Eco-epidemiology of aquatic ecosystems: Separating chemicals from multiple stressors

Leo Posthuma [a,f,*], Scott D. Dyer [b], Dick de Zwart [a,c], Katherine Kapo [d], Christopher M. Holmes [d], G. Allen Burton Jr. [e]

[a] RIVM, Centre for Sustainability, Environment and Health, P.O. Box 1, 3720 BA Bilthoven, The Netherlands
[b] The Procter & Gamble Company, Cincinnati, OH, USA
[c] DdZ Ecotox, Odijk, The Netherlands
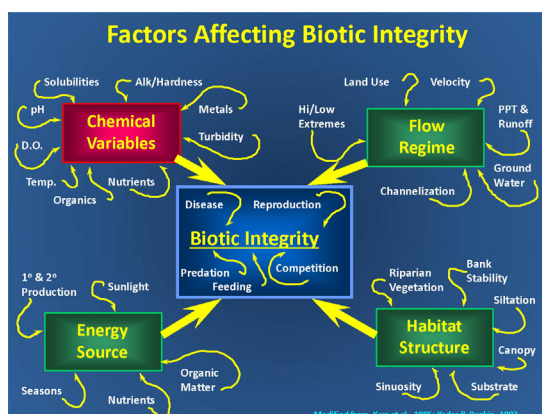[d] Waterborne Environmental, Inc., Leesburg, VA, USA
[e] School of Natural Resources & Environment, University of Michigan, Ann Arbor, MI 48109, USA
[f] Department of Environmental Science, Institute for Water and Wetland Research, Radboud University Nijmegen, Heyendaalseweg 135, 6525 AJ Nijmegen, The Netherlands

HIGHLIGHTS

- Recent 'non-toxic environment' and 'good ecological status' goals require mixture analyses.
- Human-dominated aquatic systems have multiple chemical, physical and biological stressors.
- Differences in chemical mixture impacts can be predicted by a simple proxy.
- Stressor impacts can be ranked across sites and within sites.
- Contributions of mixtures and individual chemicals to ecological impacts can be diagnosed.
- Stressor ranking results support and enhance alternative management strategies.

GRAPHICAL ABSTRACT

ABSTRACT

A non-toxic environment and a good ecological status are policy goals guiding research and management of chemicals and surface water systems in Europe and elsewhere. Research and policies on chemicals and water are however still disparate and unable to evaluate the relative ecological impacts of chemical mixtures and other stressors. This paper defines and explores the use of eco-epidemiological analysis of surveillance monitoring data sets via a proxy to quantify mixture impacts on ecosystems. Case studies show examples of different, progressive steps that are possible.

Case study data were obtained for various regions in Europe and the United States. Data types relate to potential stressors at various scales, concerning landscape, land-use, in-stream physico-chemical and pollutant data, and data on fish and invertebrates. The proxy-values for mixture impacts were quantified as predicted (multi-substance) Potentially Affected Fractions of species (msPAF), using Species Sensitivity Distribution (SSD) models in conjunction with bioavailability and mixture models.

* Corresponding author.
  E-mail address: leo.posthuma@rivm.nl (L. Posthuma).

Multiple stress
Stressor ranking

The case studies summarize the monitoring data sets and the subsequent diagnostic bioassessments. Variation in mixture toxic pressures amongst sites appeared to covary with abundance changes in large (50–86%) percentages of taxa for the various study regions. This shows that an increased mixture toxic pressure (msPAF) relates to increased ecological impacts. Subsequent multi-stressor evaluations resulted in statistically significant, site-specific diagnosis of the magnitudes of ecological impacts and the relative contributions of different stress factors to those impacts. This included both mixtures and individual chemicals. These results allow for ranking stressors, sites and impacted species groups. That is relevant information for water management.

The case studies are discussed in relation to policy and management strategies that support reaching a non-toxic environment and good ecological status. Reaching these goals requires not only focused sectoral policies, such as on chemical- or water management, but also an overarching and solution-focused view.

1. Introduction

Many surface waters are ecologically impaired due to combinations of stressors of different kinds (Burkhead, 2012; EEA, 2012; ETC/ICM, 2012). This is at variance with policy goals to safeguard water as a key heritage and as vital and scarce resource (EC, 2000; Carpenter et al., 2011; Hoekstra and Mekonnen, 2012; Hoekstra and Wiedmann, 2014; U.S. Senate, 2014). Mixtures of toxic chemicals are a specific cause for concern, as they cause widespread exposures of biota and associated risks to aquatic life (Schwarzenbach et al., 2006; USEPA, 2009; Malaj et al., 2014; Diamond et al., 2015; Stehle and Schulz, 2015).

The goal for a non-toxic environment has been formulated explicitly relatively recently (EC, 2014), although environmental regulations around the globe have also aimed at this. In this recent formulation it implies however: absence of impacts of pollutant mixtures on human health and the environment. This novel specified goal requires assessments of mixtures next to separate chemicals (EC, 2012), of evaluating impacts rather than risks (Bjørn et al., 2014; Zijp et al., 2014), of accounting for chemicals that are not subject to regulation or monitoring (Sjerps et al., 2016) and consideration of multiple stressor influences such as water scarcity (Heugens et al., 2001). Of particular interest here is flow, as this can be a direct stressor as well as a modifier of chemical concentrations (Carpenter et al., 2011; Kuzmanović et al., 2016; Sabater et al., 2016). Finally, landscape and land-use patterns act as overarching drivers of ecological status and impairments (Allan, 2004; Cervantes-Yoshida et al., 2015). Due to these characteristics, the non-toxic environment goal is more specific than before. Assessment of the presence of a non-toxic aquatic environment requires multi-stressor diagnostic approaches, in which the role of toxic chemicals can be delineated on a land- and riverscape scale. This differs from classical chemical risk assessments.

The classical approaches for assessing water quality and ecological risk of chemicals are often disjointed in terms of research and regulation. The literature on bioassessment (Metcalfe, 1989; Barbour et al., 1999; Vander Laan et al., 2013) does not connect well with the chemical risk assessment literature (Schwarzenbach et al., 2006; Van Leeuwen and Vermeire, 2007), so that e.g. transgression of risk-based safe criteria are not predictive of mixture impacts (e.g., Carvalho et al. (2014)). Separate regulatory frameworks exist in Europe for chemicals (REACH, concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals) and surface waters (WFD, the Water Framework Directive), though REACH liaises to water policies (Article 2.4), while water policies address chemicals too via defining Good Chemical Status (EC, 2000; EC, 2003). A similar situation can be found in other countries (for an overview of complexes of pollutant-related regulatory frameworks within and across jurisdictions: see Geiser (2015)). This disjointed state is likely due to a number of reasons, such as the scientific complexity of an infinite number of chemical-biota interactions (Hendriks, 2013), the associated 'curse of dimensionality' where large numbers of chemical parameters reduces the statistical power of bioassessments to (near) zero (Bellman, 1961), and simply disparate disciplinary focuses. It is therefore not surprising that mixture impact assessments are under-represented in bioassessments (Metcalfe, 1989; Norris et al., 2007; Friberg et al., 2011), though recent studies increasingly try to bridge the gap (Chiogna et al., 2016).

In this paper we aim to continue bridging the gap between bioassessments and mixture impact assessments as shown in Fig. 1. We do this by eco-epidemiological analysis of surveillance monitoring data sets. The term 'eco-epidemiology' was coined in 1984 (Bro-Rasmussen and Løkke, 1984). It was introduced as an approach to diagnose the relative importance of chemical impacts on ecosystems subject to multiple stress conditions. Since 1984, the concept has matured into an approach that uses bioassessments to characterize ecological impairments, followed by causal analyses to determine causes, with linkages to risk assessment and decision making (Suter et al., 2007). In our studies, the surveillance monitoring data sets were supplemented by a proxy for mixture impacts on ecosystems. The aim was to improve our understanding of multi-stressor impacts in aquatic ecosystems, and subsequently advancing chemical- and water management to attain the goals of a non-toxic environment and good ecological status. We address continental to local spatial variability while diagnosing the relative role of mixtures of chemicals in shaping local species assemblages

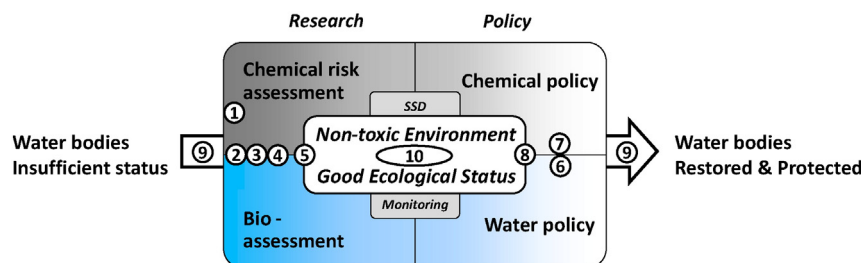**Fig. 1.** The nexus of research and policies in chemical and water management, with the central policy goals of a non-toxic environment and a good ecological status, in the context of a solution-focused assessment paradigm (water bodies restored and protected in a sustainable way, see Zijp et al. (2016)). The numbers reflect the schematic position of the Case Studies presented in this manuscript (see Table 1).

in the presence of multiple stressors. We focus on prioritization. This can concern ranking impact magnitudes across water bodies (Norris et al., 2007), ranking stressor types within water bodies (Fedorenkova et al., 2012), and/or ranking ecotoxicity of chemicals within a mixture (Brack et al., 2015). All such prioritizations are relevant for deriving (cost) effective water quality management plans. In addition, we address 'emerging chemicals' (UNEP, 2013) and options for predicting net impacts of chemical use (Murray et al., 2010; Pal et al., 2010; Pistocchi, 2014).

In order to improve our understanding of multi-stressor impacts in aquatic ecosystems and forward the decision support towards reaching a non-toxic environment, we describe and evaluate eco-epidemiological methods that include methods to diagnose and predict the impacts of chemical mixtures. We start from aquatic monitoring data, as those are being collected in vast numbers (e.g., via surveillance monitoring) triggered by existing law (EC, 2000; U.S. Senate, 2014). Survey data at various scales are required to answer questions like (a) what are the states of our water systems?, (b) do these states change in time or vary in space, positively or negatively, or comparatively?, and (c) if so, what drives these changes or differences? We explore the state of art regarding the diagnostic and prognostic analysis of multiple and emerging stressors footprints in surface waters over large geographical areas, such as national, states or catchments. Working at this scale has three motives. One, gradients of stressor data need be 'long' when statistical power is to be reached in bioassessments (large inter-site variability across sites). Two, for the same reason the gradients of degree of ecological impacts need be 'long'. Three, policy solutions are required on the scale of a safe chemical economy (Geiser, 2015) in relation to safe water supplies (Hoekstra and Wiedmann, 2014), which are large-scale issues.

2. Case study data sources and approaches

2.1. Current pollutants assessment

Current assessments of chemicals often focus on separate chemicals for environmental protection (European Communities, 2003) or for the characterization of environmental status (EC, 2000), while whole effluent toxicity tests are applied to monitor and control pollutant mixture emissions. Expanding on these common methods, we propose to add methods to diagnose and predict impacts of chemical mixtures in the context of co-occurring stressors based on monitoring data. These methods focus directly on ecosystems and species occurring in field conditions that are exposed to mixtures and multiple stresses under field conditions, while encompassing inter-species interactions. We apply a suite of methods in exposure and effects assessment, as described in the case studies, and earlier introduced by De Zwart et al. (2006) and Kapo et al. (2014).

2.2. Bioassessments of mixture impacts

The case study results presented in this paper expand on earlier analyses of the same data sets, and provide – more than the separate studies published earlier – an holistic overview of subsequent approaches and result types that can be obtained to provide the insights required according to Fig. 1. The case studies are causal assessments. A causal assessment should deliver technical support for decision making, and often refers to a particular situation, system or place, sensu Norton et al. (2014). The studies considered here, however, all concern the analysis of landscape-level monitoring data on species assemblages, in order to delineate spatial variability in ecological impairment and its probable causes, with special emphasis on mixture impacts. Full causal analyses are not made, as those would require additional assessment steps (Norton et al., 2014), which are beyond the illustrative nature of the case studies.

2.3. Eco-epidemiology of chemical mixture impacts

Though the term eco-epidemiology dates back to 1984 (Bro-Rasmussen and Løkke, 1984), eco-epidemiological studies remained scarce for some decades. Often, chemical risks are evaluated via checking transgressions of ambient Water Quality Criteria of individual compounds (i.e., the PEC/PNEC-ratio type of evaluation, with PEC = Predicted Environmental Concentration and PNEC = Predicted No Effect Concentration), though scientific evidence suggests that such transgressions do not necessarily predict ecological impacts (Solomon and Takacs, 2002; De Zwart et al., 2009; Burton et al., 2012; Carvalho et al., 2014) — that is: transgressions may trigger management action while there are no ecological impacts. In the mid-2000's, however, the amounts of monitoring data had increased vastly, and the application of Species Sensitivity Distribution (SSD) modelling in conjunction with exposure- and mixture models reduced the 'curse of dimensionality' (De Zwart et al., 2006; Kapo and Burton, 2006). An SSD model is a probability distribution describing the sensitivity of multiple species exposed to a hazardous compound (Van Straalen and Denneman, 1989; Posthuma and De Zwart, 2014), at a certain impact level, using e.g. the EC50 as sensitivity test endpoint. Bioavailable concentrations of chemicals can be expressed in a proxy named 'toxic pressure'. The toxic pressure estimates the fraction of species that is probably affected to a certain extent by an ambient exposure level. It is expressed as the Potentially Affected Fraction$_{Endpoint}$ of species affected to a certain extent, e.g., PAF$_{EC50}$ for single compounds, or as multi-substance PAF (msPAF) for mixtures (De Zwart and Posthuma, 2005). This proxy reduces the number of stressor variables from many (all chemicals separate) to one per sample (total mixture toxic pressure), or to a few representing main compound groups (e.g., msPAF$_{PAHs}$, msPAF$_{Metals}$, et cetera). A local toxic pressure (msPAF$_{EC50}$) of 58% has the conceptual meaning that – when 100 species would have been tested to obtain sensitivity data for a chemical – 58% would exhibit >50% effects if the tested species would individually be exposed to this exposure (excluding ecological interactions). We acknowledge that SSD-modelling is not based on ecological considerations and that model outcomes not predict impacts in ecological terms, following key arguments in the debate on SSD-modelling in risk assessment (e.g., Forbes and Calow (2002) vis a vis methods such as population- of food web modelling (e.g., Forbes et al. (2008)). Supporting the development of the latter, we follow a lower-tier pragmatic approach in using the toxic pressure proxy for two reasons. First, the latter models cannot yet be applied in eco-epidemiology. Second, though mixture toxic pressure should be seen as a characteristic of the water sample, it is quantitatively related to ecological impacts (e.g., Posthuma and De Zwart (2012)).

After replacing the pollutant concentrations by the (mixture) toxic pressure data, the statistical analyses in the case studies were executed according to the various approaches common in aquatic bioassessments (Sponseller et al., 2001; Allan, 2004; Leps et al., 2015). Details of the steps made are provided with the case studies.

2.4. Case study data sources, approaches and limitations

Data sources for case studies are described with the case study results. The typical steps in data collection and management consist of the collection of monitoring data, the imputation of abiotic and biotic parameters, the expansion of the data set with additional landscape and land use data, adding information of scale (like proximity of land use to the water bodies), addition of mixture toxic pressure proxies (Kapo et al., 2014), and screening of the resulting data set. The mixture toxic pressures were based on measured or predicted concentrations (e.g., via the iSTREEM® model (www.iStreem.org)). When observed or predicted concentration data lack, toxicity was also evaluated by proxy variables, such as broad-scale usage data (e.g., from agricultural census), land cover (to identify agricultural land) and proximity to streams (e.g., agriculture within 100 m). Notably, the derivation of

mixture toxic pressures was tailored to the exposure ranges in the area under investigation, that is: msPAF$_{NOEC}$ or msPAF$_{EC50}$ at lower or higher exposure concentration. Ecological impacts were quantified in a variety of ways in the different case studies, e.g. abundance per taxon, trait frequencies, biodiversity indices or metrics quantifying deviation from expected conditions (Lenat and Resh, 2001; Baird et al., 2008; Herman and Nejadhashemi, 2015; Pilière et al., 2016). The resulting data set was screened, by checking in various ways for redundant variables, e.g., via analysis of correlations and of Variance Inflation Factors (Kline, 1998; O'Brian, 2007). Variables that would invalidate diagnostic inferences were removed from the statistical analyses. The data sets of the case studies are all statistically appropriate for evaluating the relative contribution of chemicals and their mixture to ecological impacts, as they show a large variability of the (mixture) toxic pressure across sites and no- or limited covariation to other stress factors.

Causal studies that are based on large-scale monitoring data have limitations. Evidently, it is evident to collect good quality data, with special emphasis on the proper co-location of stressor and response variables (Dyer et al., 2000), and to study a data set that represents the situation under investigation, Interpretations on the role of pollutants evidently relates only to the compounds measured or modelled. Finally, the final insights show statistical associations between stressor variables and biotic variables rather than proven causation.

2.5. Case studies and water quality assessment questions

The case studies concern different problem definitions in water quality assessment and management (Fig. 1 and Table 1). They all concern the statistical analyses of large-scale monitoring data sets in which the mixture-toxic pressure approach was applied to study statistical associations between mixture and biotic parameters. The study data concern various regions in Europe and the U.S. Note that the approaches were initiated by- and relate to regulatory requirements, such as the need to use a reference concept in regulatory bioassessment schemes.

3. Case study results

All case study descriptions provide a problem definition, and summarized methods, results and implications. Note that (for ease of communication) we use cause-effect terms rather than the more appropriate terminology of statistical association. The case studies are characterized by sufficient variation of the mixture toxic pressure proxy and by lack of covariation of this proxy with other potential stressors.

3.1. Mixture toxic pressure (co)variation

3.1.1. Problem definition

Statistical associations between mixture exposure levels and biotic variables can only be meaningfully interpreted when there is sufficient variability of mixture exposure levels, and when the covariation of this parameter with other stressors is absent or limited. This case study illustrates the analysis of variation and covariation of mixture toxic pressure and other abiotic variables.

3.1.2. Methods

Monitoring data were collected for the analysis of abundances of benthic invertebrate assemblages to multiple-stress conditions in a region in the Netherlands (Posthuma and De Zwart, 2012). The data set concerns 220 benthic macroinvertebrate taxa, approximately 550 sites, and 12 abiotic variables. Mixture toxic pressures were calculated, based on 45 measured compounds and SSD$_{EC50}$-modelling. Variation and covariation of the predictors was analysed.

3.1.3. Results

Cumulative distribution functions of the predictors are shown in Fig. 2. An ideal data set would exhibit 'diagonal' distributions. The mixture toxic pressure distribution for this data set shows (bold line) wide variability of toxic pressure between samples, as well as the distributions of other predictors (continuous or stepwise distributions). Analysis of the Variance Inflation Factors showed that the covariation of mixture toxic pressure with the other predictors would not invalidate inferences. This implies that statistical associations between mixture exposures and biotic parameters based on these data are unbiased.

3.1.4. Discussion

This case study is just one example of inter-site variability of mixture toxic pressure on a landscape scale, and limited covariation with other stressors. When study areas are large, experience has learned that the toxic pressures of individual chemicals, subgroups of chemicals and of total mixtures varies across sites too, while the probability of covariation with other stressors will reduce (they are not invalidating the inferences in the case studies presented below). The requirement that the chemical pollution proxy has no- or limited covariation with other stressor variables likely requires data from large study areas, and critical evaluation of both statistical- or non-statistical (logical) sources of covariation between mixture toxic pressure and other stressors.

Table 1
Case studies used to illustrate problem definitions, approaches and outputs types in the eco-epidemiological analysis of monitoring data. The case studies present sequential steps, shown via the numbers in Fig. 1. Case 1 concerns the data set, cases 2–4 concern the scientific interpretation of the concept of toxic pressure (fundamental for the further cases), and the cases 5–10 concern analyses made in the context of decision support.

| Case #	Case identification	Problem definition
1	Mixture toxic pressure (co)variation	What is the current distribution of mixture toxic pressure variability? A statistical association between mixture toxic pressure and biotic parameters requires sufficient variability of this metric and low covariation with other predictors.
2	Mixture exposures and species abundance changes – pollutant stress	Species sensitivity differences imply expected different species abundance changes; can these be elucidated with Pearson correlation analyses?
3	Mixture exposures and species abundance changes – multi-stressor	As in the case with pollutant stress, but now with multi-stress. Can chemical pollution responses be elucidated with multi-variate statistical models?
4	Calibrating mixture toxic pressure to ecological impacts	The mixture toxic pressure proxy can be derived from laboratory test data, but it requires calibration to observed impacts. What does the calibration look like?
5	Developing preliminary working hypotheses	Scale matters in bioassessments, and this also hold for chemical mixture impacts. Can preliminary data explorations support a proper diagnostic analysis of the monitoring data set?
6	Ranking sites and stressors	Can data analyses result in site-specific information that supports a useful ranking of impact magnitudes and of the relative importance of probable causes?
7	Ranking chemicals within mixtures	Ibidem for chemicals or chemical groups within impacted sites?
8	Reference sites and mixture toxic pressure	Are reference sites, used in the regulatory context of water management, indeed non-toxic, and also relatively unaffected by other stressors?
9	Abatement policies and abatement success	Can retrospective evaluation of monitoring data elucidate the presence of success of investments in abatement, and show reduced mixture impacts?
10	Restoration target	Is causal analysis sufficient, or is attention needed for alternative restoration and management scenarios?
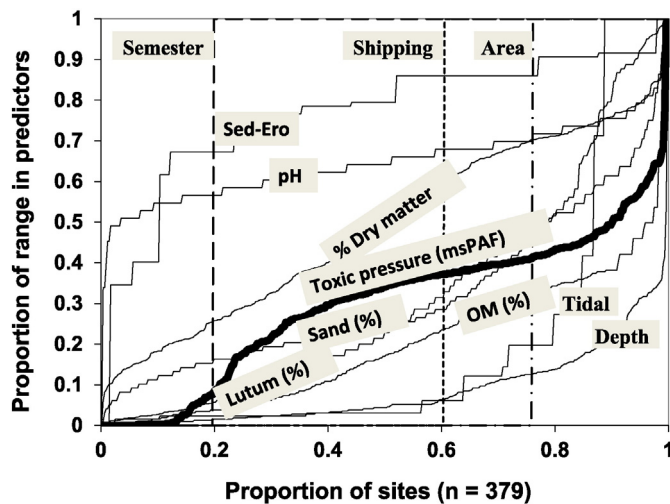
Fig. 2. Cumulative distributions of predictor variability for a sampling area. Names of metrics refer to stressor variable types as follows: 'Semester' relates to sampling time (spring or fall); 'Shipping' considers presence or absence of shipping; 'Sed-Ero' consider the sedimentation-erosion balance; pH represents acidity; '%dry matter', 'sand%', 'lutum%', and 'OM%' represent the percentages dry matter, sand, clay and organic matter of the sediments; 'tidal' represents the influence of tide, and 'depth' represents depth of sampling. Toxic pressure (%) represents the msPAF$_{EC50}$ derived from measured chemical pollutant concentrations.

3.2. Mixture exposures and species abundance changes — main effect

3.2.1. Problem definition

The SSD model is based on the observation that species respond differently to exposures to single chemicals under test conditions (Stephan et al., 1985; Van Straalen and Denneman, 1989). Does this logically mean that species will also respond differently to mixture exposures under both laboratory and field conditions? What would be the possible influence of multiple stressors and ecological interactions, and could these influence species responses in the field. What do mixture responses under field conditions look like?

3.2.2. Methods

Using the same data set as in the previous example, the association between mixture toxic pressure (msPAF$_{EC50}$) and the abundances of 103 individual benthic macrofauna taxa was established and ranked.

3.2.3. Results

Rank-ordered correlation coefficients between mixture toxic pressure and taxon abundance varied widely (Fig. 3). No correlation was found for the majority of taxa (grey band). Further, as expected from the species sensitivity variation phenomenon, there was a negative correlation for various taxa, with more and less sensitively responding taxa. Additionally, and not expected from the SSD-modelling of laboratory test data on species sensitivities, some taxa exhibited an opportunistic response. There were no strong and highly significant associations between toxic pressure variability and abundances of taxa.

3.2.4. Discussion

Diagnostic analyses of monitoring data can unveil a suite of direct, taxon-specific statistical associations between mixture exposure and abundance changes, part of which (abundance increases) are unexpected from the species sensitivity variability phenomenon. However, similar results were repeatedly found in other mixture impact assessment studies (De Zwart et al., 2008; De Zwart et al., 2009), and they are also typical for other stress parameters (Sundermann et al., 2015). These results imply that SSD-modelling yields a quantitative proxy which is expressed in terms of ecological impact magnitude (fraction of species potentially affected), and that a limited level of ecological responses
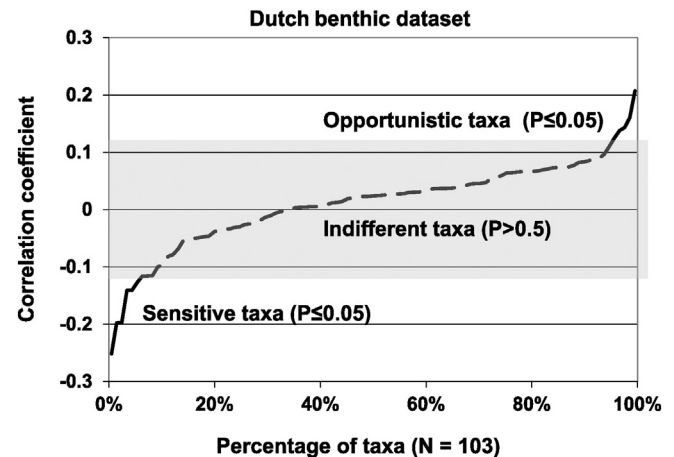

Fig. 3. Rank-ordered Pearson correlation coefficients between mixture toxic pressure (msPAF$_{EC50}$) and taxon abundance for 103 taxa.

(and response variability) can be found, though it needs be noted that this so far neglects multiple-stress influences. Moreover, the method does not predict which species exhibit abundance reductions, nor does it predict the presence of abundance increases in some species. The latter are likely attributable to ecological interactions, e.g., due to reduced interspecies competition for resources, or due the loss of a sensitive predator causing abundance increases in prey. When studying impacts of chemicals in natural systems, the variability of response patterns across taxa should be acknowledged.

3.3. Mixture exposures and species abundance changes – multiple stress

3.3.1. Problem definition

The two previous examples showed the presence of multiple stresses and of multiple types of species responses to mixture exposures, with many species apparently indifferent to mixture exposure. The latter may, however, be misleading, due to other stressors masking the recognition of impacts of mixtures in direct correlation analyses. Do multistress conditions mask mixture impacts?

3.3.2. Methods

Various data sets (references in caption of Table 2) were analysed using statistical models allowing for a multiple-stress interpretation. In the examples, GLMs (Generalized Linear Models) were fitted to the monitoring data. The models all had the shape of log($abundance$) = ($a*msPAF + a'*msPAF^2$) + ($b*predictor\ B + b'*predictor\ B^2$) + … + ε, resulting in estimates of a, a', b, b' (et cetera). The fitted model allows for linear as well as quadratic (optimal, minimum) response shapes and their combinations. A typical example of the variables in the studied data sets is shown in Fig. 2.

3.3.3. Results

Analyses of the various data sets showed that the abundances of a high proportion of taxa (50–86% of the taxa, at the levels of species, genus and family) are related to mixture toxic pressure in surface water bodies. The abundances of the taxa are similarly frequent co-determined by other physical and chemical stressor variables (data in cited studies). Taxon abundances are determined by multiple stresses.

3.3.4. Discussion

In comparison to the Pearson correlations of the various data sets studied (see above), the percentages of taxa with a significant toxic pressure – abundance association is high, with a minimum estimate of 50% (fish, species level) and a maximum value of nearly 90% (macrofauna, genus level). The repeated high percentages of taxa responding to toxic pressure is surprising, as many large-scale evaluations do not

Table 2

Percentages of taxa (different taxonomic levels) for which a GLM-term (linear and/or quadratic) marks a significant influence of mixture toxic pressure (determined as msPAF$_{EC50}$ or msPAF$_{NOEC}$), for various subsets of measured or modelled chemical concentrations. References are: (1) (Posthuma and De Zwart, 2006); (2) (De Zwart et al., 2008); (3) (De Zwart et al., 2009), (4) (Posthuma and De Zwart, 2012); (5) (Posthuma et al., 2016). Note: data set (4) analysed taxa abundance data when found in >10 samples across the Netherlands, but also a sub-set of taxa with a statistically robust data structure (taxa found in all sampling sites). 'Var.' reflects that the number of chemical pollutants is large, covering a wide variety of compounds.

Study area	Response variable	msPAF	Compounds	Other stressors	Ref.
Ohio (US)	Fish (n = 96 taxa, species level)	55% (NOEC)	Household	43%–99%	1
Ohio (US)	Fish (n = 96 taxa, species level)	50% (NOEC)	Metals + ammonia	43%–99%	1
U.K.	Macrofauna (n = 76 taxa, family level)	71% (EC50)	Metals + ammonia	55%–73%	2
U.K.	Macrofauna (n = 76 taxa, family level)	56% (EC50)	Pesticides (model)	55%–73%	2
Belgium	Macrofauna (n = 64 taxa, mostly family level)	59% (EC50)	Var. (n = 335)	66%–80%	3
Netherlands	Benthic invertebrates (n = 103 taxa, genus level)	74% (EC50)	Var. (n = 45)	72%–89%	4
Netherlands	Macrofauna (n = 308 taxa, genus level)	59% (EC50)	Var. (n > 1750)	38%–65%	5
Netherlands	Macrofauna (n = 110 taxa, robust genus level data)	86% (EC50)	Var. (n > 1750)	64%–84%	5

suggest such clear and high numbers for toxicity (e.g., (EEA, 2012)). The results show the presence of (combinations of) linear and quadratic terms for mixture toxic pressure as well as for other stressors. The net abundance changes of species, genera and families are thus commonly a result of multiple-stressor influences which vary amongst taxa (shown by different compositions of the GLM-models). In conclusion, increased mixture toxic pressure is a proxy that predicts increased change in ecosystems, associated with species-specific downward-, neutral- and upward abundance changes. It predicts degree, not kind of impact, and the evidence suggests that such impacts are present in majority of taxa at current ambient chemical exposure levels.

3.4. Calibrating the mixture toxic pressure proxy to impacts

3.4.1. Problem definition

The relationships between predicted impact (e.g., msPAF$_{EC50}$) and observed impacts have not yet been calibrated to field impact observations, neither for taxon-specific abundance changes, nor for the observed fraction of species affected. How does the predicted mixture impact proxy relate to observed ecological impacts? This case study shows analyses based on surveillance data, but also for two well-defined contamination gradients. The latter approach reduces the influences of possible confounding factors.

3.4.2. Methods

Landscape study data are similar to those in Section 3.1 (Posthuma and De Zwart, 2012). The relationships between predicted impacts and observed impacts were based first on taxon abundance values predicted by the GLMs in combination with the monitoring data (taxon-specific response variation), and then by calculating the percentages of species showing >50% abundance change (summarizing the taxon-specific responses). Validation study data were obtained from two terrestrial pollution gradients.

3.4.3. Results

Species abundance patterns and observed fraction of species affected (responding >50% in abundance change, upward or downward) are shown in Fig. 4. The two graphs (upper and lower) are for the same data set, and show (bottom) the variance in taxon-specific abundance changes, and (top) their aggregation in the observed fraction of species affected (>50% abundance change in the field). The association between predicted and observed fraction of taxa negatively affected at the EC50-level (top, continuous line) is surprisingly close to the ideal 1:1 line. Note that the fraction of species with an increased abundance response type (dotted line) responds at lower proxy levels than the decrease abundance response pattern.

3.4.4. Discussion

The mixture toxic pressure proxy relates to a complex set of responses, which are true at the same time: (1) increasing toxic pressure relates to a variety of neutral and up- and downward abundance

changes of separate taxa, and (2) to associated net fractions of species affected. This was also found in a similar analysis of a nationwide monitoring data set for the Netherlands (Posthuma et al., 2016). In the current study, pollution seems to trigger opportunistic species to rise in abundance prior to triggering negative abundance change responses on other species. A similar divergence of taxon-specific response patterns to chemical exposures is not unique for this case study. It has also been found in gradient studies (Fig. 5), where the impact of confounding factors is low.

3.5. Developing preliminary working hypothesis

3.5.1. Problem definition

Chemical mixture exposures in surface waters result from point- (Waste Water Treatment Plants, WWTPs) and non-point sources (pesticide or fertilizer use in the landscape upstream) in combination with
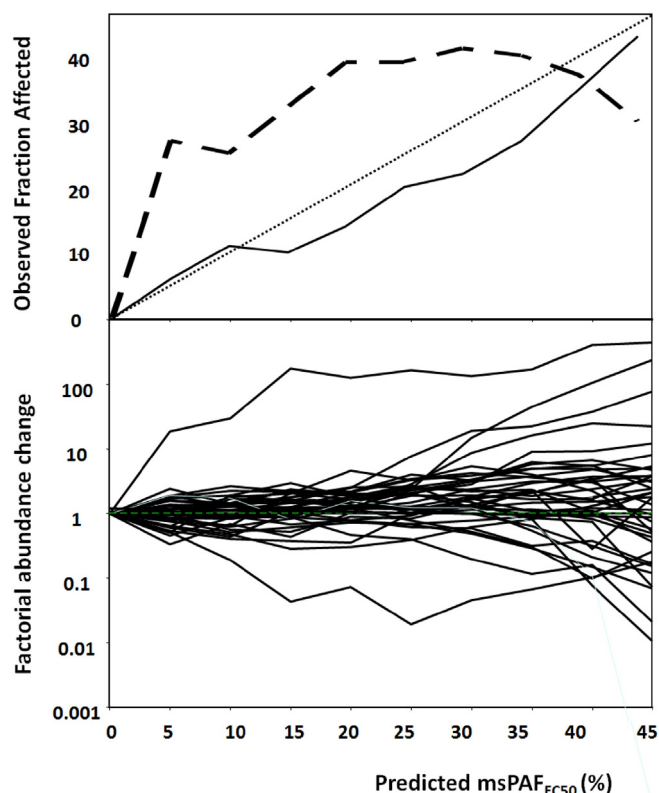

Fig. 4. Calibration of the msPAF$_{EC50}$-proxy to the observed factorial abundance change patterns for 45 taxa (bottom) and the observed fraction of species affected (top). The continuous and broken lines (top figure) are the fractions of species with negative and a positive abundance change, respectively. Dotted line: ideally expected 1:1 relationship between predicted and observed affected fraction.
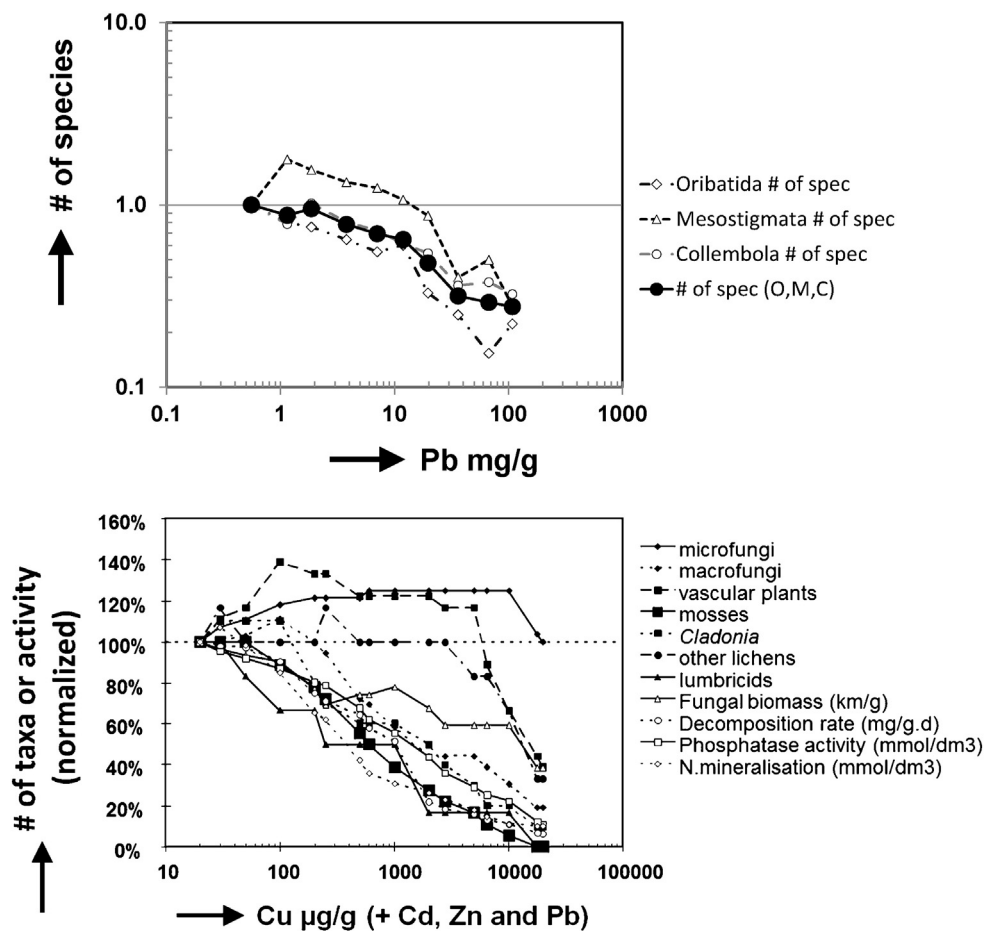
Fig. 5. Terrestrial gradient studies corroborate the abundance change patterns found in Fig. 4. Top: a lead contamination gradient at Kastad, Norway. Data from Låg et al. (1970) and Hågvar and Abrahamsen (1990). Bottom: a mixed-metal gradient at Gusum, Sweden. Data from (Tyler, 1984). Results summarized from Posthuma (1997) and Posthuma et al. (2001).

hydrological and chemical processes (dilution, breakdown, et cetera). Chemical exposures and impacts thus expectedly vary across sampling sites, with an expected key role for multiple stress (see earlier case study). In line with bioassessment experiences, the scale of observations and of causal analysis is important to consider for chemical pollution too (Mykrä et al., 2007; Li et al., 2012). Hence, the recognition of mixture impacts in large monitoring data sets is also expected to improve when emission-concentration-impact analyses are considered at scales relevant for the chemical groups of interest.

3.5.2. Methods

Monitoring data were collected for Ohio (a state covering seven ecoregions) between 2000 and 2008 (Kapo et al., 2014), starting with 3000 sites, of which 1917 sites had data for all parameters. Mixture toxic pressures were determined for (separately) pesticides (n = 156), "down-the-drain" consumer product chemicals (n = 7, household product constituents), pharmaceuticals (n = 49), estrogens (n = 3) and conventional pollutants (n = 11, metals, ammonia, and nitrite), based on measured and modelled concentrations. Conditional inference forest analysis (Strobl et al., 2007) was applied separately at the state-wide and ecoregional levels (n = 3) to provide initial evaluations of the relative influences of the various mixture toxic pressures and other stressors on biological communities. The initial evaluations provide information on associations between stressors and community status, but do not yet provide diagnostic information on ecological impacts per site and the possible causal factors of those impacts (see case study below). Therefore, the results of these analyses are helpful to recognize important patterns in the data, which can thereafter be used in

the refined diagnostic studies (example below). Response metrics were the Index of Biotic Integrity (IBI, based on 12 fish assemblage metrics) and the invertebrate community index (ICI, based on 10 invertebrate assemblage metrics).

3.5.3. Results

The analysis showed an important role of scale in the analysis of emission-concentration-impact hypotheses. As a major result, the data analyses showed that there was a strong influence of physical habitat quality on biological community condition at the state-wide and ecoregional scale compared with most other stressors. Given those influences, the mixture toxic pressure of conventional pollutants was negatively associated with biological community condition, although its relative influence at the state-level appeared to be lower compared to other variables such as habitat and other water chemistry factors (e.g., conductivity, biological oxygen demand, and related parameters). At a more refined scale, of the ecoregions, the relative influence of the mixture toxic pressure for conventional pollutants appeared to be markedly higher, e.g., for the Western Allegheny Plateau. This outcome is in line with the fact that the region is known for acid mine drainage and industrial pollution, and it confirms the results of an earlier study which was executed with an older monitoring data set (De Zwart et al., 2006). The mixture toxic pressure for 'down-the-drain' consumer products, which was highly correlated with toxic pressure of pharmaceuticals, had a relatively low influence compared with other stressors at both the state and ecoregion scales. The greatest potential negative influence of chemical emissions typical for human settlements on biological communities was found in the Erie Drift Plain ecoregion, an

area with high population densities. The mixture toxic pressure of pesticides (evaluated for 159 compounds) was highly correlated with a GIS-derived percentage agricultural land use. Further details on the outcomes of the preliminary analyses are provided by Kapo et al. (2014).

3.5.4. Discussion

The role of chemicals in shaping biological community conditions likely varies due to land use differences. Neglecting scale and landscape issues, and looking only at a state-wide scale, may imply a statistical 'dilution' of an impact-signal for mixture effects in the data set. That is, impacts that are present in the data set are not traced for statistical reasons (e.g., the clear effect of metal mixtures at the ecoregion scale where metal exposures mainly occur was hardly recognized at the state scale). The preliminary diagnostic analysis generated initial stressor-response information with a recognized role of scale. The preliminary analysis results can offer insights into the possible mechanisms with which landscape variables likely influence biological condition. Some locally relevant aspects (emission sources) induce high inter-site stressor variability, while others can be addressed via large-scale land-use proxies for initial recognition of stress patterns. The initial data exploration illustrated here serves the purpose of integrating and optimizing available raw data, delineating initial stressor-response associations over large geographies (e.g., state and ecoregion), and laying the groundwork for subsequent steps such as more refined diagnostic assessment targeted on developing, and eventually testing, stressor-response hypotheses at the local (site-specific) level. Delineating the impacts of chemical mixtures requires not only a metric like mixture toxic pressure, but – for the sake of finding impacts when they are there, and for implementing solutions – also consideration of the different scales at which emission-, exposure- and impact phenomena occur.

3.6. Ranking sites and stressors

3.6.1. Problem definition

Upon collating the data set for the analysis of mixture impacts, and of preliminary scale-dependent approaches to investigate mixture impacts in a hypothesis-driven way, the final aims of eco-epidemiological analyses are: (1) the site-specific characterization of the relative importance of chemical mixtures vis a vis other stressors in causing ecological impacts, including consideration of the opportunistic responses (described in Sections 3.2, 3.3 and 3.4), and (2) the ranking of the relative importance of chemical groups or individual chemicals within sites (see Section 3.7).

3.6.2. Methods

Various methods have been proposed to derive site-specific information on impact magnitudes and probable causes. Two of these methods were applied first on Ohio monitoring data collected prior to the year 2000 (i.e., 1990–1995) (De Zwart et al., 2006; Kapo and Burton, 2006; Kapo et al., 2008a) and had shown site-specific impact magnitudes (expressed as species expected but absent based on RIVPACS-modelling, see Section 3.8) and their probable causes. The Weight of Evidence/Logistic Regression (WOE/LR) method was applied here to investigate mixture impacts using the more recent Ohio (2000–2008) data set. The latter data set is a vast expansion of the Ohio (1990–1995) data set, examining both direct impacts to flowing water (i.e., local catchment) and for entire upstream areas (Kapo et al., 2014). As illustration, the differences between the older and the newer Ohio data set are: (1) number of variables, after quantifying mixture toxic pressures: from 18 to 32; (2) the numbers of chemicals considered in the mixture toxic pressure variables: from 11 to 226; (3) the numbers of sites, from 1572 to 3000; and (4) the number of sites with complete data from 695 to 1917. The analyses of the newer data set considered the IBI of fish as the response variable. In addition, the Effect and Probable Cause (EPC) method was applied on the earlier data set (De Zwart et al., 2006) to investigate impact magnitudes and

probable causes for taxa showing increased abundances. The method can yield two types of pie diagrams for each sampling site, in which pie sizes either represent the fraction of species expected but absent or the fraction unexpected but present, and slice sizes to which probable causes the changes are attributed by statistical associations.

3.6.3. Results

The WOE/LR analysis characterized catchments (data 2000–2008) by relative probability of biological impact (Fig. 6), including the potential stressors hypothesized for each study catchment. The analysis output enabled prioritization of sites of interest, that is: the identification of catchments where mixture toxic pressure, such as those for metals or effluent constituents, was most likely to have an impact on biota (example for fish community condition in Fig. 6). Geographic trends in stressor-response can be visualized from the analysis, such as urban-impacted catchments in populated metropolitan areas such as Cleveland (enlargement in the figure noted by "C"), and catchments impacted by metals toxicity in the mining-impacted region of the Western Allegheny Plateau (south-east), corroborating the findings of the two earlier studies with the 1990–1995 surveillance data set.

The EPC analysis of the Ohio (1990–1995) data set had earlier resulted in 695 site-specific Effect and Probable Cause (EPC) diagrams for sensitive (species expected, but absent) responses (De Zwart et al., 2006), but the role of stressors in relation to opportunist responses (Figs. 3 to 5) was not yet analysed. To sketch an overview of both response types, EPC-results for the data set were collected for sensitive and opportunistic responses and then averaged according to land use. Note that the changes of taxa abundances for the averaged EPC's relate to a small number of chemicals represented in the mixture toxic pressure proxy of this data set (see above). Land uses were defined as the proportionally dominant land use (% surface cover) locally draining to the sampling site. Summary results are shown in Table 3. The EPCs for sensitive and opportunistic changes appeared to be of similar sizes (impact magnitude) and probable causes (attribution of impact to grouped predictor types). The relative role of mixtures in causing local ecological impacts is shown to be low on average (1% and 2%). However, local values up to 30% were found for the Ohio (1990–1995) case study (in line with the findings of the Sections 3.2 and 3.3), suggesting a skewed distribution of the relative role of mixture exposures to ecological impacts. When it is considered to implement abatement strategies, it is key to acknowledge the variability in the mixture proxy values for the mixture impacts across sites (like shown in Fig. 2).

3.6.4. Discussion

The case studies demonstrate that it is possible to derive outputs which provide insight in spatial differences in local impact magnitudes and the local set of probable causes, by analysing existing monitoring data with a combination of ecotoxicological and statistical modelling in a GIS-data management context. This type of output is needed in the context of the ambitions shown in Fig. 1, as the approaches and the results link bioassessment with chemical risk assessment, while yielding a basis of prioritizing sites and stressors per site for management. In deriving a statement on a toxic- or a non-toxic environment, the numbers of studied compounds and their representativity for the potential mixture problems, needs be specifically considered. Moreover, the methods focus on effect types captured in the ecotoxicity data (like growth, reproduction), so that specific attention may be needed for pollutant groups with a specific exposure pathway or mode of action. It is noteworthy that the approaches were applied not only to measured chemical concentrations, but also to predicted ones. This may be a valuable asset, especially when this is done with validated models. With such models, the diagnostic assessment is applied to a case where the exposure to pollutants is – in fact – known better than after measurements of water samples taken in a small set of sampling events.
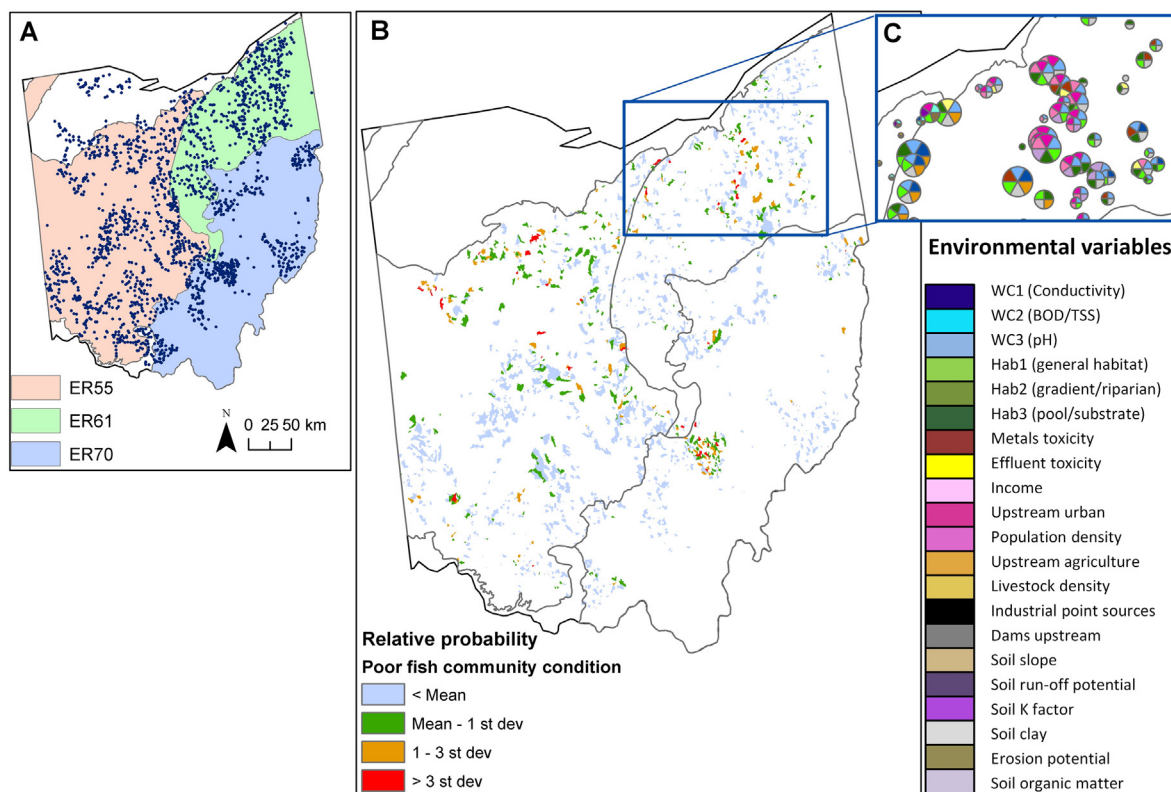
Fig. 6. WOE/LR spatial analysis of multiple-stressor influences, based on Ohio (2000–2008) data (Kapo et al., 2014). (A) Catchments (centroid points shown) with biological and environmental data were analysed for three ecoregions. This generated (B and C) catchments (shaded polygons) symbolized by (relative) impact probability (for "poor" Index of Biotic Integrity condition for fish as defined by Ohio EPA). In detail (C) the relative roles of potential local environmental variables (colours) contributing to the relative impact probability (pie size, relative impact probability). Further detail can be found in diagnostic studies below.

3.7. Ranking chemicals within mixtures

3.7.1. Problem definition

The previous case study provides insights in the relative relevance of local mixtures in causing local ecological impacts. It seems that in that step the focus on individual toxic chemicals is lost in the analyses, while the role of specific pollutants or pollutants groups are worthwhile to know in the context of programs of measures. However, the information on specific pollutants can be obtained from the underlying data, the toxic pressures of the individual compounds. Therefore, the local impact attributed to a mixture can be dis-aggregated into compound groups and compounds of specific interest.

3.7.2. Methods

Case study data for the Scheldt river were collected and analysed using the EPC-method (De Zwart et al., 2009). The case study data concerns a study on 335 chemicals (metals and organic compounds of industrial origin and pesticides), the concentrations of which were recalculated in mixture toxic pressures, msPAF$_{EC50}$, 13 predictors (including msPAF), 64 macroinvertebrate taxa, and 972 sites (31 used as reference). The EPC-method was applied to quantify the pie size (species expected, but missing as compared to the references) and slice sizes (fraction of species expected but missing attributed to a predictor variable). Thereupon, the mixture impacts were disaggregated in order to rank the relative contributions of chemicals to the impacts, looking at toxic pressures for subgroups of chemicals (metals, inorganic other compound, organics and pesticides).

3.7.3. Results

The study identified the highest impacts of chemical mixtures on macroinvertebrates in the Nete sub-catchment of the Scheldt catchment. The disaggregation of the mixture toxic pressure data clearly identifies the relative importance of metals in relation to ecological impacts (Fig. 7). The Nete sub-catchment is known to drain an area characterized by long-term metal smelter activities.

Table 3
Effect and Probable Cause (EPC) results for the state of Ohio (distinguishing various land uses), considering both sensitive species responses (top) and opportunistic responses (bottom), based on the data of De Zwart et al. (2006). Percentages are for impact magnitude (EPC pie size), where they reflect species lost or gained as compared to reference conditions, and the relative contribution of a subgroup of stressors. The last column summarizes unexplained variance in species abundance changes (down and upward).

| Response	Land use	Impact	Ranking of stressor type (predictors in subgroups) relevance to impact						
		Magnitude	PhysChem	Habitat	Flow	Land use	Toxicity	Effluents	Unexplained
Sensitive	Nature	41%	26%	17%	5%	6%	2%	1%	42%
Sensitive	Rowcrop	36%	25%	19%	7%	7%	2%	1%	39%
Sensitive	Urban	54%	22%	14%	11%	7%	1%	2%	43%
Opportunist	Nature	37%	27%	18%	5%	7%	1%	1%	42%
Opportunist	Rowcrop	40%	31%	16%	8%	4%	1%	1%	39%
Opportunist	Urban	31%	24%	15%	10%	6%	1%	0%	42%
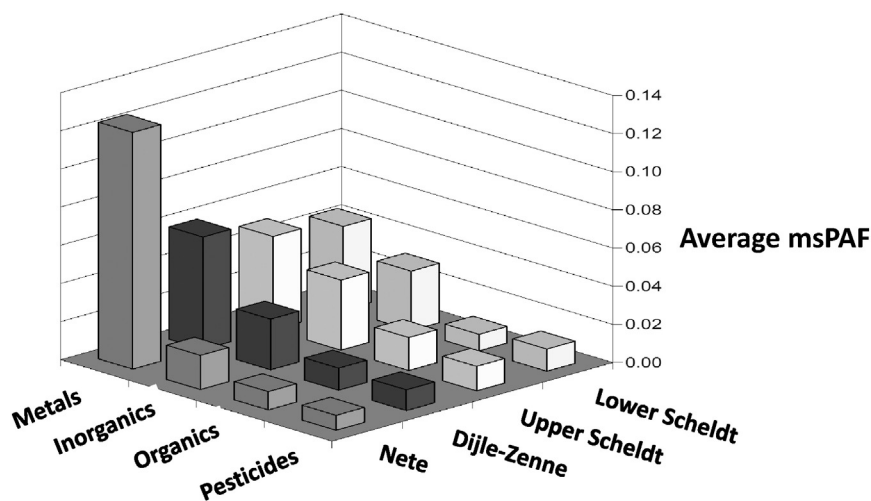
Fig. 7. Ecological impact analysis for the Scheldt basin in Belgium: disaggregation of the mixture impact results into the relative contributions of chemicals (here: chemical groups, per sub-catchment).

3.7.4. Discussion

The EPC and the spatial WOE/LR analysis approaches, while considered to yield exploratory stressor-response associations, delineate geographic patterns in stressor-response patterns (magnitude of impacts, as compared to the references) and site-specific attribution of impacts to probable causes (individual stressor variables, amongst which toxic pressure). The outputs support management prioritization options, subsequently: (1) sites in the landscape, (2) stressor within sites (including mixtures), and (3) individual chemicals or chemical groups within sub-catchments or sites. All these types type of information are useful in the design and evaluation of optimal management strategies, site selection strategies for field monitoring, as well as achieving a better understanding of chemical risk in a multiple-stressor context.

3.8. Reference conditions and mixture toxic pressure

3.8.1. Problem definition

Regulatory frameworks for water quality management require the use of benchmarks or reference conditions defined by physical, chemical, and ecological characteristics, against which impacts are judged (EC, 2000; Stoddard et al., 2006). Given the array of water typologies, there is an array of reference site types (e.g., large rivers, small lakes, et cetera). The European Water Framework Directive (EC, 2000) defines references as sites with 'no or minimal anthropogenic stress' (judged by totally or nearly undisturbed conditions for hydromorphological-, physicochemical-, and biological-quality elements, concentrations of synthetic pollutants close to zero or below the limit of detection of advanced analytical techniques, and concentrations of non-synthetic pollutants in the range of natural background levels). In using the reference concept, the sites assigned as reference sites may not be in reference conditions, due to e.g. droughts or floods. Checking the appropriate use of reference sites and conditions is crucial for diagnostic success, as illustrated in this case study.

3.8.2. Methods

Both the EPC-method and the WOE/LR analysis apply the reference concept. In Ohio, the reference sites are commonly identified via best professional judgment (De Zwart et al., 2006), but in this case study a reference condition approach was used, defined by the maximum fish assemblage integrity score (Index of Biotic Integrity, IBI > 46, (Karr, 1981)), related to high variability of biosurvey data for the set of reference sites. This case study explores whether these references may be affected by low-level chemical mixture exposures and other stressors

using the Ohio (2000–2008) data (Kapo et al., 2014), the characteristics of which were detailed in Section 3.5.

3.8.3. Results

The number of sites with IBI-scores >46 was 944, which could be grouped in 25 types of references using RIVPACS-type (River Invertebrate Prediction and Classification System) modelling of species occurrences (Clarke et al., 2003). Bray-Curtis dissimilarities (Bray and Curtis, 1957) between the RIVPACS-predicted species composition and the observed species were calculated for the 944 sites. Predicted and observed fish fauna for the reference sites were considered similar for local Bray-Curtis index values <0.2, chosen to define 'good' reference conditions at the time of sampling, and else dissimilar, defining poorer reference conditions (all at IBI > 46). Thereupon, potential influences of stressor variables on reference conditions was investigated by evaluation of statistical differences between the two groups of sites for each stressor variable, including the mixture toxic pressure proxies for some chemical groups. This analysis showed that various predictors significantly differed between the two subgroups of sites, probably causing slight deviations from (expected) reference conditions (be it with IBI > 46). As an example, poor predictability of reference site fauna regarding urban land use predictors was related to higher average income of people upstream and lower waste water treatment plant numbers, but also to a lower number of permit violations, and a lower percentage of annual effluent loads. Regarding instream habitat, poor predictability of reference fauna appeared significantly related to poor values of all habitat predictors. Surprisingly, a significantly poorer prediction of reference fauna was related to a lower rather than a higher upstream use of pesticides (upstream-caused local toxic pressure). These outcomes imply that sites, despite being judged as being in reference conditions, may be influenced by a multitude of (upstream-caused) stressors/predictors. The net effect could not easily be delineated and assigned to a single cause, as all factors act simultaneously. The results show, however, that reference sites and models should be scrutinized for consistency, accuracy and optional causes for bias invisible in expert judgments (such as chemical mixtures at low concentrations) prior to diagnostic assessment.

3.8.4. Discussion

The use of reference sites or conditions in regulatory assessments of water quality status requires a check on, amongst others, the presence of possible (but hidden) chemical stressors. These may be locally present due to (far) upstream emission sources, and may not be discerned by expert judgment, especially given the phenomenon of 'hidden

mixture impacts' explained in Sections 3.2 and 3.3. The problem of defining, checking and using the reference concept (either as sites, conditions or both) is particularly problematic for some water types, such as large lowlands rivers, when those are almost all subject to various sources of man-made stress; such a problem may be more widespread than only in clearly human-influenced systems, as discussed by Ellis (2015)). In the example, we expected poorer reference conditions in relation to higher metrics for the use of various chemical groups, but we found the opposite ('good' reference conditions at higher parameter values for the use of chemicals). In this case, the diagnostic analysis shown in Section 3.6 may thus be slightly biased by hidden stressor- and impact patterns in the set of reference sites used to delineate impacts and their probable causes. In that study, however, this bias is low as compared to another source of bias that was found, related to weather conditions. That is, reference sites (selected by best professional judgment, and checked for hidden stressor influences as in this example) may experience uncommon or extreme ecological conditions, like extreme droughts and flashiness. Inspection of the fish data for the expert-selected sites showed many instances of low IBI in relation to extreme weather events (unpublished results). It can be concluded that a check on reference site definition, and reference conditions, is key to perform prior to a diagnostic analysis of mixture and multiple-stress impacts.

3.9. Abatement policies and abatement success

3.9.1. Problem definition

Monitoring data sets that cover a long time period can be used for retrospective evaluation of abatement investments. This situation holds for both Ohio data sets (1990–1995 and 2000–2008), especially in view of investments in WWTPs. Ecological conditions and water quality downstream of WWTPs in urban areas were significantly affected in the first period (Dyer and Wang, 2002). Here we explore whether WWTP-investments resulted in improved ecological conditions.

3.9.2. Methods

The "before-after" comparison considered 301 co-located sites involving a WWTP facility, with a categorization in sites downstream or not downstream of the WWTP, as well as in urban or rural sites. The impact metrics were IBI, species richness and fish abnormalities ("DELT": deformities, erosions, lesions and tumours). In addition, the values of the potential stressor metrics were compared.

3.9.3. Results

The observed "before-after" changes are summarized in Table 4. Given unequal numbers of upstream and downstream-, and urban and rural sites, and different numbers of measured parameters, the state-wide fish condition metrics (IBI, DELTS, species richness) showed a moderate but statistically significant increase over the decade, showing a general improvement in fish condition over the 1990–2008 time

period. The improvement in IBI and species richness was however only significant for sites in rivers downstream of WWTP discharges, and fish abnormalities significantly declined for both downstream and non-downstream sites, though the average decline was over 3 times greater for downstream sites. These observations suggest a positive impact of wastewater treatment within the given state-wide trend. IBI change further differed between urban and rural sites, with a higher average IBI improvement for urban sites. The state-wide trend on improvement in species richness was largest (and significant) for rural downstream sites and on improvement of DELT abnormalities the trends were significant in both rural and urban sites. Of the potentially causal factors, total phosphorous showed a significant increase over time on the state level and an increase for downstream sites. The mixture toxic pressure reduced significantly on the state level and for up- and downstream sites, although no significant trend could be delineated for urban downstream sites (as expected) due to few co-located data points with measurements (n = 18). Given these trends, approximately 60% of the improvements (≥2 IBI increase) in downstream sites resulted in the upgraded classification of the site to a higher OEPA biological condition status (OEPA 1988 biocriteria ranges), and approximately 40% of the improvements remained in the same status. The majority of the status upgrades were "good" to "exceptional" status (32% of all improvements) and from "fair" to "good" (17% of all improvements).

3.9.4. Discussion

The data analysis showed a trend of a state-wide improvements of fish community conditions, within which investments in WWTPs appeared to have resulted in an improvement of the biological condition, especially downstream of WWTPs. The improved biological condition coincided, amongst others, with reduced mixture toxic pressures. Especially, as the incidence of abnormalities is considered to be an indicator of impact of industrial and municipal wastewater discharges on fish communities (Ohio EPA, 1988), there may be a link between reduced toxic pressure and reduced abnormalities. However, the data set which could be constructed for the before-after comparisons contained insufficient data to disentangle mixture toxic pressure and other factors showing changes, amongst which total phosphorous. As part of the studied facilities were operating at an advanced treatment level already by 1996, separate information on the relative role of reduced toxic pressure in causing impacts on fish would require both evaluation of earlier (pre-1990) data, as well as a larger data set (see Case Study 1).

3.10. Restoration target

3.10.1. Problem definition

As a final step in Fig. 1, the implementation of abatement strategies should result in the protection or restoration of water bodies, given the understanding of the stressor-response relationships derived from bio-assessments that involve mixture impact assessment. Early attention for the definition of the desired end-status (after management) is key for

Table 4
Mean change in of abiotic and biotic variables based on a before-after comparison for sites up- and downstream from WWTPs in Ohio. "*" = Significant result (p< 0.05), according to Wilcoxon Rank test; N = number of co-located data points).

Sites	Stressor variables				Response variables		
	Δ msPAF (%)	Δ BOD (mg/L)	Δ Total P (mg/L)	Δ TSS (mg/L)	Δ IBI	Δ Species richness	Δ DELTs
All sites (N = 301)	−6.0* (N = 112)	−0.69 (N = 115)	+0.39* (N = 125)	+3.66 (N = 141)	+1.9* (N = 301)	+0.81* (N = 301)	−0.75* (N = 301)
Not downstream of WWTP(s) (N = 98)	−12.6* (N = 29)	−0.59 (N = 34)	−0.12 (N = 40)	−5.36 (N = 44)	+1.2 (N = 98)	+0.05 (N = 98)	−0.27* (N = 98)
Downstream of WWTP(s) (N = 203)	−3.7* (N = 83)	−0.74 (N = 81)	+0.63 (N = 85)	+7.7 (N = 97)	+2.3* (N = 203)	+1.2* (N = 203)	−0.97* (N = 203)
Rural sites downstream (N = 152)	−7.2* (N = 65)	−0.97* (N = 62)	+0.85 (N = 63)	−16.9* (N = 74)	+1.6* (N = 152)	+1.3* (N = 152)	−1.03* (N = 152)
Urban sites downstream (N = 51)	+8.9 (N = 18)	+0.01 (N = 19)	−0.03 (N = 22)	+87.17 (N = 23)	+4.3* (N = 51)	+0.79 (N = 51)	−0.81* (N = 51)

designing sustainable chemical- and water management plans (Zijp et al., 2016). The restoration approaches for impacted sites can be target-oriented, e.g., by considering specific ecosystem services (UN, 2003), and their underlying ecological phenomena. This case study is an early example of such a target-oriented evaluation of mixture effects using monitoring data.

3.10.2. Methods

The WOE/LR analysis for Ohio (2000–2008 data set) discussed previously (Section 3.6) was applied to evaluate stressor-response relationships for various metrics, representing different emphasis on optional restoration endpoints: for the community-level (represented by IBI), for food web base stability (represented by proportion of omnivores), insect food base (proportion of insectivores) and top predators (proportion of carnivores). Conditional inference forest analysis provided the initial assessment of relative variable importance (see Section 3.5), followed by WOE/LR analysis to delineate relative impact probability and local stressor-response hypotheses. Thereupon, it was investigated whether different restoration endpoints relate to different stressor rankings, and thus to potentially different restoration strategies.

3.10.3. Results

The results are schematically summarized in Fig. 8, based on underlying GIS-maps of impact magnitudes and probable causes (like those shown in Fig. 6). In the underlying GIS-maps, the association strength of a stressor variable cannot only be associated to a chosen reference (like in Fig. 6), but also to a defined restoration endpoint of any kind, defined in terms of an assemblage metric. This process asks first for an evaluation of a desired future status (e.g., the reference conditions, or emphasis on an optional ecosystem service), and thereafter an analysis evaluating whether the optional status definitions can be reached, given species-stressor relationships derived in the statistical analyses. The outcomes can be spatially plotted (like. Fig. 6), but for the current case study the outcomes were summarized schematically (Fig. 8). Different stressor variables have different association strengths with the different restoration endpoint definitions, and this shown by higher association strengths for stressors plotted at the figure borders. The Figure shows that different restoration targets would imply handling and reducing different stressor combinations. Overall habitat condition in the Eastern Cornbelt Plains ecoregion was, in the example results, strongly

Fig. 8. Example of visualization of potential stressors in terms of impacts to ecosystem service delivery, given the outcomes of data analyses (sensu the case WOE/LR case study, Section 3.6) applied to the Eastern Cornbelt Plains ecoregion (ER55). Mixture toxic pressure is e.g. related strongest to 'insect community ecosystem services' and 'fish biodiversity and overall ecosystem quality'.

predictive across all trophic-level endpoints, and would require management attention for all restoration targets. Mixture toxic pressure relates more strongly to insect and fish biodiversity, so that this stressor would require more attention to restore desired aspects of ecosystem services, such as fishing or biological community status. If a management goal is the protection of fishing resources, the identification of the stressors most strongly related to top predators may provide the most effective guidance for deriving a management strategy.

3.10.4. Discussion

This case study explores the possible implications of the eco-epidemiological analyses of aquatic monitoring data, by emphasizing restoration target evaluations for water bodies, next to causal analyses. That is, the diagnostic analyses yield an understanding of the stressor-response relationships (including the potential role of pollutant mixtures), and this understanding can also be used in a prognostic context for water system management purposes. This alternative use of the data analyses directly links to the concept of solution-focused risk assessment. Solution-focused risk assessment is a methodological approach to risk assessment that puts alternative management strategies upfront in the assessment process. This idea was generated upon a large evaluation of risk assessment practices, which suggested that the utility of assessment outcomes for risk management could be improved by formulating optional solution scenarios early in the assessment process and by evaluating them (U.S. NAS, 2009; Abt et al., 2010; Finkel, 2011; Zijp et al., 2016).

4. Discussion

4.1. Towards screening-level diagnosis

This paper presents analyses of existing surveillance monitoring data sets that bridge bioassessment and chemical risk assessment, and that build on previous publications on case study data that are all based on the concept of ecoepidemiology (Bro-Rasmussen and Løkke, 1984; Suter et al., 2007). The bridging of bioassessment and (chemical) risk assessment is needed to support reaching the new policy goal of a non-toxic environment and good ecological status, and expands on a current discussion on using Weight of Evidence methods in risk assessment (Ågerstrand and Beronius, 2016). Statistically significant relationships were demonstrated between ecological status metrics and a toxicity metric that quantifies the relative toxicity of ambient mixtures, or specific pollutants or pollutant groups, to aquatic life (mixture toxic pressure). The various methods described here provide insight into site-specific stressors as well as geographic trends in biological impacts, with each approach evaluating the data slightly differently and communicating different aspects of information. Though various eco-epidemiological methods were applied in the case studies, past comparison of the EPC and WOE/LR approaches (Kapo et al., 2008b) indicated general agreement between the methodologies in stressor identification results.

The case studies were not aimed to describe results for the studied monitoring data set per se, but were chosen to as to provide an overview of potentially useful approaches at the interface of chemical risks and water quality assessment, management and policies. That is, the first case study illustrates the need to check the monitoring data for 'signal' (mixture toxic pressure variation) and lack of covariation (which would complicate inferences from final diagnostic results). The next set of case studies (Sections 3.2 to 3.4) illustrates that mono-variate correlations between current ambient pollutant exposures and species abundances may reveal an apparently limited degree of association, but that multiple-stress analyses show that a significant association exists between mixture toxic pressure and taxa responses for a high percentage of taxa (50–86% in the studied cases). Further analyses showed, that the changes associated to increased mixture toxic pressure are – simultaneously – (1) a highly variable response of different taxa
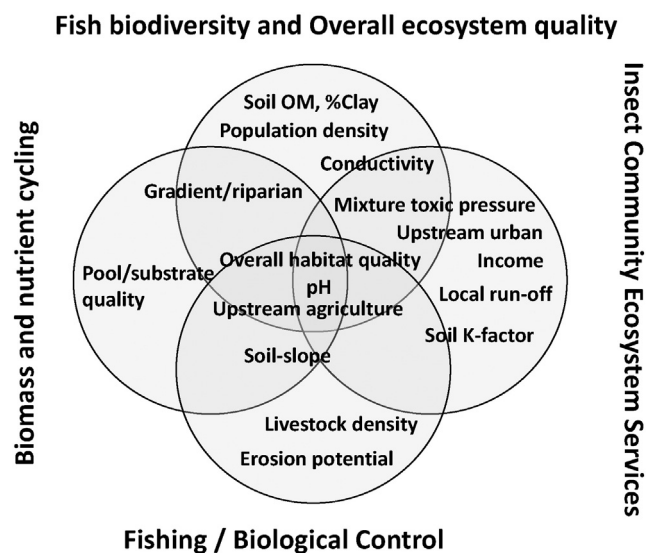
(sensitive, neutral, opportunistic changes were found), which is in line with the species sensitivity distribution phenomenon, and (2) an increased net change of a biodiversity metric such as species number. The final set of case studies showed that the use of the mixture toxic pressure proxy in the eco-epidemiological analysis of monitoring data eventually yields insights into key information for water quality management: (1) the option of ranking sites regarding impact magnitudes, of (2) ranking stressors within sites (to delineate relative contributions of stressor variables to impacts, Sections 3.6 and 3.7), and of diagnosing specific pollutant groups (Section 3.7). Regarding the use of reference sites and -conditions, the case studies further show that it is important to check the reference conditions of the reference sites (at the time samples were taken), for obtaining unbiased diagnostic results (Section 3.8), that the methods can be helpful to explore the success of past abatement strategies (Section 3.9), and that diagnostic approaches can also help to explore alternative management scenarios (Section 3.10). The last step is key, as it requires an explicit choice for water quality management to focus on restoration (to the reference conditions as defined) or to alternative future scenarios, which may involve a focus on particular ecosystem services. The set of approaches helps to bridge the gaps between bioassessments and chemical risk assessments, to support (solution-focused) sustainable chemical- and water policies (Fig. 1). Finally, diagnostic tools are eventually needed to provide evidence for reaching the goal of the non-toxic environment.

While the case studies provide useful examples for attaining the policy goals, they are not ideal. The data were collected following regulatory surveillance monitoring schemes, which are not optimized to represent the sites and hydrologic events for diagnosing multi-stress and chemical impacts. Moreover, as many chemical compounds were not monitored, the results may not provide full insights into toxic exposures. Averaged values (e.g., Table 4) likely distract from the true spatial distribution and magnitudes of chemical impacts. For management purposes, however, the types of approaches that have been illustrated deliver an exploratory (screening-level) yet cost-effective approach to support watershed management and/or the design of targeted higher-tier field studies to better delineate stressor–impact relationships (Kapo et al., 2014).

4.2. The case-study lessons

Despite the aforementioned limitations, the case studies present clear, repetitive and novel findings, with associated literature observations.

4.2.1. Mixture stress varies

Explorations of mixture toxic pressure variability on a landscape scale consistently show high inter-site variability. Although Fig. 2 presents a relatively uniform distribution, other cases of national or regional monitoring data result in a more skewed distribution, with lower numbers of sites with increased toxic pressure levels (see also Fig. 4 in Malaj et al. (2014)).

4.2.2. Mixture stress has impacts

The collated studies show that a high (50–86%) percentage of species of which the abundance is influenced by mixtures. These numbers are unexpectedly high in the context of bioassessment reports, where habitat and land-use and various other stressors are often identified as priority stressors. They are, however, not unexpectedly high when compared to risk assessment expectations. Many ecotoxicological analyses predict and report that chemicals and their mixtures pose threats to aquatic systems, be it due to common modes of action (baseline toxicity) or highly specific modes of action (e.g., pharmaceuticals). A recent risk analysis example at the European scale illustrates this (Malaj et al., 2014). With the case study results, both chemical risk assessment (predictive) and aquatic bioassessment (retrospective) results point to the same conclusion: current chemical exposures cause impacts on

aquatic ecosystems, thereby affecting many species, with impacts varying across sites. It is important to note that ecoregions, sub-regions and sites vary widely regarding impact causes and magnitudes. In most cases, the site-specific evaluation is not only the most important, but also the most informative. Regional to continental (averaged) assessments have high levels of uncertainty, averaging-out the site-specific phenomena.

4.2.3. Prioritizations for management are possible

The skewed distributions of mixture toxic pressures over an area implies that priority management attention to chemicals needs to become focused on sites, e.g. in River Basin Management Planning, given the calibration of the mixture toxic pressure proxy on observed ecological impacts. Neither chemicals nor other stressors should be considered separate from each other when considering restoration. At some sites, mixtures may be dominant stressors, but at many sites they may act as next-limiting factor for restoration. In addition to ranking sites and stressors, it is likely that river-basin specific pollutants or pollutant groups can be identified (Fig. 7) and prioritized for abatement. The bioassessment results on the dominance of a relatively small selection of chemicals in a local mixture are similar to those of controlled experiments. Such experiments have shown that mixture impacts can often be attributed to few chemicals only (Backhaus and Karlsson, 2014). This offers practical perspectives for the prioritization of abatement strategies, contrasting to the impractical consequence of assuming that an assessment should always cover the infinite number of thinkable chemical-species interactions.

4.2.4. Assessment methods bridge disciplines and regulations and yield summary insights

The disciplines and policies shown in Fig. 1 apply different methods and they yield different outcomes. However, chemical and water assessments can all consistently apply a single model (SSDs) for assessing chemical problems. This option was first proposed by Van Straalen and Denneman (1989). SSDs are the common basis to derive safe chemical criteria, as well as to derive the impact proxy for exposed ecosystems. This common basis is valuable for policy adoption of the methods in the different contexts as well as for risk and impact communication. Rather than presenting a high number of criteria-exceedance observations (e.g., Vijver et al. (2008)), the mixture toxic pressure can be summarized as a convenient summary overview of possibly mixture toxicity-impacted areas (De Zwart, 2005; De Snoo and Vijver, 2012). Areas where the mixture toxic pressure is negligible, despite a large number of possible chemicals present, can be subject to maintained protection, and highly 'coloured' (high toxic pressure) sites can be subject to more intensive chemical emission reduction and restoration. Next to landscape and the riverscape maps in GIS, chemical concentration maps can show 'chemoscapes' of chemicals, but the mixture toxic pressure map can provide spatial insights in the ecological 'impactscape'. These ideas are in line with working hypotheses that are being developed to explore whether chemical impacts on the landscape scale could be related to major drivers of chemical emissions, with typical emission 'fingerprints' and impacts such as those of agricultural land use, or the built environment, or upstream combinations of emissions under a mixed land-use scenario. Eco-epidemiological results can also support post-market analyses aimed to ascertain that the use of chemicals that are allowed on the market via current risk assessment protocols is indeed safe, either for individual compounds or for mixtures.

4.2.5. Assessment methods can explore the relevance of emerging chemicals

Currently, most chemicals found in surface waters cannot be judged as Water Quality Criteria (Sjerps et al., 2016), while chemical emissions expectedly rise (UNEP, 2013). In the case studies, part of the toxic pressures in water bodies were derived from modelling, e.g., with iSTREEM®. This modelling, in conjunction with growing numbers of

ecotoxicity data allows for a further step. That is: the generation of predicted impact maps for many chemicals. To this end, as an example, U.S. EPA's aquatic database contains data from 365,981 tests with 470,343 associated result records, covering 11,016 chemicals and 11,799 tested species, (http://cfpub.epa.gov/ecotox/help.cfm?sub=recentadditions_1215, consulted 24 Feb. 2016). Local toxic pressures were derived for emerging chemicals from emission-based predicted environmental concentrations in combination with newly derived SSD models, beyond the limited number of SSDs used in chemical policies. Other disciplines have indeed developed additional SSD_{EC50}'s, e.g. in the context of the comparative Life Cycle Assessment of chemicals (Harbers et al., 2006; Rosenbaum et al., 2008), and making use of patterns across SSD's (De Zwart, 2002; Hendriks et al., 2013). Example studies have developed this idea further in the format of chemical footprints – a judgment of the ratio of net ecotoxic impact potential of emitted chemicals vis a vis the water volume of a region (Bjørn et al., 2014; Zijp et al., 2014). This may be the basis for net chemical impact judgment of regions, whereby a completely novel regulatory judgment approach might be adopted such as 'no export of net mixture impacts to neighbouring regions' (see e.g. Hoornweg et al. (2016)).

4.3. Versatility of study endpoints

Bioassessments can focus on a multitude of measurement endpoints (in principle from 'omics' data, via fish deformities (DELTS), species counts and trait frequencies (Pilière et al., 2016) up till trait frequencies and functional metrics, see Table 2). These endpoints vary dramatically in their importance to ecological impacts and in our ability to measure them in an efficient and effective manner. Odum (1992) suggested to focus on populations for the evaluation of impacts, but the eco-epidemiological approaches can be used for all data types. It is critical to link "impacts" to those that are of regulatory importance (e.g., population, community and functional), although some sensitive methods may be developed as early-warning approach. Our case studies can thus be seen as a few examples of a quickly developing set of approaches.

4.4. The role of species sensitivity distributions

In all case studies, expected impacts have so far been quantified based on the use of SSDs. We note that this model is subject to scholar debate since its invention (Forbes and Forbes, 1993; Hopkin, 1993; Forbes and Calow, 2002; Posthuma et al., 2002) up till now, be it with a focus in that debate on the utility and accuracy of SSDs to estimate protective criteria. In that discussion context, the issue at stake is an honest 'level playing field' evaluation of chemical hazards across chemicals prior to their formal registration and use. In the case studies, however, the SSD model is used differently. SSDs are here used to quantify a proxy for which it logically holds that increasing proxy values are an expression of an increased potential of an environmental sample to cause ecological harm. The proxy is – in fact – a recalculation of predicted or observed concentrations into hazard units, 'trained' by laboratory observations on effects in tested species. Thus, it quantifies a characteristic of an environmental sample, in that it predicts which fraction of tested species would exhibit effects when reared in a polluted sample. It reflects the environment's Hazard Potential (as the counterpart of the concept of Hazardous Concentrations used in the derivation of protective criteria). Just like any other environmental parameter (e.g., acidity or salinity), ecological studies can establish the relationship between parameter values and ecological phenomena. This has been done repeatedly for msPAF, for highly different studies (example in Section 3.4), like a calibration to the degree of genetic adaptation in soil arthropods (Posthuma, 1992), and to responses in aquatic ecosystems (Posthuma and De Zwart, 2006; Posthuma and De Zwart, 2012; Posthuma et al., 2016). Despite differences amongst studies, it appeared that within studies (*ceteris paribus*) increased proxy values appear to

relate to increased ecological impacts. The model predictions are thereby of degree, not kind.

The statistical aphorism of George Box "*All models are wrong, but some are useful*" (Box, 1979) applies here, and – following Box again – the question of interest is: "*Is the model illuminating and useful?*". The linking of chemical risk assessment to bioassessment can be supported by SSDs due to the observed meaning of the toxic pressure proxy: SSD results shed light ("illuminate") ecological effects of pollutant exposures (given the case study results, especially in Sections 3.2 to 3.4, but also 3.6 and 3.7). This focus on usefulness of the SSD-model for decision making is different from the set of reasons via which the SSD-model is criticised or rejected, e.g., for its lack of ecological input, via unjustified reification (the fallacy of misplaced concreteness, interpreting the model and its output as a fact rather than as a proxy which is relevant for rankings that are useful for decision making), or via arguments and prescribed guidances that are solely applicable to the context of environmental protection and safe criteria setting. Whether or not SSD-outputs are the best way to provide decision support on environmental pollution, rejection of the model for the above reasons may result in alternative model approaches that are conceptually weaker (e.g., a criteria-based judgment) and that lack the versatility to judge the myriad of chemical-biota interactions. Our results show that increased mixture toxic pressure relates to increased impacts – which is in general supportive of the use of SSDs as lower-tier model in practice-oriented risk assessments and decision support, especially regarding prioritization (of sites, of stressors, and of compounds within mixtures). Although SSDs can thus be supportive for decision making, they are in our view certainly not a panacea for all environmental management problems.

4.5. Towards solutions

Bioassessments that include toxic chemicals can improve the understanding of the causes of impacts in aquatic ecosystems at various spatial scales, involving regional issues like land use, catchment-scale phenomena like hydrology and proximity to economical activities, and in-stream parameters for water quality and ecology. The case studies show in which ways the understanding of polluted water systems can increase, in the context of the solution-focused assessment paradigm (Fig. 1). Although our analyses illustrate directions taken in diagnosing chemical mixture impacts under field conditions, there are still many improvements needed, not the least in the quality and evaluation of the unique sets of monitoring data being collected for water management (Hering et al., 2010), and further innovative technical approaches. An intriguing option of the latter would be to fill data gaps on various stressors via application of SSD-type modelling for stressors other than chemicals, like engineered nanomaterials (Garner et al., 2015) or radiation (Garnier-Laplace et al., 2006). And also to study interactive effects of e.g. climate change and endocrine disrupting chemicals (Brown et al., 2015).

Irrespective of further developments, our findings underscore the large-scale and highly diverse phenomenon of chemicals that are actually co-shaping species assemblages in aquatic ecosystems. We have witnessed over the last decades that observed effects of chemicals have diminished since the start of environmental policies, but our diagnosis of current impact levels (supported by similar risk analyses, (Malaj et al., 2014)) shows that more efforts are needed. Chemical regulations aim to reduce hazardous emissions, and water policies aim at water quality improvements. But both signal the presence of a pollution problem, suggesting that further actions and a possible overarching reframing of the problem may be worthwhile. Amongst others, (Geiser, 2015) undertook such an effort recently. The holistic reframe, referred to as a 'new chemical economy' (Schwarzman, 2015), describes a suite of opportunities that start with safe chemical design, and evaluate abatement opportunities from many angles. These angles are of various kinds: 'hardware' (chemicals, environment, exposures, impacts, geography, hydrology, spatio-temporal aspects, etc.), 'software' (regulations

and management tools, but also teaching) and 'mindware' (a deliberate evaluation of risks and impacts and alternative management options aimed at sustainable management), and operate across disciplines and current frameworks (Ministerie I&M, 2014). Regarding 'hardware', the impact-focused approach can be further developed to support choices for the most sustainable chemical-use scenarios given societal developments like the growth in green chemistry, the biobased economy and the re-use of materials in the circular economy. Regarding 'software', the case studies substantiate how the regulatory link between chemical and water assessments could look like (as present in REACH via Article 2.4 and the Water Framework Directive in its concept of Good Chemical Quality), and they cover the goals of a non-toxic environment and good ecological status simultaneously. Regarding 'mindware', a solution-focused orientation used next to causal analysis supports finding sensible but as yet unexpected abatement options (Brack et al., 2015; Zijp et al., 2016), which has (amongst others) already resulted in an analysis towards safe spatial planning of chemical emission sources vis a vis sensitive functions (Coppens et al., 2015). While we acknowledge the major improvements that have been gained, are gained, and will be gained by continued application of existing chemical- and water policies, we wholeheartedly support the concluding remark of the risk study of Malaj et al. (2014) based on our independently collected impact-based results: "[…] *more effort is necessary to integrate and advance these regulations toward the reduction of toxic pollution. Our study suggests that a paradigm change in chemical regulation and management is required to achieve a holistic approach, which assesses the toxic pressure as a whole rather than from individual chemicals, and complements specific case studies by large-scale analyses.*"

Acknowledgments

References

Abt, E., Rodricks, J.V., Levy, J.I., Zeise, L., Burke, T.A., 2010. Science and decisions: advancing risk assessment. Risk Anal. 30, 1028–1036.

Ågerstrand, M., Beronius, A., 2016. Weight of evidence evaluation and systematic review in EU chemical risk assessment: foundation is laid but guidance is needed. Environ. Int. 92–93, 590–596.

Allan, J.D., 2004. Landscapes and riverscapes: the influence of land use on stream ecosystems. Annu. Rev. Ecol. Evol. Syst. 35, 257–284.

Backhaus, T., Karlsson, M., 2014. Screening level mixture risk assessment of pharmaceuticals in STP effluents. Water Res. 49, 157–165.

Baird, D.J., Rubach, M.N., Van den Brink, P.J., 2008. Trait-based ecological risk assessment (TERA): the new frontier? Integr. Environ. Assess. Manag. 4, 2–3.

Barbour, M.T., Gerritsen, J., Snyder, B.D., Stribling, J.B., 1999. Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish. second ed. .

Bellman, R.E., 1961. Adaptive Control Processes: A Guided Tour. Princeton University Press, Princeton, NJ, USA.

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

Box, G.E.P., 1979. Robustness in the strategy of scientific model building. Technical Summary Report #1954. Mathematics Resea rch Center, University of Wisconsin, Madison.

Brack, W., Altenburger, R., Schüürmann, G., Krauss, M., López Herráez, D., van Gils, J., et al., 2015. The SOLUTIONS project: challenges and responses for present and future emerging pollutants in land and water resources management. Sci. Total Environ. 503–504, 22–31.

Bray, J.R., Curtis, J.T., 1957. An ordination of upland forest communities of southern Wisconsin. Ecol. Monogr. 27, 325–349.

Bro-Rasmussen, F., Løkke, H., 1984. Ecoepidemiology — a casuistic discipline describing ecological disturbances and damages in relation to their specific causes; exemplified by chlorinated phenols and chlorophenoxy acids. Regul. Toxicol. Pharmacol. 4, 391–399.

Brown, A.R., Owen, S.F., Peters, J., Zhang, Y., Soffker, M., Paull, G.C., et al., 2015. Climate change and pollution speed declines in zebrafish populations. Proc. Natl. Acad. Sci.

Burkhead, N.M., 2012. Extinction rates in north American freshwater fishes, 1900-2010. Bioscience 62, 798–808.

Burton, G.A., De Zwart, D., Diamond, J., Dyer, S., Kapo, K.E., Liess, M., et al., 2012. Making ecosystem reality checks the status quo. Environ. Toxicol. Chem. 31, 459–468.

Carpenter, S.R., Stanley, E.H., Vander Zanden, M.J., 2011. State of the world's freshwater ecosystems: physical, chemical, and biological changes. Annu. Rev. Environ. Resour. 36, 75–99.

Carvalho, R.N., Arukwe, A., Ait-Aissa, S., Bado-Nilles, A., Balzamo, S., Baun, A., et al., 2014. Mixtures of chemical pollutants at European legislation safety concentrations: how safe are they? Toxicol. Sci. 141, 218–233.

Cervantes-Yoshida, K., Leidy, R.A., SM, C., 2015. Contemporary land change alters fish communities in a San Francisco Bay Watershed, California, U.S.A. PLoS One 10, e0141707.

Chiogna, G., Majone, B., Cano Paoli, K., Diamantini, E., Stella, E., Mallucci, S., et al., 2016. A review of hydrological and chemical stressors in the Adige catchment and its ecological status. Sci. Total Environ. 540, 429–443.

Clarke, R.T., Wright, J.F., Furse, M.T., 2003. RIVPACS models for predicting the expected macroinvertebrate fauna and assessing the ecological quality of rivers. Ecol. Model. 160, 219–233.

Coppens, L.J.C., van Gils, J.A.G., ter Laak, T.L., Raterman, B.W., van Wezel, A.P., 2015. Towards spatially smart abatement of human pharmaceuticals in surface waters: defining impact of sewage treatment plants on susceptible functions. Water Res. 81, 356–365.

De Snoo, G.R., Vijver, M., 2012. Bestrijdingsmiddelen en water kwaliteit [Plant protection products and water quality] Leiden, the Netherlands .

De Zwart, D., 2002. Observed regularities in SSDs for aquatic species. In: Posthuma, L., Suter II, G.W., Traas, T.P. (Eds.), Species Sensitivity Distributions in Ecotoxicology. Lewis Publishers, Boca Raton, FL, USA, pp. 133–154.

De Zwart, D., 2005. Ecological effects of pesticide use in The Netherlands: modeled and observed effects in the field ditch. Integr. Environ. Assess. Manag. 1, 123–134.

De Zwart, D., Dyer, S.D., Posthuma, L, Hawkins, C.P., 2006. Predictive models attribute effects on fish assemblages to toxicity and habitat alteration. Ecol. Appl. 16, 1295–1310.

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

De Zwart, D., Posthuma, L., Gevrey, M., Von Der Ohe, P.C., De Deckere, E., 2009. Diagnosis of ecosystem impairment in a multiple-stress context-how to formulate effective river basin management plans. Integr. Environ. Assess. Manag. 5, 38–49.

De Zwart, D., Posthuma, L., Pemberton, E., 2008. Understanding Ecological Impacts in Rivers in England and Wales and Identifying Their Possible Causes. Part 1 — The Effect and Probable Cause (EPC) Method. Environment Agency.

Diamond, J., Munkittrick, K., Kapo, K.E., Flippin, J., 2015. A framework for screening sites at risk from contaminants of emerging concern. Environ. Toxicol. Chem. 34, 2671–2681.

Dyer, S.D., Wang, X., 2002. A comparison of stream biological responses to discharge from wastewater treatment plants in high and low population density areas. Environ. Toxicol. Chem. 21 (5), 1065–1075.

Dyer, S.D., White-Hull, C.D., Carr, G.J., Smith, E.P., Wang, X., 2000. Bottom-up and top-down approaches to assess multiple stressors over large geographic areas. Environ. Toxicol. Chem. 19, 1066–1075.

EC, 2000. Directive 2000/60/EC of the European parliament and of the council of 23 October 2000 establishing a framework for community action in the field of water policy. Off. J. Eur. Communities 327, 1–72.

EC, 2003. Proposal for a Regulation of the European Parliament and of the Council Concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (Reach), Establishing a European Chemicals Agency and Amending Directive 1999/45/EC and Regulation (EC) on Persistent Organic Pollutants. European Commission, Brussels, Belgium.

EC, 2012. The Combination Effects of Chemicals. Chemical Mixtures. European Commissions, Brussels, Belgium.

EC, 2014. Living Well, Within the Limits of Our Planet. General Union Environment Action Programme. 2020.

EEA, 2012. European Waters — Assessment of Status and Pressures.

Ellis, E.C., 2015. Ecology in an anthropogenic biosphere. Ecol. Monogr. 85, 287–331.

ETC/ICM, 2012. Ecological and Chemical Status and Pressures of European Waters. European Topic Centre on Inland, Coastal and Marine Waters, Prague, Czech Republic.

Communities, European]–>E., 2003. Technical Guidance Document on Risk Assessment in Support of: Commission Directive 93/67/EEC on Risk Assessment for New Notified Substances, Commission Regulation (EC) No 1488/94 on Risk Assessment for Existing Substances, and Directive 98/8/EC of the European Parliament and of the Council Concerning the Placing of Biocidal Products on the Market Part II .

Fedorenkova, A., Vonk, J.A., Lenders, H.J.R., Creemers, R.C.M., Breure, A.M., Hendriks, A.J., 2012. Ranking ecological risks of multiple chemical stressors on amphibians. Environ. Toxicol. Chem. 31, 1416–1421.

Finkel, A.M., 2011. Solution-focused risk assessment: a proposal for the fusion of environmental analysis and action. Hum. Ecol. Risk. Assess. 17, 754–787.

Forbes, T.L., Forbes, V.E., 1993. A critique of the use of distribution-based extrapolation models in ecotoxicology. Funct. Ecol. 7, 249–254.

Forbes, V.E., Calow, P., 2002. Species sensitivity distributions revisited: a critical appraisal. Hum. Ecol. Risk. Assess. 8, 473–492.

Forbes, V.E., Calow, P., Sibly, R.M., 2008. The extrapolation problem and how population modeling can help. Environ. Toxicol. Chem. 27, 1987–1994.

Friberg, N., Bonada, N., Bradley, D.C., Dunbar, M.J., Edwards, F.K., Grey, J., et al., 2011. Biomonitoring of human impacts in freshwater ecosystems. The good, the bad and the ugly. Adv. Ecol. Res. 44, 1–68.

Garner, K.L., Suh, S., Lenihan, H.S., AA, K., 2015. Species sensitivity distributions for engineered nanomaterials. Environ. Sci. Technol. 49 (9), 5753–5759.

Garnier-Laplace, J., Della-Vedova, C., Gilbin, R., Copplestone, D., Hingston, J., Ciffroy, P., 2006. First derivation of predicted-no-effect values for freshwater and terrestrial ecosystems exposed to radioactive substances. Environ. Sci. Technol. 40, 6498–6505.

Geiser, K., 2015. Chemicals Without Harm. Policies for a Sustainable World. MIT Press (456 pp).

Hågvar, S., Abrahamsen, G., 1990. Microarthropoda and enchytraeidae (Oligochaeta) in naturally lead-contaminated soil, a gradient study. Environ. Entomol. 19, 1263–1277.

Harbers, J.V., Huijbregts, M.A.J., Posthuma, L., Van de Meent, D., 2006. Estimating the impact of high-production-volume chemicals on remote ecosystems by toxic pressure calculation. Environ. Sci. Technol. 40, 1573–1580.

Hendriks, A.J., 2013. How to deal with 100,000+ substances, sites, and species: overarching principles in environmental risk assessment. Environ. Sci. Technol. 47, 3546–3547.

Hendriks, A.J., Awkerman, J.A., de Zwart, D., Huijbregts, M.A.J., 2013. Sensitivity of species to chemicals: dose-response characteristics for various test types (LC50, LR50 and LD50) and modes of action. Ecotoxicol. Environ. Saf. 97, 10–16.

Hering, D., Borja, A., Carstensen, J., Carvalho, L., Elliott, M., Feld, C.K., et al., 2010. The European water framework directive at the age of 10: a critical review of the achievements with recommendations for the future. Sci. Total Environ. 408, 4007–4019.

Herman, M.R., Nejadhashemi, A.P., 2015. A review of macroinvertebrate- and fish-based stream health indices. Ecohydrol. Hydrobiol. 15, 53–67.

Heugens, E.H.W., Hendriks, A.J., Dekker, T., Van Straalen, N.M., Admiraal, W., 2001. A review of the effects of multiple stressors on aquatic organisms and analysis of uncertainty factors for use in risk assessment. Crit. Rev. Toxicol. 31, 247–284.

Hoekstra, A.Y., Mekonnen, M.M., 2012. The water footprint of humanity. Proc. Natl. Acad. Sci. 109, 3232–3237.

Hoekstra, A.Y., Wiedmann, T.O., 2014. Humanity's unsustainable environmental footprint. Science 344, 1114–1117.

Hoornweg, D., Hosseini, M., Kennedy, C., Behdadi, A., 2016. An urban approach to planetary boundaries. Ambio 1–14.

Hopkin, S.P., 1993. Ecological implications of 95% protection levels for metals in soil. Oikos 66, 137–141.

Kapo, K.E., Burton, G.A.J., De Zwart, D., Posthuma, L., SD, D., 2008a. Quantitative lines of evidence for screening-level diagnostic assessment of regional fish community impacts: a comparison of spatial database evaluation methods. Environ. Sci. Technol. 42, 9412–9418.

Kapo, K.E., Burton Jr., G.A., 2006. A geographic information systems-based, weights-of-evidence approach for diagnosing aquatic ecosystem impairment. Environ. Toxicol. Chem. 25, 2237–2249.

Kapo, K.E., Burton Jr., G.A., Zwart, D., Posthuma, L., Dyer, S.D., 2008b. Quantitative lines of evidence for screening-level diagnostic assessment of regional fish community impacts: a comparison of spatial database evaluation methods. Environ. Sci. Technol. 42, 9412–9418.

Kapo, K.E., Holmes, C.M., Dyer, S.D., de Zwart, D., Posthuma, L., 2014. Developing a foundation for eco-epidemiological assessment of aquatic ecological status over large geographic regions utilizing existing data resources and models. Environ. Toxicol. Chem. 33, 1665–1677.

Karr, J.R., 1981. Assessment of biotic integrity using fish communities. Fisheries 6, 21–27.

Kline, R.B., 1998. Principles and Practice of Structural Equation Modeling. Guilford Press, New York, NY, USA.

Kuzmanović, M., López-Doval, J.C., De Castro-Català, N., Guasch, H., Petrović, M., Muñoz, I., et al., 2016. Ecotoxicological risk assessment of chemical pollution in four Iberian river basins and its relationship with the aquatic macroinvertebrate community status. Sci. Total Environ. 540, 324–333.

Låg, J., Hvatum, O.O., Bolviken, B., 1970. An occurence of naturally lead-poisoned soil at Kastad near Gjovik, Norway. Nor. Geol. Unders. 266, 141–157.

Lenat, D.R., Resh, V.H., 2001. Taxonomy and stream ecology — the benefits of genus- and species-level identifications. American Benthological Society]–>J. N. Am. Benthol. Soc. 20, 287–298.

Leps, M., Tonkin, J.D., Dahm, V., Haase, P., Sundermann, A., 2015. Disentangling environmental drivers of benthic invertebrate assemblages: the role of spatial scale and riverscape heterogeneity in a multiple stressor environment. Sci. Total Environ. 536, 546–556.

Li, F., Chung, N., Bae, M.J., Kwon, Y.S., Park, Y.S., 2012. Relationships between stream macroinvertebrates and environmental variables at multiple spatial scales. Freshw. Biol. 57, 2107–2124.

Malaj, E., von der Ohe, P.C., Grote, M., Usseglio-Polatera, P., et al., 2014. Organic chemicals jeopardize the health of freshwater ecosystems on the continental scale. Kühne R. Proceedings of the National Academy of Sciences. Mondy CP.

Metcalfe, J.L., 1989. Biological water quality assessment of running waters based on macroinvertebrate communities: history and present status in Europe. Environ. Pollut. 60, 101–139.

Ministerie I&M, 2014. Bewust Omgaan met Veiligheid: Rode Draden. Een proeve van een IenM-breed afwegingskader veiligheid. Ministerie van Infrastructuur en Milieu, Den Haag.

Murray, K.E., Thomas, S.M., Bodour, A.A., 2010. Prioritizing research for trace pollutants and emerging contaminants in the freshwater environment. Environ. Pollut. 158, 3462–3471.

Mykrä, H., Heino, J., Muotka, T., 2007. Scale-related patterns in the spatial and environmental components of stream macroinvertebrate assemblage variation. Glob. Ecol. Biogeogr. 16, 149–159.

Norris, R.H., Linke, S., Prosser, I.A.N., Young, W.J., Liston, P., Bauer, N., et al., 2007. Very-broad-scale assessment of human impacts on river condition. Freshw. Biol. 52, 959–976.

Norton, S.B., G.W.I., S., SM, C., 2014. Ecological Causal Assessment. CRC Press, Boca Raton, FL. USA.

O'Brian, R.M., 2007. A caution regarding rules of thumb for variance inflation factors. Qual. Quant. 41, 673–690.

Odum, E.P., 1992. Great ideas in ecology for the 1990's. Bioscience 42, 532–545.

Ohio EPA, 1988. Biological Criteria for the Protection of Aquatic Life (Volume II). Ecological Assessment Section, Ohio EPA, Columbus OH.

Pal, A., Gin, K.Y.-H., Lin, A.Y.-C., Reinhard, M., 2010. Impacts of emerging organic contaminants on freshwater resources: review of recent occurrences, sources, fate and effects. Sci. Total Environ. 408, 6062–6069.

Pilière, A., Verberk, W., Gräwe, M., Breure, A.M., Dyer, S.D., Posthuma, L., et al., 2016. On the importance of trait interrelationships for understanding species environmental responses of stream invertebrates. Freshw. Biol. 61, 181–194.

Pistocchi, A., 2014. Perspectives: The Challenge of Cumulative Impacts and Planetary Boundaries. GIS Based Chemical Fate Modeling. John Wiley & Sons, Inc., pp. 475–479.

Posthuma, L., 1992. Genetic Ecology of Metal Tolerance in Collembola. Vrije Universiteit.

Posthuma, L., 1997. Effects of toxicants on population and community parameters in field conditions, and their potential use in the validation of risk assessment methods. In: Straalen, N.M., Løkke, H. (Eds.), Ecological Risk Assessment of Contaminants in Soil. Springer US, Boston, MA, pp. 85–123.

Posthuma, L., De Zwart, D., 2006. Predicted effects of toxicant mixtures are confirmed by changes in fish species assemblages in Ohio, USA, rivers. Environ. Toxicol. Chem. 25, 1094–1105.

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., 2014. Species Sensitivity Distributions. Encyclopedia of Toxicology. third ed. 4. Elsevier Inc., Academic Press, pp. 363–368.

Posthuma, L., De Zwart, D., Keijzers, R., Postma, J., 2016. Watersysteemanalyse met de Ecologische Sleutelfactor Toxiciteit. Deel 2. Calibratie: toxische druk en ecologische effecten op macrofauna. STOWA, Amersfoort, the Netherlands.

Posthuma, L., Schouten, A.J., Van Beelen, P., Rutgers, M., 2001. Forecasting effects of toxicants at the community level. Four case studies comparing observed community effects of zinc with forecasts from a generic ecotoxicological risk assessment method. In: Rainbow, P.S., Hopkin, S.P., Crane, M. (Eds.), Forecasting the Environmental Fate and Effects of Chemicals. John Wiley, Chichester, UK, pp. 151–175.

Posthuma, L., Suter, G.W.I., Traas, T.P., 2002. Species Sensitivity Distributions in Ecotoxicology. CRC-Press, Boca Raton, FL, U.S.A.

Rosenbaum, R.K., Bachmann, T.M., Gold, L.S., Huijbregts, M.A.J., Jolliet, O., Juraske, R., et al., 2008. USEtox — the UNEP-SETAC toxicity model: recommended characterisation factors for human toxicity and freshwater ecotoxicity in life cycle impact assessment. Int. J. Life Cycle Assess. 13, 532–546.

Sabater, S., Barceló, D., De Castro-Català, N., Ginebreda, A., Kuzmanovic, M., Petrovic, M., et al., 2016. Shared effects of organic microcontaminants and environmental stressors on biofilms and invertebrates in impaired rivers. Environ. Pollut. 210, 303–314.

Schwarzenbach, R.P., Escher, B.I., Fenner, K., Hofstetter, T.B., Johnson, C.A., von Gunten, U., et al., 2006. The challenge of micropollutants in aquatic systems. Science 313, 1072–1077.

Schwarzman, M.R., 2015. A new chemical economy. Science 349, 1175.

Sjerps, R.M.A., Vughs, D., Van Leerdam, J.A., Ter Laak, T.L., Van Wezel, A.P., 2016. Data-driven prioritization of chemicals for various water types using suspect screening LC-HRMS. Water Res. 93, 254–264.

Solomon, K.R., Takacs, P., 2002. Probabilistic risk assessment using species sensitivity distributions. In: Posthuma, L., Suter II, G.W., Traas, T.P. (Eds.), Species Sensitivity Distributions in Ecotoxicology. Lewis Publishers, Boca Raton, FL, USA, pp. 285–313.

Sponseller, R.A., Benfield, E.F., Valett, H.M., 2001. Relationships between land use, spatial scale and stream macroinvertebrate communities. Freshw. Biol. 46, 1409–1424.

Stehle, S., Schulz, R., 2015. Agricultural insecticides threaten surface waters at the global scale. Proc. Nat. Acad. Sci.

Stephan, C.E., Mount, D.I., Hansen, D.J., Gentile, J.H., Chapman, G.A., Brungs, W.A., 1985. Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses. US EPA ORD ERL, Duluth MN, pp. 1–97.

Stoddard, J.L., Larsen, D.P., Hawkins, C.P., Johnson, R.K., Norris, R.H., 2006. Setting expectations for the ecological condition of streams: the concept of reference condition. Ecol. Appl. 16, 1267–1276.

Strobl, C., Boulesteix, A.N., Zeileis, A., Hothorn, T., 2007. Bias in random forest variable importance measures: illustrations, sources and a solution. BMC Bioinf. 8.

Sundermann, A., Leps, M., Leisner, S., Haase, P., 2015. Taxon-specific physico-chemical change points for stream benthic invertebrates. Ecol. Indic. 57, 314–323.

Suter, G.W., Cormier, S., Norton, S.B., 2007. Ecoepidemiology and causal analysis. In: Suter, G.W. (Ed.), Ecological Risk Assessment, second ed. CRC Press, Boca Raton, FL, USA, pp. 39–68.

NAS, U.S., 2009. Science and Decisions: Advancing Risk Assessment. The National Academies Press.

Senate, U.S., 2014. Water Resources Reform and Development Act of 2014.

Tyler, G., 1984. The impact of heavy metal pollution on forests: a case study of Gusum, Sweden. Ambio 13, 18–24.

UN, 2003. Millenium Ecosystem Assessment. Ecosystems and Human Well-Being. A Framework for Assessment. Island Press, Washington, DC, USA.

UNEP, 2013. Global Chemicals Outlook. Towards Sound Management of Chemicals. United Nations Environment Programme, Geneva, Switzerland.

USEPA, 2009. The National Study of Chemical Residues in Lake Fish Tissue. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

Van Leeuwen, C.J., Vermeire, T.G., 2007. Risk Assessment of Chemicals. An Introduction. second ed. Springer, Dordrecht, the Netherlands.

Van Straalen, N.M., Denneman, C.A.J., 1989. Ecotoxicological evaluation of soil quality criteria. Ecotoxicol. Environ. Saf. 18, 241–251.

Vander Laan, J.J., Hawkins, C.P., Olson, J.R., Hill, R.A., 2013. Linking land use, in-stream stressors, and biological condition to infer causes of regional ecological impairment in streams. Freshw. Sci. 32, 801–820.

Vijver, M.G., Van't Zelfde, M., Tamis, W.L.M., Musters, K.J.M., De Snoo, G.R., 2008. Spatial and temporal analysis of pesticides concentrations in surface water: pesticides atlas. J. Environ. Sci. Health B 43, 665–674.

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

Zijp, M.C., Posthuma, L., Wintersen, A., Devilee, J., Swartjes, F.A., 2016. Definition and use of solution-focused sustainability assessment: a novel approach to generate, explore and decide on sustainable solutions for wicked problems. Environ. Int. 91, 319–331.