.cfm

Vallotton N, Price PS. 2016. Use of the maximum cumulative ratio as an approach for prioritizing aquatic coexposure to plant protection products: A case study of a large surface water monitoring database. *Environ Sci Technol* 50:5286–5293.

Vörösmarty CJ, McIntyre PB, Gessner MO, Dudgeon D, Prusevich A, Green P, Glidden S, Bunn SE, Sullivan CA, Liermann CR, Davies PM. 2010. Global threats to human water security and river biodiversity. *Nature* 467:555–561.

Zijp MC, Posthuma L, Van de Meent D. 2014. Definition and applications of a versatile chemical pollution footprint methodology. *Environ Sci Technol* 48:10588–10597.