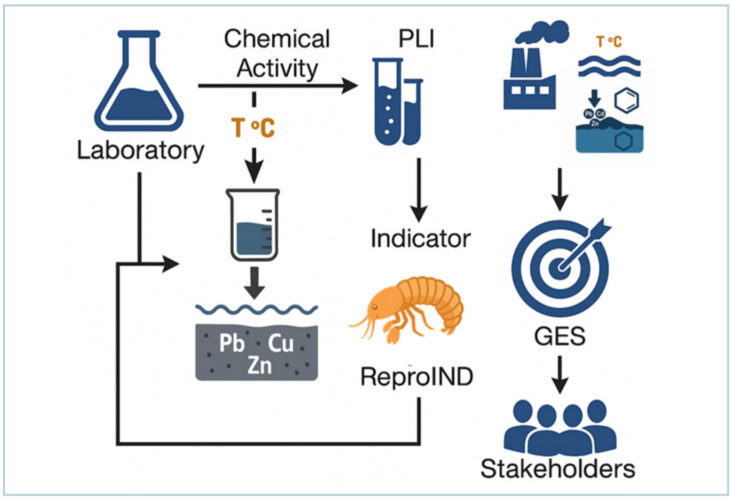
MILJÖ | NATURVÅRDSVERKET FORSKNING

Cumulative effects: Incorporating Nonchemical StrEssors into ReproIND indicator

Project INSERT: integration of monitoring data, field and experimental research, and indicator development

Elena Gorokhova, Sebastian Abel, Gastón Alurralde, Ann-Kristin Eriksson Wiklund, Anna Sobek, Yves Saladin, Sophie Steigerwald and Brita Sundelin



RAPPORT 7208 | DECEMBER 2025

Cumulative effects: Incorporating Nonchemical StrEssors into ReproIND indicator

Project INSERT: integration of monitoring data, field and experimental research, and indicator development

by Elena Gorokhova, Sebastian Abel, Gastón Alurralde,
Ann-Kristin Eriksson Wiklund, Anna Sobek, Yves Saladin,
Sophie Steigerwald and Brita Sundelin

SWEDISH ENVIRONMENTAL
PROTECTION AGENCY

The Swedish Environmental Protection Agency
Phone: + 46 (0)10-698 10 00
E-mail: registrator@naturvardsverket.se
Address: Naturvårdsverket, SE-106 48 Stockholm, Sweden
Internet: www.naturvardsverket.se

ISBN 978-91-620-7208-7
ISSN 0282-7298

Print: Arkitektkopia AB, Bromma 2025
Cover illustration: Elena Gorokhova

Preface

This report Cumulative effects: Incorporating Nonchemical StrEssors into ReproIND indicator presents the results from one of seven projects funded within the research call “Cumulative effects on the environment”. With this research call the Swedish Environmental Protection Agency and the Swedish Agency for Marine and Water Management aimed to advance development of methods and analytical tools for qualitatively and quantitatively assessing the burden of cumulative effects on ecosystems in environmental assessments.

The project has been financed by the Swedish Environmental Protection Agency’s environmental research fund.

This report is written by Elena Gorokhova, Sebastian Abel, Gastón Alurralde, Ann-Kristin Eriksson Wiklund, Anna Sobek, Yves Saladin, Sophie Steigerwald and Brita Sundelin (Stockholm University).

The report has been reviewed for scientific quality by Ingela Dahllöf (University of Gothenburg) and for practical relevance by Johan Gustafsson (Swedish Agency for Marine and Water Management) and Elisabeth Nyberg (Swedish Environmental Protection Agency).

The authors are responsible for the content of the report.

Stockholm, December 2025

Marie Uhrwing
Head of Department
Sustainable Development Department

Contents

Preface 3

Summary 9

Sammanfattning 10

1. Introduction 11
1.1 Advancing environmental assessments with chemical activity metrics 11
1.2 Project activities 12
1.3 Overarching hypotheses and objectives 13

2. Material and methods 15
2.1 Experimental methods 16
2.1.1 Dose metrics and chemicals 16
2.1.2 Test species 18
2.1.3 The exposure systems 18
2.1.4 Cumulative impacts evaluated and test factors 19
2.2 Field studies 20
2.3 Statistical modelling 21
2.3.1 Data transformation, univariate tests and missing data 22
2.3.2 Dose-response analysis 22
2.3.3 Multivariate modelling 22
2.3.4 PLS-SEM 22
2.3.5 Decision tree modeling 23

3. Results 24
3.1 Experimental evidence for suitability of chemical activity as a dose metric 24
3.1.1 Experimental systems 24
3.1.2 Cumulative effects in acute and chronic exposures 25
3.2 Effects of contaminants and abiotic factors on ReproIND 27
3.3 DT analysis of GES targets 29
3.4 Pressure-indicator relationships 31

4. Discussion 33
4.1 Chemical activity in ecotoxicity testing 33
4.2 Evidence for mixture effects from the field studies 35
4.3 ReproIND responses to regional variability in contaminant and environmental conditions 35

5. Conclusions and recommendations 38
5.1 Further developments in ReproIND implementation and GES setting 38
5.2 Linking chemical and biological indicators 39

6. Acknowledgements 41

7. References 42

8. Publications and data availability 46
8.1 Peer-reviewed publications 46
8.2 Reports 46
8.3 Preprints 46
8.4 Master theses 47
8.5 Scientific conference presentations 47
8.6 Stakeholder presentations 48
8.7 Open access data 48

Tables

Table 1. Overview of the test/target species, experimental factors and variables used in the experimental and field studies. Corresponding objectives (O1-O4 as outlined above) are indicated for each activity 16

Table 2. Overview of the data sources and projects used to harvest data on environmental contaminants, environmental factors and ReproIND components (frequency of embryo aberrations in the population of *Monoporeia affinis* and frequency of females carrying more than one aberrant embryo) 21

Table 3. GLM output for GES effect on the relationship between chemical activity (LogPAHa) and TOC. Sigma-restricted parameterization, Effective hypothesis decomposition; Std. Error of Estimate: 0.6 055 32

Figures

Figure 1. Overview of the project activities 13

Figure 2. Contribution of PAHs to total chemical activity in sediments, based on data from the Effect Screening Study (ESS) and two additional studies that included measurements of contaminants beyond PAHs. The contaminants used for the total chemical activity calculations are listed within each panel. Each data point is a study site, and the short horizontal line indicates the mean value .. 17

Figure 3. Experimental systems for the waterborne (A) and sediment (B) exposure to PAHs. Passive dosing with silicone as a carrier was used in exposures with *Daphnia* (A). Sediment-bound exposure was conducted using chironomids and earthworms (B). In addition, the chemical activity-based loading procedure for HOCs in artificial sediments was developed (C and D), where polyethylene bags filled with PAH crystals and water (C) were placed in vessels containing a peat-water mixture to saturate the peat (D). 18

Figure 4. Growth response of earthworms and chironomids exposed to sediment with varying PAH activities. Earthworm growth decreased with increasing chemical activity (Ea_{50}: 0.01–0.07), while chironomid growth remained largely unaffected, reflecting differences in PAH uptake between the two species 25

Figure 5. Dose-response curves for *Daphnia magna* immobilization rate (a proxy for mortality) exposed to the PAH mixture at chemical activity range of 0.01–0.1. The daphnids were grown at either 20 oC or 25 oC and exposed to the PAHs at the same temperature for three degree-days. At 20 oC, this corresponded to 72 h (blue) and at 25 oC, to 60 h. In addition, the daphnids from the 25 oC treatment, were exposed to the same dose range during 72 h. The median lethal doses (La50) are shown on the insert (mean and 95%-CI) 26

Figure 6. Response of individual protein content in *Daphnia magna* exposed to the PAH mixture at chemical activity range of 0.01–0.1 with or without zinc 26

Figure 7. Variability of the reproductive aberrations across sites grouped according to their contaminant load. (A) Hierarchical clustering of the sampling sites according to the contaminant load (Ward linkage), yielding four clusters (C1 to C4), and (B) Violin plots for the proportion of the females carrying more than 1 aberrant embryo. The width of each curve corresponds with the approximate frequency of data points in each region. The densities are additionally annotated with the median value and interquartile range, shown as a black boxplot within each violin plot. Significant differences between the groups detected by the unpaired t-test are indicated by an asterisk ($p < 0.05$). Note that C1 cluster was comprised by only 2 stations from Sundsvall with extremely high contamination levels for both HOCs and metals 28

Figure 8. The PLS-SEM model linking reproductive aberrations (ReproHealth) to biotic (TrophDiv, trophic diversity, and GenDiv, genetic diversity) and abiotic (EXPOSURE, chemical contaminants in the sediment, and physical, PHYS, environmental conditions). The paths link different latent variables (LVs circles). The biological LVs are shown as white circles, the chemical (EXPOSURE) and non-chemical (PHYS) LVs are shown in brown, and the indicators (the measured parameters representing some aspects of specific LVs) are hidden (denoted as [+]) for clarity. The EXPOSURE model had a three first-order LVs shown in light brown and representing different types of contaminants and pollution indices. 29

Figure 9. Decision Tree (DT) analysis of contaminants and environmental factors influencing GES derived from ReproIND values across 55 study sites used for model validation (red: non-GES, blue: GES). The decision tree model illustrates the relationship between contaminant pressure using composite variables for metals and PAHs (Metal-PLI and LogPAHa, chemical activity for PAHs, log-transformed) and GES status based on ReproIND outcomes. The model identifies critical contaminant thresholds, including LogPAHa (–5) and Metal-PLI (0.5) as key factors influencing GES outcomes, along with non-contaminant factors, bottom depth (75 m) and salinity (6.5). These findings suggest the necessity to adjust GES targets for reproductive aberrations, particularly in deeper sites, to enhance the detection of anthropogenic stressors and revise EQS for metals at salinities below 6.5 30

Figure 10. Variability in chemical activity (LogPAHa) and Metal-PLI values grouped for the sites classified as non-GES and GES. The output of the statistical comparison between the groups is shown below the x axis. The box-and-whisker plots show the mean, median (yellow line), 25 % - and 75 %, minimum, and maximum values for each parameter 31

Figure 11. Relationships between TOC in sediment and (A) chemical activity (LogPAHa) and (B) PLI for metals (Metal-PLI) for sites classified as GES and non-GES. The linear regression is significant for LogPAHa, with significantly lower intercept in the non-GES group (Table 3). By contrast, Metal-PLI was not clearly related to TOC, albeit this regression was weakly significant for the non-GES sites (Table 3B). The regression lines (panel A) are shown with their 95%-confidence bands, and the horizontal lines corresponding to Metal-PLI of 1 and 0.5 (panel B) indicate levels generally accepted as potential toxicity threshold (PLI = 1; Tomlinson et al. 1980) and the level identified by our DT model (PLI = 0.5) as increasing likelihood of adverse effects when HOCs are present at chemical activities exceeding 0.00 001 (i.e., LogPAHa >–5) 32

Figure 12. Summary of key decision points identified by the decision tree model (Figure 8) for assessing GES based on ReproIND outcomes. The figure highlights critical environmental thresholds, including bottom depth, Metals-PLI, and chemical activity (CA), that influence GES classification. Recommendations based on these decision points emphasize the importance of stratifying ReproIND data by site depth, prioritizing monitoring at sites with Metals-PLI >0.5, addressing the combined effects of metal pollution and HOCs, particularly at high chemical activity levels, and accounting for salinity effects on metal toxicity 39

Summary

Pollution in marine environments, such as the Baltic Sea, poses significant challenges to ecosystem health. Moreover, environmental variability related to temperature, salinity, or nutrient levels can alter the bioavailability and toxicity of chemical contaminants, leading to modified biological responses and influencing the effectiveness of environmental assessments. However, risk assessments are not equipped to consider the cumulative effects of contaminant mixtures and various environmental factors, failing to capture the full impact of real-world exposures. This gap was addressed in INSERT by investigating composite toxicity potential indices for non-polar hydrophobic organic contaminants (chemical activity) and trace metals (pollution load index; PLI) as dose metrics to evaluate the mixture effects of organic contaminants in environmental exposures and considering variable nonchemical conditions.

We focused on three main activities. First, we investigated the cumulative impact of chemical and nonchemical stressors on biological responses in standard ecotoxicological test systems. These experiments combined chemical activity with relevant stressors to enhance our understanding of their biological effects. Second, we applied the composite toxicity potential indices to ReproIND *(Reproductive aberrations in amphipods)*, an indicator for the holistic assessment of the Baltic Sea health in the context of MSFD implementation. This involved refining the GES targets by integrating chemical activity and PLI with regionally specific environmental variables to ensure that these targets accurately reflect the unique conditions and stressors in the Baltic Sea. Third, we engaged with expert groups and projects involved in HELCOM-guided biological effect assessments and collaborated with OSPAR experts.

In addition to the scientific findings, our study resulted in several key recommendations to enhance the practical value of contaminant screening surveys, monitoring and status assessment in the Baltic Sea, both nationally and internationally. We also demonstrated the value of using long-term monitoring data on the health status of sentinel species to integrate chemical and biological criteria for MSFD. In addition to advancing ReproIND implementation across the environmental gradients of the Baltic Sea, this integrated approach supports ongoing revisions of monitoring and assessment strategies for biological effects both nationally (SEPA/SWAM) and internationally (HELCOM, ICES, OSPAR).

Sammanfattning

Föroreningar i Östersjön och andra marina miljöer, utgör risker för ekosystemens hälsa. Dessutom kan naturliga variationer i temperatur, salthalt eller näringsnivåer påverka biotillgängligheten och toxiciteten hos kemiska föroreningar, vilket i sin tur påverkar tillförlitligheten i bedömningen av miljörisker och biologisk respons. Riskbedömningar är dock inte designade (utrustade) för att beakta de kumulativa (sammanlagda) effekterna av föroreningsblandningar och olika miljöfaktorer, vilket innebär att de misslyckas med att bedöma den fulla påverkan av exponeringar i naturliga miljöer. I INSERT adresserade vi detta genom att undersöka sammansatta toxicitetspotentialindicier för hydrofoba organiska föroreningar (kemisk aktivitet) och spårmetaller (föroreningsbelastningsindex, PLI) som dosmått för att utvärdera blandningseffekterna av föroreningar, samtidigt som vi tog hänsyn till varierande icke-kemiska faktorer.

Vi inriktade oss på tre centrala aktiviteter. För det första undersökte vi den samlade påverkan av kemiska och icke-kemiska stressfaktorer på biologiska reaktioner i standardiserade ekotoxikologiska testsystem. Dessa experiment kombinerade kemisk aktivitet med relevanta stressfaktorer för att förbättra vår förståelse av deras biologiska effekter. För det andra tillämpade vi de sammansatta toxicitetspotentialindexen på ReproIND *(reproduktiva avvikelser hos amfipoder)*, en indikator som används för en helhetsbedömning av Östersjöns hälsotillstånd i samband med genomförandet av havsmiljödirektivet (MSFD). Detta innebar att förfina GES-målen genom att integrera kemisk aktivitet och PLI med regionalt specifika miljövariabler för att säkerställa att målen baserad på ReproIND bättre speglar de unika förhållandena och stressfaktorerna i Östersjön. För det tredje samarbetade vi med expertgrupper och projekt inom ramen för Helcom som arbetar med biologiska effektbedömningar, och vi samarbetade även med experter från Ospar.

Förutom de vetenskapliga resultaten ledde vår studie till flera viktiga rekommendationer för att förbättra det praktiska värdet av föroreningsundersökningar, övervakning och statusbedömning i Östersjön, både på nationell och internationell nivå. Vi visade också värdet av att använda långsiktiga övervakningsdata om hälsostatus hos indikatorarter för att integrera kemiska och biologiska kriterier i arbetet med havsmiljödirektivet (MSFD). Utöver att främja implementeringen av ReproIND över Östersjöns miljögradienter stödjer detta integrerade tillvägagångssätt även de pågående översynerna av övervaknings- och bedömningsstrategier för biologiska effekter, både nationellt (NV/HaV) och internationellt (Helcom, Ices, och Ospar).

1. Introduction

The European Union Marine Strategy Framework Directive (MSFD) mandates a systematic assessment of the environmental status of European regional seas to ensure they achieve Good Environmental Status (GES). This involves developing and utilising various tools to define and measure GES targets, specifically focusing on Descriptor 8 (D8), which addresses the impact of contaminants on marine environments. According to the directive, member states must demonstrate that contaminant concentrations are at levels that do not lead to pollution effects. However, assessing environmental quality is primarily based on chemical (Criterion 8.1) and less on biological (Criterion 8.2) indicators that reflect the contaminant pressure in marine ecosystems. These indicators simplify complex ecological processes for evaluating ecosystem health and formulating hypotheses about contaminant effects. It's essential, therefore, that these indicators provide a practical snapshot of how environmental stressors affect ecosystems, and even though they represent broader ecological processes, their target values (GES targets) accurately reflect critical environmental features to enable effective management.

1.1 Advancing environmental assessments with chemical activity metrics

Amphipods are well-established sentinels for monitoring due to their widespread distribution and representation of various ecological groups, i.e., sediment-dwelling, littoral, and nektonic species. Embryo aberration analysis for assessing reproductive and developmental toxicity was established in 1994 (Sundelin and Eriksson, 1998), underscoring the importance of monitoring reproductive and developmental effects of contaminant exposure for ecologically meaningful contaminant assessment in complex pollution scenarios.

ReproIND, an indicator of embryo aberrations in amphipods (HELCOM, 2023 a), is one of the few available indicators for biological effect assessment in the region. This indicator, first implemented by the Swedish National Marine Monitoring Program (SNMMP) and then accepted by HELCOM, has been recently recommended by SGEFF expert group (Study Group on developing new guidelines for the monitoring of biological EFFects of contaminants; HELCOM-OSPAR) to support MSFD biological effect monitoring for MSFD implementation. Thus, this indicator is integral to the holistic assessments of the Baltic Sea's health (HELCOM, 2023 b, 2018) and has been used in HOLAS II (2011–2016) and HOLAS III (2016–2021) assessment cycles.

Exposure to environmental contamination involves chemical mixtures and environmental factors like temperature, salinity, and oxygen, leading to complex effects often overlooked in traditional risk assessments. Despite existing frameworks for mixture effect assessments in aquatic environments (Holmes et al., 2018; Redman et al., 2017), their broad implementation is limited by mixture complexity, insufficient mechanistic data, regulatory constraints, and uncertainties in predicting cumulative effects (Beyer et al., 2014). Additionally, most assessments rely on chemical data, with little integration of biological evidence, hindering meaningful implementation of Descriptor 8 in the MSFD, and underscoring the need for habitat classification with target values based on combined chemical and biological criteria.

To distinguish contaminant impacts from other stressors, both field and experimental research are needed. Identifying confounding factors helps develop models that separate chemical from non-chemical causes of biological effects, such as embryo abnormalities. This approach clarifies contaminant effects and supports better management. Integrating toxicity indices like Toxic Units (TU), Toxic Equivalency Quotient (TEQ), and Hazard Index (HI) enhances exposure assessment and improves understanding of cumulative toxicity in complex environments (Swain, 2024; Van den Berg et al., 1998; Warne, 2003).

Chemical activity is an emerging metric in environmental toxicology that provides a unified framework for evaluating the impact of chemical contaminants on ecosystems (Gobas et al., 2018). Unlike traditional concentration-based assessments, which measure individual chemical levels, chemical activity integrates the cumulative effects of multiple substances by quantifying their combined potential for toxicity. This metric accounts for baseline toxicity caused by various neutral organic chemicals at low concentrations, reflecting their collective impact on biological systems. Chemical activity is particularly valuable for assessing mixture toxicity and identifying excess toxicity (i.e., specific toxicity not caused by narcosis) that may not be apparent from individual concentration measurements (Mackay et al., 2011). Despite these advantages and the solid theoretical background, chemical activity is not practically adopted in contaminant effect assessments, which rely on concentration measurements of selected contaminants. In INSERT, we explored how chemical activity representing the toxicity potential of nonpolar hydrophobic contaminants (HOCs) can be integrated into the current risk assessment framework to better link exposure risks with biological effects.

1.2 Project activities

INSERT addressed the gaps in understanding cumulative effects in Baltic sediments by using chemical activity as the dose metric for HOCs. Given the significance of metals as contaminants, the interactive effects of HOC and metal exposure were also included using chemical activity and either single metal exposure (experimental studies) or the Pollution Load Index (PLI; (Tomlinson et al., 1980)) in sediment (field studies) as predictors of biological effects. By integrating chemical and nonchemical stressors, the study refined ReproIND for holistic contaminant assessment in the Baltic Sea (Figure 1).

Our work was structured as follows:

- Exploratory laboratory tests used daphnids, chironomids and earthworms as standard test species in ecotoxicology. Their sensitivity to pollutants, short life cycles, established testing protocols, and experimental versatility allowed controlled studies of contaminant-stressor interactions.
- ReproIND was validated and refined for MSFD implementation by integrating composite chemical exposure indices (chemical activity for HOCs and PLI for metals) and nonchemical factors (temperature, salinity, and others). Field data

on Baltic Sea amphipods and advanced statistical modelling quantified these relationships. Based on the outcomes of this work, we generated recommendations for ReproIND improvements and proposed an approach to integrate chemical and biological effect assessments.

- We collaborated with HELCOM, OSPAR, and national research projects to improve biological effect monitoring and ReproIND implementation in the Baltic Sea. These efforts have been instrumental in conceptualising how nonchemical factors can be accounted for in setting GES values to enhance HELCOM-guided biological effect monitoring in the Baltic Sea.

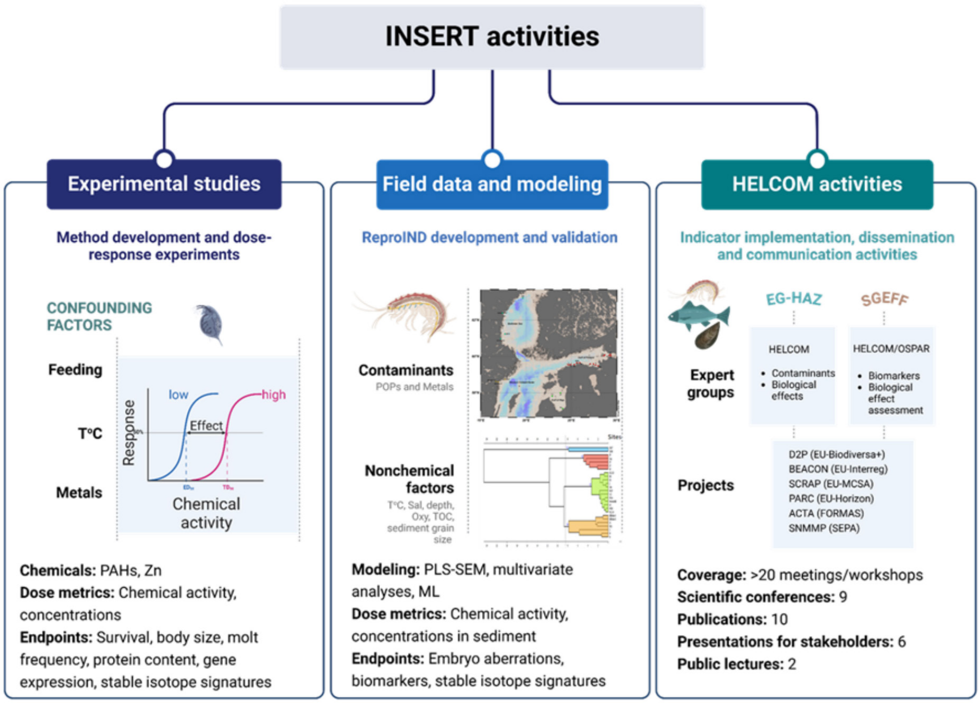


Figure 1. Overview of the project activities.

1.3 Overarching hypotheses and objectives

The overall aim was to improve assessments of cumulative effects and environmental status in the Baltic Sea by integrating chemical and biological measures. The hypotheses (H1–H4) test the utility of composite toxicity indices, refine the ReproIND indicator, and evaluate nonchemical environmental influences. The objectives (O1–O6) address these hypotheses through targeted laboratory and field studies, robust diagnostics, stakeholder engagement, and practical recommendations for monitoring programs.

Hypotheses

- **H1: Composite toxicity potential indices and biological effects:** The chemical activity and PLI estimates provide an adequate measure of the combined toxic potential of chemical mixtures. Chemical activity effectively captures baseline toxicity (narcosis) and helps identify cases of excess toxicity (specific toxicity).
- **H2: ReproIND revision:** When integrated with chemical activity estimates and adjusted for influential environmental variables, ReproIND offers a more accurate assessment of developmental toxicity in amphipods and improves the classification of environmental status with regard to chemical exposure (i.e., GES assessment).
- **H3: Nonchemical factors:** Environmental factors, such as temperature, salinity, and nutrition-related variables significantly modulate the biological responses observed in the test animals, necessitating their integration into environmental assessments to refine GES values.

Objectives

- **O1: Evaluate the chemical activity as a dose metric in ecotoxicological assays:** Investigate the applicability and accuracy of the chemical activity for assessing the combined toxicity of HOC mixtures in water- and sediment-based assays. Select relevant HOCs for these mixtures using practical considerations and the contaminant data harvested by the field studies.
- **O2: Identify influential nonchemical factors:** Conduct laboratory and field studies to identify the influence of nonchemical factors on biological responses related to growth and reproduction and quantify their impact. Use field observations in amphipods from different basins to maximise the variability of the environmental gradients in the statistical models. Develop robust statistical diagnostics to differentiate between chemical and nonchemical drivers of embryo aberrations, facilitating more precise attribution of biological responses to contaminants in situ.
- **O3: Refine ReproIND targets:** Validate and refine the existing GES targets for the ReproIND indicator (i.e., the acceptable background values for the frequency of malformed embryos in the population and the frequency of females carrying malformed embryos) in the context of the MSFD framework by integrating chemical activity, PLI, and region-specific environmental variables.
- **O4: Engage stakeholders:** Collaborate with HELCOM, OSPAR, and national research projects on contaminant and biological effect assessments to integrate INSERT findings into current risk assessment frameworks. Support the revision and implementation of GES targets and promote the integration of chemical and biological effect components of D8/MSFD.
- **O5: Improve effect-based monitoring of contaminants in the Baltic Sea:** Provide actionable recommendations for designing and enhancing monitoring programs for biological effects of contaminants, ensuring the efficient integration of chemical and nonchemical parameters in the monitoring programs.

2. Material and methods

In our studies, we systematically explored the use of chemical activity to assess contaminant impacts on animal responses to chemical mixtures, developing methods and employing diverse experimental and sampling designs. The methods included:

- *Laboratory studies* identified biological responses in the standard test species (daphnids, chironomids and earthworms; Table 1) to the chemical exposure, evaluate confounding factors influencing these responses, and develop statistical models to quantify the relative contribution of the significant predictors of the observed effects. Experimental systems using passive dosing (waterborne exposure) were developed for dose-response studies of HOC mixtures utilizing climate-controlled chambers to regulate exposure conditions and laboratory animal cultures. These studies evaluated chemical activity effects under chemical (zinc co-exposure) and nonchemical (nutrition, temperature) confounding factors. Rigorous measures were taken to measure exposure levels, ensuring the integrity and accuracy of the experimental results.
- *Field* studies evaluated the ReproIND indicator (amphipod embryo aberrations and frequency of affected females) using data from *Monoporeia affinis* collected across the Baltic Sea (Bothnian Sea, Northern Baltic Proper, Western Gotland Basin, Gulf of Riga, Gulf of Finland). Additional data were gathered from published sources and open-access repositories (ICES DOME, SHARKweb, GitHub). Machine learning analyzed the combined effects of chemicals and environmental factors (temperature, salinity, oxygen, sediment), refining environmental status classification, contaminant diagnostics, and understanding of nonchemical stressors.

Table 1. Overview of the test/target species, experimental factors and variables used in the experimental and field studies. Corresponding objectives (O1–O4 as outlined above) are indicated for each activity.

Activity, objectives	Exposure chemical(s)	Nonchemical factors	Target species	Endpoints
Experimental studies (O1 and O2)	PAH mixture, zinc	Feeding, temperature	*Daphnia magna*	Survival, body size (body length and individual protein content), moult frequency, gene expression, stable isotope signature (^{13}C)
Experimental studies (O1)	PAH mixture		*Lumbriculus variegatus, Chironomus riparius*	Uptake of PAHs, relative growth
Field studies, basin-specific effects of individual contaminants and nonchemical factors (O2)	Various POPs, trace metals	Temperature, salinity, oxygen, bottom depth, TOC, chlorophyll *a*, sediment PSD	*Monoporeia affinis*	Fecundity, embryo aberration frequency, body size (protein content), genetic diversity, SI signature, population abundance
Field studies, decision tree modelling (O3)	Chemical activity (based on PAHs), metal pollution index (based on trace metal concentrations)	Temperature, salinity, oxygen, bottom depth, TOC	ReproIND based on *Monoporeia affinis*	GES condition based on embryo aberration frequency

2.1 Experimental methods

2.1.1 Dose metrics and chemicals

Hydrophobic organic contaminants, such as PAHs, preferentially dissolve in non-polar environments like lipids, leading to their accumulation in lipid tissues and phospholipid membrane layers and narcosis caused by non-specific disruption of the proper functioning of the cell membrane (generally thought of as the site of toxic action). The reported data and theoretical considerations relating toxicity of nonpolar narcotic compounds to the logarithm of the octanol-water partition coefficient (Kow) have been used to define the target interval for baseline toxicity.

Chemical (or thermodynamic) activity is a dimensionless measure of how “active” a chemical is within a specific phase, such as lipids. It represents the ratio between the chemical’s concentration (or partial pressure) in the phase and its saturation point in that phase. Chemical activity is additive for neutral chemicals, making it a useful metric for assessing the impact of the mixtures of hydrophobic chemicals (Gobas et al., 2018). Activities in the range of 0.01–0.1 are most relevant for mixtures for mixture toxicity studies where narcosis is the primary mode of action. At these levels, chemicals are not fully saturated in lipids, but are concentrated enough to jointly induce narcotic effects (Gobas et al., 2018).

Based on field contaminant data, PAHs were selected as representative HOCs for experimental studies. Since chemical activity-based experiments depend on overall chemical activity rather than individual substance identity, a consistent mixture of four PAHs–acenaphthene, fluorene, fluoranthene, and phenanthrene–was used across experiments. All chemicals were high purity (≥ 98%) and prepared following our published methods (Abel et al., 2024; Steigerwald et al., 2024).

PAH and metal concentrations were reported for all field stations with ReproIND data (Figure 2). However, only the Effect Screening Study (Förlin et al., 2019; Gorokhova et al., 2023) provided a comprehensive selection of HOCs, while the other two studies (Kolesova et al., 2024; Löf et al., 2016) reported only PAHs and PCBs. For consistency, chemical activity was calculated solely from PAHs, acknowledging potential underestimation due to omitted HOCs.

To address this underestimation and evaluate PAHs' contribution to total chemical activity, we analyzed chemical activity distribution using the most comprehensive multi-HOC dataset (Figure 2). This analysis demonstrated that PAHs typically accounted for over 90 % of the total chemical activity (median 93.2 %; 95 %-confidence interval: 82.3 %–98.1 %). A notable exception was the Gulf of Finland, where PCBs contributed substantially (> 50 %). Therefore, we believe that our chemical activity values based on PAHs are fairly representative for the HOCs load in the sediment.

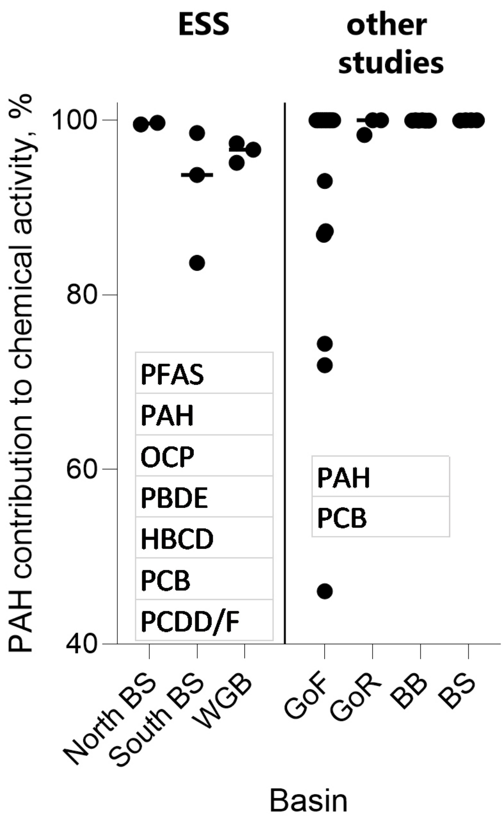


Figure 2. Contribution of PAHs to total chemical activity in sediments, based on data from the Effect Screening Study (ESS) and two additional studies that included measurements of contaminants beyond PAHs. The contaminants used for the total chemical activity calculations are listed within each panel. Each data point is a study site, and the short horizontal line indicates the mean value.

2.1.2 Test species

Daphnia magna (environmental pollution test clone 5, the Federal Environment Agency, Berlin, Germany) were used in the acute tests. They were cultured (~ 10–15 ind./L) in M7 medium and fed with a mixture of green algae grown in culture following OECD guidelines (OECD, 2004). The medium was changed weekly, and cultures were kept at 20 °C ± 1 °C and a 16 L:8 D h photoperiod.

For the chronic sediment-water exposures, earthworm *Lumbriculus variegatus* and *Chironomus riparius* were used, both are standard test species for assays with sediment (OECD, 2023, 2007). The cultures were obtained via collaboration with Joensuu University, Finland.

2.1.3 The exposure systems

We further developed and used two exposure systems to address different environmental contexts where PAHs are commonly found (Figure 3). The waterborne exposure system replicates conditions in aquatic environments to study PAH toxicity and bioavailability in water. The sediment exposure system simulates PAH accumulation in sediments, allowing for stable contaminant levels and impacts in test vessels with sediment as the primary source of chemical activity. In both systems, the target chemical activity range was 0.01–0.1, and PAH concentrations in the medium were verified using HPLC-PDA; the results were converted into temperature-adjusted chemical activities and used in the dose-response analysis.

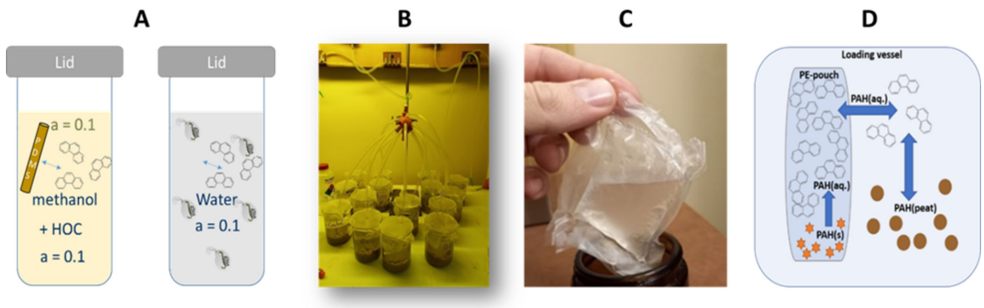


Figure 3. Experimental systems for the waterborne (A) and sediment (B) exposure to PAHs. Passive dosing with silicone as a carrier was used in exposures with *Daphnia* (A). Sediment-bound exposure was conducted using chironomids and earthworms (B). In addition, the chemical activity-based loading procedure for HOCs in artificial sediments was developed (C and D), where polyethylene bags filled with PAH crystals and water (C) were placed in vessels containing a peat-water mixture to saturate the peat (D).

- *Waterborne exposure.* Passive dosing is a method used in aquatic toxicology to maintain stable concentrations of hydrophobic organic chemicals in water (Figure 3 A). This technique is particularly useful for studying HOCs and other poorly water-soluble compounds (Smith and Jeong, 2021). A stock solution was used to prepare loading solutions for passive donors (silicone rods) to maintain stable exposure to the PAH mixture at five levels of total chemical activities. The silicone rods were thoroughly cleaned, soaked in a methanol-loading solution, and then equilibrated with the exposure medium.
- *Sediment exposure.* Artificial sediment was prepared and loaded according to OECD 218 and 225 using quartz sand (Figure 3 B). The earthworms *(Lumbriculus variegatus)* and chironomids *(Chironomus riparius)* were exposed to a chemical activity gradient (0.01–0.12) for 28 and 12 days, respectively, according to the OECD guidelines. Food for the earthworms (*Urtica* sp. powder) was mixed into the sediment prior to the experiment, whereas for the chironomids, aquaria fish food was provided three times a week during the exposure. Exposure confirmation was done by measuring the bioavailable fraction of PAHs in the sediment using the coated jar method (Jahnke et al., 2014) and thereafter converting the concentrations to chemical activity.
- *Chemical-activity based loading of sediment* (Abel et al., 2024). A chemical activity-based loading procedure for HOCs in artificial sediments was developed by saturating peat with PAHs (Figure 3 C, D). Polyethylene bags filled with PAH crystals and water were placed in vessels containing a peat-water mixture (Figure 3 C) to saturate the peat (Figure 3 D). These vessels were regularly sampled, and chemical activity was calculated. Equilibrium was achieved within 5–17 weeks, depending on the specific PAH kinetics. The target activity range in the sediment was adjusted by mixing loaded and clean peat. Unfortunately, time was insufficient to use the sediment exposure for cumulative impact testing.

2.1.4 Cumulative impacts evaluated and test factors

The cumulative impacts in controlled experimental systems were assessed using waterborne exposures with *Daphnia magna* as the test species. The experiments were designed to evaluate the following effects:

- *Animal nutritional status.* The nutritional status of *Daphnia* neonates was manipulated by feeding or starving them before the start of the experiment to assess its impact on the toxicity dose. The OECD test was conducted using immobilization, individual protein content and growth as endpoints. These were measured after 72 hours of PAH exposure and following a 48-hour recovery period with food in clean media (Saladin, 2023; Steigerwald et al., 2024).
- *Temperature.* Virtually all physiological functions are temperature-dependent. Moreover, the actual impact of a contaminant at a given chemical activity could vary with temperature because as the temperature rises, the solubility, volatility, and diffusion of contaminants through the cell membranes can change, altering their concentration in different phases, such as water, membranes and intracellular lipids. We used a combination of the degree-days (DD, proxy for physiological age) approach to delineate thermal stress *per se* from the temperature-induced toxicity, and corrected chemical activities for temperature. The

rationale was as follows. In standard ecotoxicity testing, the temperature is held optimal, and the exposure time is usually fixed to, e.g., 24, 48, or 72 h, which relies on the assumption of the same rate of ageing at the constant temperature. However, ageing and metabolic activity are temperature-dependent, which requires the standardization of the exposure time and temperature using the DD concept. Daphnids were exposed to the narcosis chemical activity range for 3 DD at 20 ˚C (72 h) and 25 ˚C (72 and 60 h). Immobilisation and individual protein content were used as endpoints.

- *Combined exposure to PAHs and heavy metals.* A combined PAH/zinc, PAH only and zinc only exposures for 72 h were applied using immobilisation and individual protein content as endpoints. The full narcosis chemical activity range at two levels of zinc exposure (nominal concentration of 1 and 2 mg/l) was used for PAHs and PAH/zinc treatments.

2.2 Field studies

The project began with three studies (Gorokhova et al., 2024; Kolesova et al., 2024; Raymond et al., 2021), each exploring these relationships at different scales and regions. A common central objective of these studies was identifying the primary drivers of reproductive disorders and/or amphipod population abundance, focusing on contaminants and environmental factors.

Further, the data gathered in each of these and other relevant studies conducted earlier (Löf et al., 2016) or within other projects (Gorokhova et al., 2023; Martella et al., 2023) were collated. Missing data on abiotic factors were harvested from open data repositories (ICES DOME; https://dome.ices.dk/datsu/; amphipod reproductive success) and SHARKweb (https://sharkweb.smhi.se/hamta-data/; abiotic parameters). Finally, a few international research projects also contributed (Table 2). This dataset was then used in a modeling study (Gorokhova, 2024) to identify association rules and thresholds, with the goal of refining GES targets based on ReproIND, chemical pressures, and abiotic environmental variables across several basins (Bothnian Bay, Bothnian Sea, Western Gotland Basin, and Gulfs of Finland and Riga). The primary data used in the model are summarised in Appendix I and are publicly available via Zenodo hub.

The biological effect data included reproductive aberrations in *M. affinis* (two ReproIND components: frequency of malformed embryos in population and frequency of females carrying malformed embryos; (HELCOM, 2023 a)). The methods for this analysis are described elsewhere (Löf et al., 2016; Sundelin and Eriksson, 1998).

Chemical activity was calculated with Frank Gobas calculator [https://www.sfu.ca/rem/toxicology/our-models/activity-calculator.html] using PAH concentrations reported in the original studies (NAP, ACNE, FLE, PA, ANT, FLU, PYR, BAA, CHR, BBF, BKF, BAP, DBAHA, BGHIP, and ICDP). Sediment TOC was set to 2 % and temperature to 10 °C for the chemical activity calculations.

The Pollution Load Index for metals (Tomlinson et al., 1980) was calculated as a geometric average of contamination factors and served as a proxy for metal toxicity in the sediment for the nine metals that were reported consistently in the source datasets (As, Cd, Co, Cr, Cu, Hg, Ni, Pb and Zn). For PLI calculations, we used the regional reference values for these metals in sediment (SEPA, 1999).

The environmental variables (temperature, salinity, oxygen, TOC and bottom depth at the site) used in the statistical models were derived from the source datasets; in a few cases where any of these variables were not reported, they were harvested from the SHARKweb or extracted from Data Assimilation System (DAS; (Sokolov et al., 1997)), using linear interpolation/extrapolation. Briefly, DAS is a program that constructs a 3 D grid of hydrographic and chemical monitoring data from all countries around the Baltic. Environmental data were extracted from the relevant year (i.e., the year before *M. affinis* collection) by interpolating/extrapolating the measured values with 1 and 2 nautical mile resolutions in northward and eastward directions, respectively.

Table 2. Overview of the data sources and projects used to harvest data on environmental contaminants, environmental factors and ReproIND components (frequency of embryo aberrations in the population of *Monoporeia affinis* and frequency of females carrying more than one aberrant embryo).

Source	Basin	Contaminant variables	Environmental variables	Project
Löf et al., 2016	Bothnian Bay, Bothnian Sea	Metals, PAH, PCB	T °C, Depth, Sal, Oxy, TOC	BEAST project funded jointly by BONUS+ Program (FP7/2007–2013, Grant agreement 217246) and the Swedish EPA;
Förlin et al., 2019; Gorokhova et al., 2023	Bothnian Bay, Bothnian Sea, Western Gotland Basin	Metals, PAH, PCB, PCDD/F, PFAS, OCP, HBCD, BT, PBDE	T °C, Depth, Sal, Oxy, TOC, PSD	Effect Screening Study, Swedish EPA; 2018–2020
Raymond et al., 2021	Bothnian Sea, Western Gotland Basin	Metals, PAH	T °C, Depth, Sal, Oxy, TOC	King Carl XVI 50-year Foundation for Science, Technology and Environment, ReproIND (FORMAS; 2017-00864) and INSERT (SEPA) projects.
Martella et al., 2023	Bothnian Sea, Western Gotland Basin	Metals, PAH	T °C, Depth, Sal	ReproIND (FORMAS; 2017-00864) and AmphiDNA (FORMAS, 2019-01157).
Kolesova et al., 2024; Strode et al., 2017	Gulf of Finland, Gulf of Riga	Metals, PAH, PCB, BT	T °C, Depth, Sal, Oxy, TOC, PSD	Several projects, including BEACON (Interreg Baltic Sea Region co-funded by the EU; S007 10/2022-10/2024) and D2P (Biodiversa+; 2021-473).

2.3 Statistical modelling

Univariate and multivariate techniques were applied to explore relationships between variables and assess sources of variability. In the experimental studies, dose-response analysis was the primary tool for examining the effects of exposure. In addition, for both experimental and field studies, we commonly used regression-based methods, particularly Generalized Linear Models (GLMs) and multivariate techniques (PERMANOVA, CAP, and DistLM) to identify significant predictors.

Structural Equation Modeling (partial least squares SEM; PLS-SEM) was used to explore direct and indirect effects between reproductive health, sediment contaminants, environmental variability, genetic diversity, and nutrition in *M. affinis* in the Western Gotland Basin and the Bothnian Sea. The particular focus was on the mediation of contaminant effects via trophic and genetic population properties on ReproIND with GES status as a primary outcome.

Finally, decision tree (DT) analysis, a machine learning technique, was employed to identify patterns within large datasets. This approach is particularly effective for identifying association rules and thresholds, and can handle multicollinearity and small sample size (being the case for our dataset), while also providing threshold estimates for the variables of interest (i.e., chemical and environmental variables) to refine GES targets.

2.3.1 Data transformation, univariate tests and missing data

Data pre-processing involved several key steps. First, the data were cleaned to remove duplicates and correct any inconsistencies. Numerical variables were standardized and outliers were identified and either transformed or excluded based on their impact on the analysis. The data transformations, such as log or square root transformations, were applied where necessary to meet the assumptions of statistical models and enhance interpretability. The Expectation-Maximization (EM) algorithm was applied to address a few missing contaminant values in the field datasets, leveraging Primer 7 software for this analysis.

2.3.2 Dose-response analysis

For the dose-response analysis, the chemical activity was used as the dose metric. Unlike traditional concentration-based assessments, chemical activity provides a measure of the fraction of a chemical's potential to exert baseline toxicity, accounting for the combined effects of the four PAH congeners within a mixture. This approach enabled us to quantify the toxicity of the PAH mixture and indicate whether a specific toxicity (i.e., a response that is not due to narcosis) is observed.

2.3.3 Multivariate modelling

Multivariate ordination techniques were utilized in the field studies to investigate the relationships between reproductive disorders, contaminants and environmental factors. For example, Hierarchical Cluster Analysis (HCA) was used to group sites with similar contaminant profiles (Kolesova et al., 2024), employing Euclidean distance and Ward's method to enhance classification accuracy. Distance-based Linear Modeling (DistLM) assessed the relative importance of environmental factors (sediment PSD, temperature, oxygen, TOC) and contaminants in explaining variations in amphipod responses (Raymond et al., 2021). This approach utilized a distance matrix for predictor selection, with permutations to test the significance of relationships.

2.3.4 PLS-SEM

PLS-SEM was used to analyze connections between reproductive health, sediment contaminants, environmental variability, genetic diversity, and nutrition in *Monoporeia affinis* in the Western Gotland Basin and the Bothnian Sea. PLS-SEM was the method of choice due to its suitability for complex, predictive models involving multiple latent variables (LV) that are impossible to measure directly and observed variables used to parameterize the LVs, allowing for the analysis of both direct and indirect effects in a single model. Its ability to manage small to medium sample sizes and accommodate data with non-normal distributions makes it ideal for studying environmental impacts and interactions, where traditional methods might fall short.

In our model (Gorokhova et al., 2024), the LVs included embryo aberration frequencies, isotopic niche metrics for resource utilization (trophic status of the population), and levels of chemical contaminants (metals, PAHs and composite indices of potential toxicity). The analysis also considered non-chemical factors such as temperature and the potential role of genetic diversity (mtDNA markers) in the trophic status and ReproIND variation.

2.3.5 Decision tree modeling

We used decision tree (DT) analysis to evaluate the relationships between contaminant exposure and ReproIND across varying environmental conditions in the Baltic Sea (Gorokhova, 2024). DTs offer several advantages for this type of analysis, including their ability to handle non-linear interactions between variables, which is essential given the intricate dynamics between contaminants and non-chemical environmental factors, such as temperature, salinity, and oxygen levels. Furthermore, DTs are well-suited for analysing relatively small sets of categorical and numerical data, making them versatile for our dataset with only 85 observations. Their interpretability also allows us to rank the key GES drivers and effect levels associated with non-GES conditions, detect thresholds, and reveal critical points where environmental pressures (i.e., chemical activity and PLI) may lead to biological impacts as well as at what environmental conditions (i.e., temperature, depth and salinity) these impacts are exacerbated. These strengths made DT a good choice for our task, offering a pragmatic framework for environmental data assessment and interpretation.

3. Results

We found experimental and field evidence for chemical activity driving biological responses in acute and chronic exposure. We further developed and validated a chemical-activity based loading method for sediments. Below, we present our results in the following order:

- In section 3.1, we outline the experimental exposure system and findings on how confounding factors (e.g., feeding, temperature and co-exposure to metals) affect the dose-response relationship to PAH exposure expressed as chemical activity in standard test organisms.
- In section 3.2, the influential contaminants and the abiotic factors affecting ReproIND in the wild populations of *Monoporeia affinis* in different basins are summarized.
- In section 3.3, the DT model results, association rules, and thresholds for the influential chemical and nonchemical variables with non-GES conditions serving as the basis for ReproIND revisions are presented.

3.1 Experimental evidence for suitability of chemical activity as a dose metric

Our experimental studies provided dose-response data for tests conducted according to OECD guidelines with chemical activity as a dose metric, advancing this approach in the ecotoxicology of HOCs (Saladin, 2023; Steigerwald et al., 2024). The developed and refined protocols also enable more robust experimental designs, allowing for more precise hypothesis testing when both chemical and non-chemical variables are involved. This suggests that standard guidelines may need updating to incorporate variability in environmental factors, ensuring that toxicity testing more accurately predicts combined effects.

3.1.1 Experimental systems

The proposed loading method based on chemical activity for sediment exposure allow greater stability and control of exposure levels in the experimental studies, leading to more accurate dose estimate (Abel et al., 2024). Additionally, adjustments for temperature effects on chemical concentration and sorption dynamics have been integrated, with temperature-corrected chemical activities ensuring realistic exposure conditions (Oliveira dos Anjos et al., 2023). Furthermore, our observations emphasize the need to consider peat and sediment diversity in sediment toxicity testing, as current OECD guidelines do not account for variations in chemical activity across different types of the organic substrate (Arboleda, 2024). This finding is particularly important for cases when the exposure level is based on the total sediment concentration and not freely dissolved concentrations or chemical activity. These developments collectively enhance the ability to apply relevant exposure conditions in sediment toxicity testing.

3.1.2 Cumulative effects in acute and chronic exposures

- *Animal nutritional status.* The nutritional status of *Daphnia* neonates significantly affected the toxicity dose for growth, protein content, and immobilization. The starved individuals could cope much better with the PAH exposure, reflected in the significantly higher La50 value (i.e., lethal activity for 50 % of the exposed animals). Moreover, starved individuals had greater recovery success (Saladin, 2023; Steigerwald et al., 2024). The feeding regime was also likely to affect the PAH uptake by earthworms and chironomids in the sediment exposure (Figure 4). While the accumulation in chironomids was low, it substantial in the earthworms. This uptake pattern was also reflected in the growth of the test species, with the earthworm growth decreasing with increasing chemical activity (Ea50, i.e., effect activity for 50 % of the exposed animals, varied between 0.01 and 0.07).

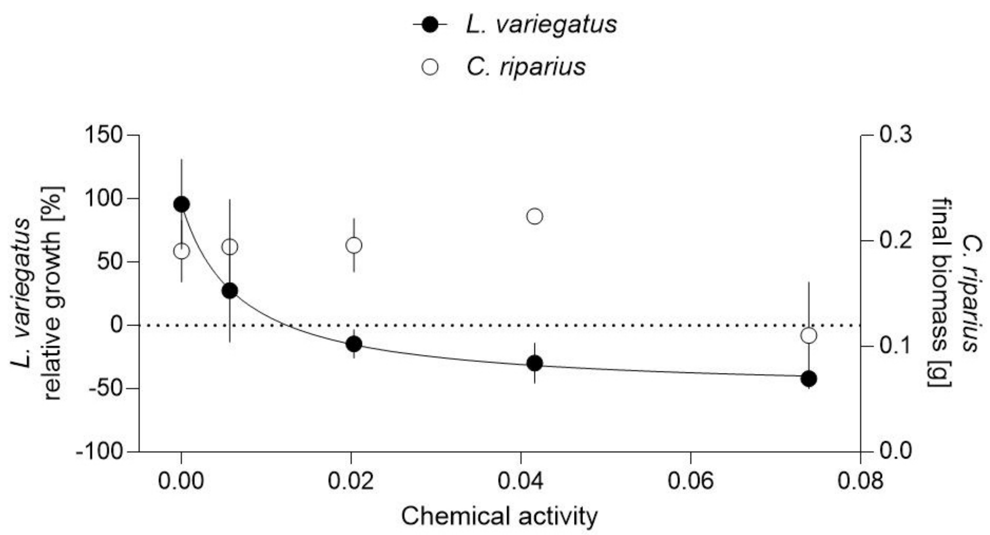


Figure 4. Growth response of earthworms and chironomids exposed to sediment with varying PAH activities. Earthworm growth decreased with increasing chemical activity (Ea_{50}: 0.01–0.07), while chironomid growth remained largely unaffected, reflecting differences in PAH uptake between the two species.

- *Temperature.* No significant temperature × dose interaction was found (Figure 5). When the exposure time was adjusted to degree days, the difference in La50 between the temperatures was not significant (La50: 0.061 vs 0.068). When the exposure time was set to 72 h, La50 was significantly higher at 25 ˚C compared to 20 ˚C (La50: 0.035 vs 0.068), which implies that the more advanced physiological age of the daphniids at the higher temperature, i.e., physiologically longer exposure, and not the enhanced toxicity at a given chemical activity were behind the higher “apparent” toxicity at the higher temperature.

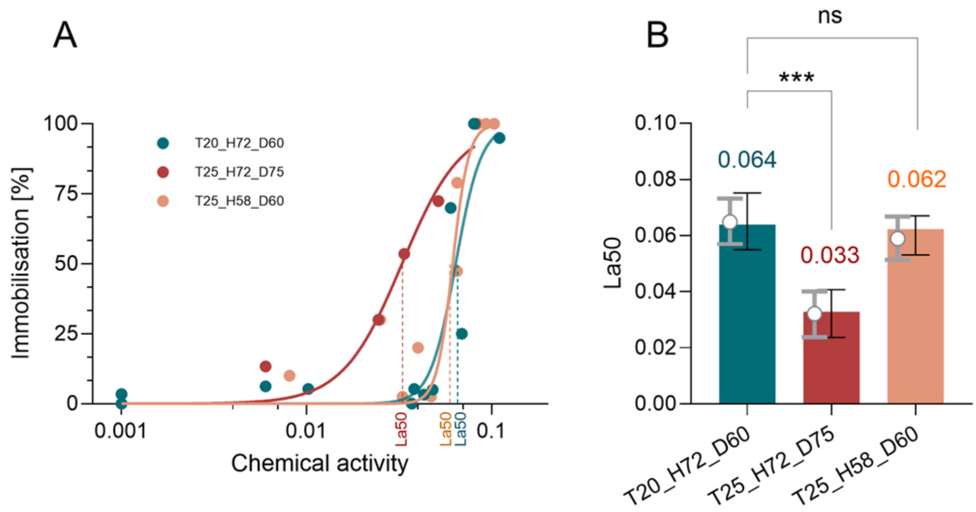


Figure 5. (A) Dose-response curves and (B) median lethal dose (La50; mean and 95 %–CI) for *Daphnia magna* immobilization rate (a proxy for mortality) exposed to the PAH mixture at chemical activity range of 0.01–0.1. The daphnids were grown at either 20 oC or 25 oC and exposed to the PAHs at the same temperature for three degree-days. At 20 oC, this corresponded to 72 h and at 25 oC, to 60 h. In addition, the daphnids from the 25 oC treatment were exposed to the same dose range for 72 h.

- *Combined exposure to PAHs and heavy metals.* When zinc (Zn) was applied at 2 mg/l, the immobilization was close to 80 % in the Zn-only treatment, thus precluding meaningful testing of the PAH/Zn mixture effect. At 1 mg/L Zn, the immobilization observed in the PAH/Zn treatment was similar to that produced by PAH alone. By contrast, the daphnid protein content was lower in the PAH/Zn mixture, suggesting a combined effect (Figure 6).

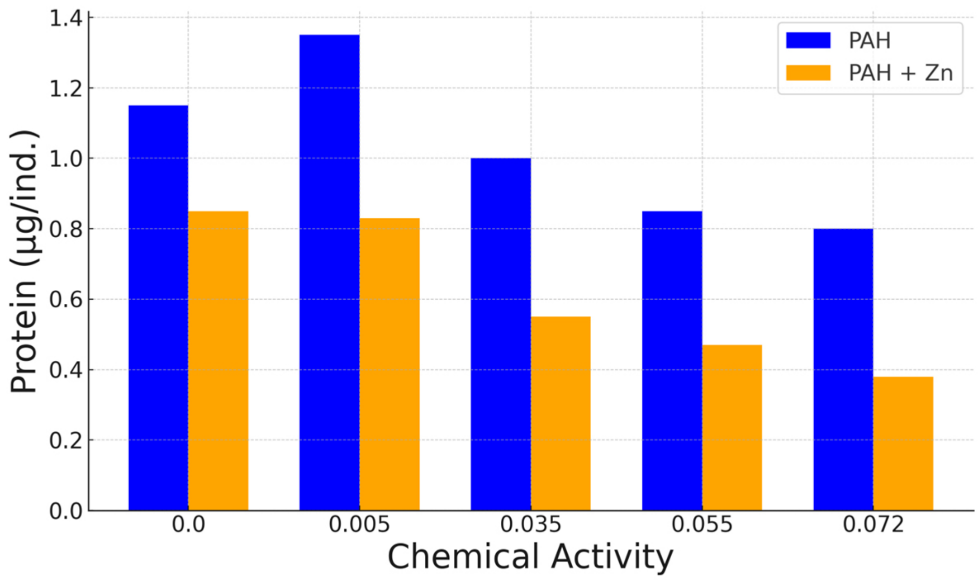


Figure 6. Response of individual protein content in *Daphnia magna* exposed to the PAH mixture at chemical activity range of 0.01–0.1 with or without zinc.

3.2 Effects of contaminants and abiotic factors on ReproIND

Three field studies contributed to the evaluation of contaminant effects and non-chemical factors on the reproductive and population characteristics of *Monoporeia affinis:*

1. *Raymond and co-workers* (Raymond et al., 2021) used monitoring data to identify the environmental predictors that best explained changes in benthic macrofaunal communities. In this dataset, PAH concentrations were classified as medium to high values according to national guidelines (Josefsson, 2017) at 4 stations out of 29; in addition, 8 stations exceeded the pristine level. Environmental predictors included sediment organic carbon (TOC), sediment concentrations of metals and PAHs, bottom water oxygen, salinity, temperature, and surface chlorophyll-a concentration. The abundance of *M. affinis*, one of the most common species in the area, was best explained by total PAH concentration as a single negative predictor. None of the metals was identified as an influential predictor.
   This finding suggests that PAH contamination can suppress the population size in sediment-living amphipods, which should be considered when benthos is used for eutrophication status assessment. This study also shows that at the reference stations sampled within the SNMMP (the macrobenthos monitoring and the aquatic environmental contaminants programs) along the Baltic Sea coast, concentrations of 11 priority PAH congeners in sediment often exceed the level characterized as “low” in the national guidelines (Josefsson, 2017), indicating that our biological effects data might overestimate the background aberration rate.
2. *Kolesova and co-workers* (Kolesova et al., 2024) analyzed data from four subbasins–the Gulfs of Finland and Riga, the Bothnian Sea, and the Western Gotland Basin–to investigate the relationship between embryo viability in *M. affinis*, HOCs (PAHs, PCBs and BTs; total concentrations, TEQ values and individual congeners present across study sites were tested as predictors) and metals (both PLI and individual metal concentrations were tested as predictors). The study’s broad geographic scope covering ecologically relevant gradients enabled a comprehensive integration of nonchemical factors, such as temperature, TOC, and oxygen, into a battery of multivariate models.
   Metals (PLI, Hg), PAHs (NAP and DBAHA), and PCBs (PCB180), were significant positive predictors of the embryo aberrations across the basins. Given the high levels of cross-correlations between different metals and PLI values as well as across congeners of PAHs and PCBs, it is likely that the detected effects are not specific to these compounds but reflected toxicity exerted by other congeners in the contaminant mixtures. Moreover, the aberration frequencies increased not only with the contaminant concentrations but also with temperature and salinity. Also, sand-dominating sediments with often low TOC were found to exert stronger toxicity than silt and clays.
   Notably, the temperature and salinity dependencies were only detectable for the entire range of the temperature values tested (T oC: 2 to 6 oC and Sal: 1.4 to 7.1; bottom layer in winter, i.e., the reproduction period of the amphipods). Within a basin, this variation was always substantially smaller and no effects across the sites was detectable. Therefore, it was concluded that having basin-specific

GES targets for reproductive aberrations in a population adapted to the local temperature and salinity fluctuations should be sufficient to account for these dependencies.

Classification of the sampling sites by cluster analysis revealed four clusters differing by the loads and compositions of co-occurring contaminants (Figure 7). Cluster 1 consisted of only two highly contaminated sites SU 57 and SU 58 (Sundsvall, Bothnian Sea), with very high levels of metals (As, Hg, Cu, and Ni), PAHs and PCBs. Cluster 2 consisted of five GoF sites with relatively high levels of BTs, whereas the other two clusters combined sites located in multiple sub-basins and characterized by diverse mixtures of PAHs, PCBs, and metals. Thus, none of the clusters represented non-polluted conditions.

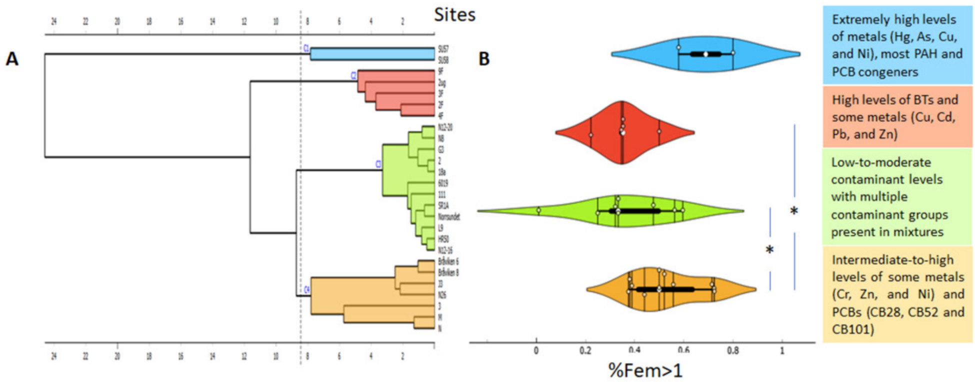


Figure 7. Variability of the reproductive aberrations across sites grouped according to their contaminant load. (A) Hierarchical clustering of the sampling sites according to the contaminant load (Ward linkage), yielding four clusters (C1 to C4), and (B) Violin plots for the proportion of the females carrying more than 1 aberrant embryo. The width of each curve corresponds with the approximate frequency of data points in each region. The densities are additionally annotated with the median value and interquartile range, shown as a black boxplot within each violin plot. Significant differences between the groups detected by the unpaired t-test are indicated by an asterisk ($p < 0.05$). Note that C1 cluster was comprised by only 2 stations from Sundsvall with extremely high contamination levels for both HOCs and metals.

3. *Gorokhova and co-workers* applied PLS-SEM that identified the key drivers of reproductive health (ReproHealth) in the Bothnian Sea (BS) and Western Gotland Basin (WGB) and established direct and indirect pressure-indicator relationships between embryo aberrations, chemical pollution (PAHs and metals), trophic properties of the amphipod populations and abiotic factors (Figure 8). The contaminant loads and non-chemical environments were substantially different between the basins. Contrary to our expectations, significantly higher contaminant levels (metals and PAHs) were observed in the BS than in WGB, with 1.5–to 6.3-fold differences for the metals and ~ 2-fold for both low- and high-molecular PAHs.

   Temperature had a significant moderating effect on ReproHealth in each sub-basin, i.e., there was a significant temperature-by-contaminants interaction for the EXPOSURE -> ReproHealth relationship. However, the effect direction was opposite for the WGB and BS models, with a strong negative effect in BS and a weak positive in WGB. This implies that the contaminant-induced toxicity was partially alleviated in BS and enhanced in WGB. Taking into account the significantly lower temperature in BS than in WGB and significantly higher levels of trace metals, one can speculate that within the optimal thermal range, the ameliorating effect of temperature might be related to more efficient detoxification by the animals as shown for fish exposed to heavy metals at different temperatures (Castaldo et al., 2021). In WGB, *M. affinis*, a glacial relict, is living close to its temperature limit, and this is why the negative effect of pollution is more likely to be amplified at higher temperatures.

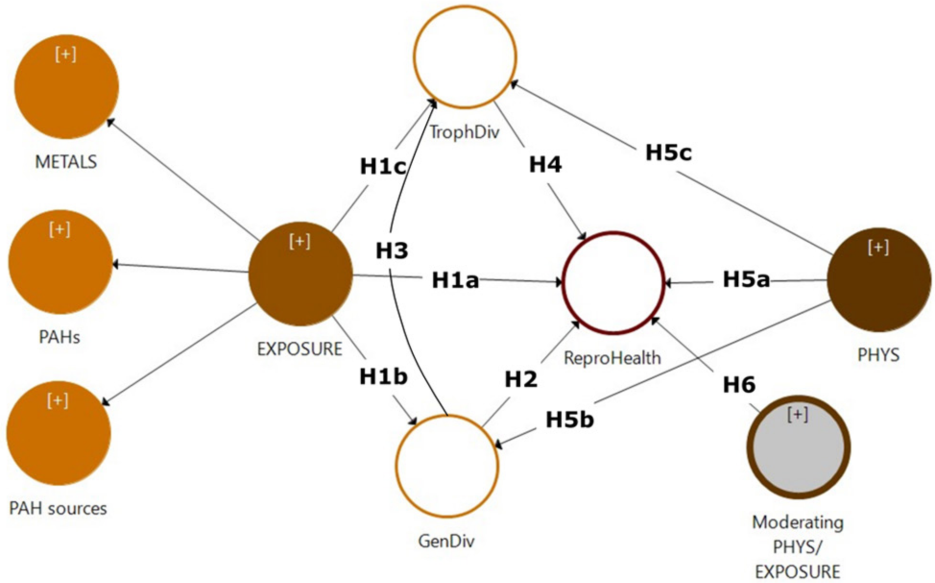


Figure 8. The PLS-SEM model linking reproductive aberrations (ReproHealth) to biotic (TrophDiv, trophic diversity, and GenDiv, genetic diversity) and abiotic (EXPOSURE, chemical contaminants in the sediment, and physical, PHYS, environmental conditions). The paths link different latent variables (LVs circles). The biological LVs are shown as white circles, the chemical (EXPOSURE) and non-chemical (PHYS) LVs are shown in brown, and the indicators (the measured parameters representing some aspects of specific LVs) are hidden (denoted as [+]) for clarity. The EXPOSURE model had a three first-order LVs shown in light brown and representing different types of contaminants and pollution indices.

3.3 DT analysis of GES targets

The decision tree model (Figure 9) highlighted the relationships between contaminant pressure based on the composite toxicity potential indices for metals (Metals-PLI) and PAHs (chemical activity) and GES status based on ReproIND outcomes. The model identified critical thresholds corresponding to Metals-PLI and chemical activity, which had the greatest influence on GES outcomes. Non-contaminant factors (bottom depth and salinity) were also identified as influential. These thresholds can guide adjustments to GES targets to enhance the detection of impacted environments.

First, the tree identifies a bottom depth of 75 m as a critical threshold influencing GES classification outcome, with increased non-GES conditions at > 75 m depth. It should be noted, however, that less than 12 % of stations in the dataset exceed this depth. The next split in the DT identifies a Metal-PLI threshold of < 0.2, suggesting that another 11 % of sites with low metal concentrations, primarily in the Neva Estuary, Gulf of Finland, are classified as non-GES.

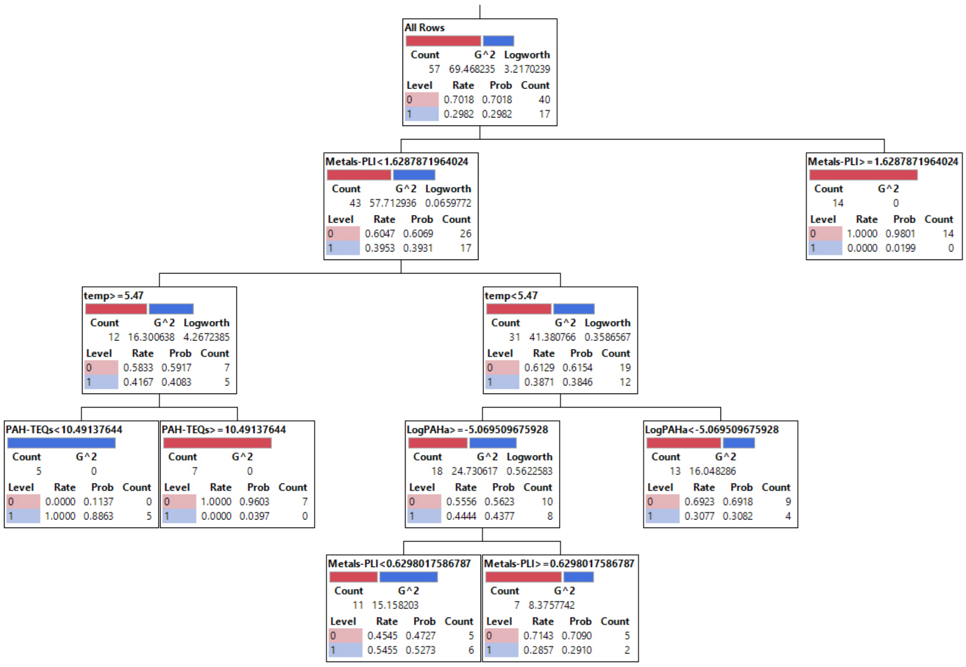


Figure 9. Decision Tree (DT) analysis of contaminants and environmental factors influencing GES derived from ReproIND values across 55 study sites used for model validation (red: non-GES, blue: GES). The decision tree model illustrates the relationship between contaminant pressure using composite variables for metals and PAHs (Metal-PLI and LogPAHa, chemical activity for PAHs, log-transformed) and GES status based on ReproIND outcomes. The model identifies critical contaminant thresholds, including LogPAHa (−5) and Metal-PLI (0.5) as key factors influencing GES outcomes, along with non-contaminant factors, bottom depth (75 m) and salinity (6.5). These findings suggest the necessity to adjust GES targets for reproductive aberrations, particularly in deeper sites, to enhance the detection of anthropogenic stressors and revise EQS for metals at salinities below 6.5.

At sites were Metals-PLI exceeds 0.2, both GES and non-GES conditions are represented. Here, chemical activity (LogPAHa) plays a significant role, with higher activity values associated with non-GES conditions (65 %). Metals are also a key factor, further classifying non-GES conditions at Metal-PLI levels exceeding 0.5. The combination of LogPAHa > −5 and Metal-PLI > 0.5 increases probability of non-GES to 86 %, which indicates the importance of different metal-PAH combinations to induce mixture effects with adverse biological impacts.

As the last split, salinity was identified as a modulating factor for sites with both high chemical activity and elevated metal pollution, with salinity < 6.5 more likely to result in non-GES status–an effect primarily driven by sites in the northernmost study area (Bothnian Bay) and the eastern Gulf of Finland. Other chemical pollutants may also contribute to the observed non-GES conditions in these low-salinity coastal areas.

The performance of the DT model was evaluated using accuracy, precision, and recall, focusing on identifying non-GES conditions. The receiver operating characteristic (ROC) curve showed an area under the curve (AUC) of 0.85, indicating significant discriminatory ability between GES and non-GES conditions. Metals-PLI was the most important predictor (0.38), followed by LogPAHa (0.28), depth (0.23) and salinity (0.15). Using stratified 5-fold cross-validation, the model achieved an accuracy of 88.2 %, with precision and recall for non-GES conditions being 0.67 and 0.77, respectively.

3.4 Pressure-indicator relationships

Despite the influence of environmental parameters that modulate contaminant effects, our analysis revealed significant relationships between contaminant load–when assessed through composite potential toxicity indices–and the GES classification, which was based on biological responses (Figures 9 and 10).

A group-wise comparison between GES and non-GES sites revealed significantly higher chemical activity (LogPAHa) and Metals-PLI values for the non-GES sites. Additionally, the non-GES group had significantly greater variability in both parameters (Figure 9). Further, GLM analysis revealed a significant positive relationship between chemical activity and TOC in the sediment (Figure 10), which is expected due to partitioning of PAHs to the organic sediment fraction. Moreover, the significantly different intercepts between the regression lines for GES and non-GES groups indicate higher PAH levels per unit of organic carbon in the sediment at the non-GES sites, i.e., when aberration frequency exceeds background levels. These findings demonstrate significant pressure-indicator relationships for ReproIND when pressure is represented by chemical activity based on PAH concentrations in sediment.

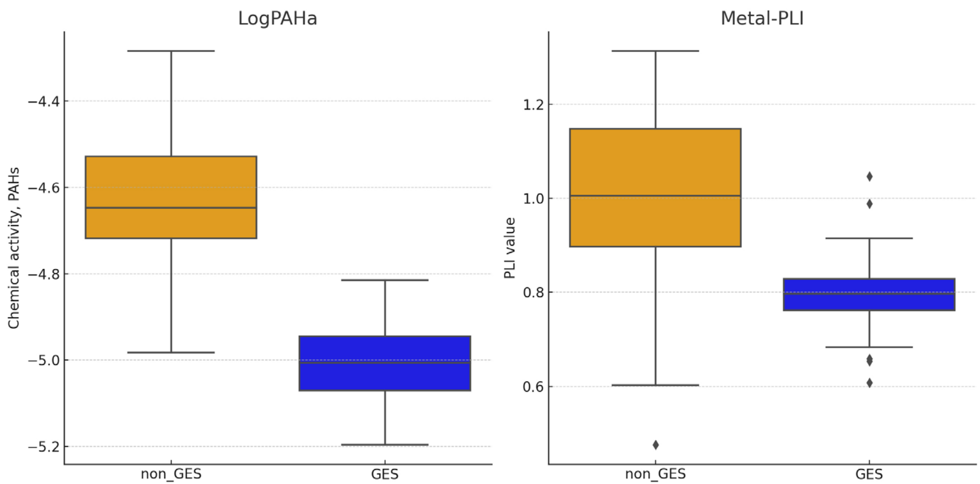


Figure 10. Variability in chemical activity (LogPAHa) and Metal-PLI values grouped for the sites classified as non-GES and GES. The output of the statistical comparison between the groups is shown below the x axis. The box-and-whisker plots show the mean, median (yellow line), 25 % - and 75 %, minimum, and maximum values for each parameter.

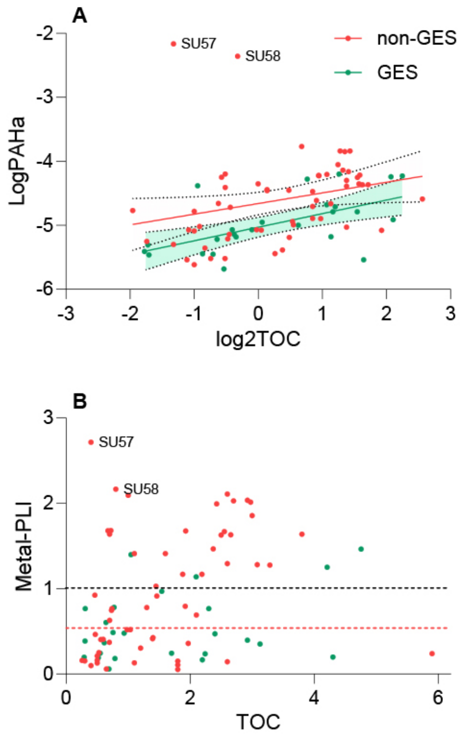


Figure 11. Relationships between TOC in sediment and (A) chemical activity (LogPAHa) and (B) metals (Metal-PLI) for sites classified as GES and non-GES. The linear regression is significant for LogPAHa, with significantly lower intercept in the non-GES group (Table 3). By contrast, Metal-PLI was not clearly related to TOC, albeit this regression was weakly significant for the non-GES sites (Table 3 B). The regression lines (panel A) are shown with their 95 %-confidence bands, and the horizontal lines corresponding to Metal-PLI of 1 and 0.5 (panel B) indicate levels generally accepted as potential toxicity threshold (PLI = 1; Tomlinson et al. 1980) and the level identified by our DT model (PLI = 0.5) as increasing likelihood of adverse effects when HOCs are present at chemical activities exceeding 0.00 001 (i.e., LogPAHa > –5).

Table 3. GLM output for GES effect on the relationship between chemical activity (LogPAHa) and TOC. Sigma-restricted parameterization, Effective hypothesis decomposition; Std. Error of Estimate: 0.6 055.

	SS	MS	F	p
Intercept	1540.348	1540.348	4 201.674	0.000 000
GES	2.391	2.391	6.521	0.012 582
LogTOC	3.231	3.231	8.814	0.003 957
GES*LogTOC	0.066	0.066	0.180	0.672 292
Error	28.962	0.367		

4. Discussion

The overarching goals of this project were to advance our understanding of the cumulative effects and multiple stressors on the biological impacts of environmental contaminants, with a particular focus on improving the environmental status assessment in the Baltic Sea. By integrating advanced chemical toxicity metrics, biological effect indicators, and accounting for nonchemical factors, we achieved several key milestones:

- enhanced the application of chemical activity as a reliable dose metric in ecotoxicity testing,
- developed a workflow for validating biological effect indicators and demonstrated significant pressure-indicator relationships for ReproIND,
- identified thresholds for co-occurring PAHs and metals associated with non-GES conditions (i.e., mixture effect outcomes),
- established thresholds for environmental variables that influence ReproIND responses to contaminants,
- proposed a novel approach for integration of chemical and biological effects data for contaminant status assessment in the Baltic Sea, with potential applicability in a variety of ecosystems.

These achievements contribute to contaminant assessment and environmental monitoring in marine ecosystems. They also support ReproIND operationalization in the Baltic Sea.

4.1 Chemical activity in ecotoxicity testing

Our experimental studies have advanced the use of chemical activity as a dose metric in HOC ecotoxicology, providing new protocols and dose-response data for both water-borne and sediment exposures with planktonic and benthic test species, compatible with current OECD guidelines. The method development supports robust experimental designs to account for chemical and non-chemical confounding factors, highlighting the need for standard guidelines updates. Also, the proposed sediment exposure techniques offer improved control over chemical activity levels, while adjustments for substrate diversity to ensure more realistic and accurate toxicity testing need more research. Our development of the chemical activity-based loading method demonstrated differences in capacity of different peat batches to sorb PAHs, implying that the same amount of chemical loaded to different peats for bioassay testing, can generate different exposure levels (Abel et al., 2024), emphasizing the need to move away from assessments based on total concentrations. Variability in organic matter properties has effects on bioavailability and potential release of chemicals from sediment to water also in the environment. This was recently demonstrated in the Gulf of Finland, where two sites with varying proportion of terrestrial versus marine organic carbon in the sediment were investigated. Hydrophobic PAHs and PCBs were found to sorb stronger to the sediment of the marine OC-dominated site with implications on bioavailability (Nybom et al., 2021).

The observed median toxicity doses in the acute immobilization tests fell within the expected chemical activity range for baseline toxicity (0.01–0.1), supporting narcosis, a non-specific toxic action, as the primary mode of action in tests with daphnids and earthworms. Additionally, the absence of immobilization at chemical activity levels below 0.01 further supports the threshold nature of narcosis, where a certain level of activity is required to elicit a response. Thus, chemical activity is a reliable predictor of baseline toxicity in ecotoxicological assessments with standard invertebrate models.

No toxicity, indicative of a specific toxic mode of action for PAHs, was observed below this threshold (including pilot trials where the chemical activity range was 0.001–0.12), reinforcing narcosis as the dominant mode of action for PAHs at high concentrations with immobilisation as the endpoint. However, the dose-response curves for protein content in *Daphnia* suggest that these responses occur at much lower chemical activity levels–at least as low as 10-fold below the expected narcosis threshold–indicating that a different mode of action is likely responsible. Thus, specific toxic mechanisms operate at lower PAH concentrations, and, possibly, induced by specific congeners, can trigger molecular or cellular responses before observable effects like immobilization occur.

Our experimental findings on the influence of nutritional status, temperature, and combined exposures on toxicity outcomes are supported by the PLS-SEM model based on the field data (Gorokhova et al., 2024). In addition to their ecological importance, they also highlight the multifaceted nature of factors that vary in experimental settings, sometimes unintentionally. For instance, the nutritional status of daphnids significantly altered their sensitivity to PAHs, with starved individuals showing higher tolerance levels compared to well-fed ones. This suggests that energy reserves and metabolic state of the test organisms can critically modulate toxic responses, making considering these variables in toxicity assessments essential.

Similarly, temperature was found to impact toxicity in ways that go beyond simple dose-response relationships. Exposure at higher temperatures result in more advanced physiological aging of organisms, which significantly impacts test outcomes due to the extended physiological time spent in the exposed environment, rather than increased toxicity. This underscores the importance of adjusting exposure times to degree days and interpreting thermal conditions in ecotoxicological studies to avoid drawing misleading conclusions about temperature-driven toxicity of specific substances.

Finally, the combined exposure to different classes of contaminants, such as PAHs and heavy metals (zinc, in our case), introduced additional layers of complexity. The interactions between these chemicals did not always follow simple additive or synergistic patterns, with some toxic effects manifesting only at specific combinations and detectable using suborganismal endpoints. This variability underscores the need for ecological risk assessments to move beyond single-chemical evaluations and instead embrace a more integrated approach that accounts for the mixture effects.

4.2 Evidence for mixture effects from the field studies

Chemical mixtures, particularly those involving PAHs and metals, significantly affect the reproductive and population health of *Monoporeia affinis*, with effects further modulated by temperature, salinity, depth, and sediment composition. The relative contributions of PAHs and metals as predictors differed across the studies, depending on the nature of the response, e.g., abundance (Raymond et al., 2021), different aspects of embryo aberrations (Kolesova et al., 2024), and GES outcome (binary variable) based on the aberration frequencies (Gorokhova et al., 2024). Moreover, in other studies relevant to INSERT but not directly conducted within this project, DNA adductome was implicated as a mechanism of developmental toxicity (Gorokhova et al., 2020), with specific adducts being associated with metals (Hg, Cd, and Zn) and the overall PAH load, but not any specific congeners (Martella et al., 2023). Therefore, applying chemical toxicity as an additive metric to integrate multiple PAHs into a single value is justified by the field observations.

While laboratory studies demonstrate that toxicity was primarily due to narcosis, with no measurable acute responses observed at chemical activities below 0.01, the chemical activity levels observed in the field suggest low relevance of these effects for field settings because chemical activity in the sediments was generally several orders of magnitude below the experimental values (Figure 10). However, despite such low values, frequent non-GES conditions were observed at chemical activities above 0.00 001 (Figure 9 A), likely associated with specific modes of action. We also should keep in mind that our chemical activity calculations focus solely on PAH activity and do not account for other HOCs that may also contribute to the observed effects. Given these limitations, we cannot conclude whether the field responses are due to the excess toxicity of PAHs, even though PAHs are known endocrine disrupters in arthropods, representing such effects (Oberdörster et al., 1999). Excess toxicity screening should be conducted through controlled laboratory experiments because field conditions involve multiple stressors and contaminants, making it difficult to isolate the effects of a single chemical. Furthermore, the precise levels and durations of exposures in situ are often unknown, unlike in the well-controlled experimental studies. These factors underscore the need for laboratory-based studies to accurately identify and quantify excess toxicity beyond baseline narcosis (Kienzler et al., 2017).

4.3 ReproIND responses to regional variability in contaminant and environmental conditions

The decision tree model (Figure 8) highlights key thresholds and interactions between contaminants and environmental factors influencing GES classification across 71 sites in the five Baltic Sea basins. It identifies critical levels of PAHs (expressed as chemical activity) and metals (expressed as their combined potential toxicity). The key finding of the model is that sites with both chemical activity exceeding 0.00 001 and metal pollution (PLI > 0.5) were more likely to be classified as non-GES (Figure 8). As metal toxicities are typically anticipated when the PLI exceeds 1 (Tomlinson et al., 1980), this finding is important because it suggests that a more conservative threshold might be needed. Moreover, more-than-additive effects are common for metal–PAH mixtures (Gauthier et al., 2014). The key interactions that can lead to more-than-additive toxicity in these mixtures include PAHs disrupting membrane integrity and thus facilitating metal uptake, potential metal–PAH complexation, metals inhibiting PAH detoxification via cytochrome P450, PAHs impairing metal detoxification via metallothionein, and enhanced reactive oxygen species production in certain metal–PAH combinations. Additionally, the mutual inhibition of detoxification processes suggests possible positive feedback among these mechanisms (Gauthier et al., 2014). All these mechanisms are particularly relevant during embryogenesis and can thus contribute to the observed interaction between the critical thresholds for chemical activity and PLI for the ReproIND outcome detected by the model.

The model has also identified bottom depth at the sampling site and salinity, as significant determinants of GES outcomes; hence, adjustments to GES targets for deeper sites and environmental quality standards (EQS) for metals in low-salinity areas are needed to better address the contaminant effects. The high likelihood of non-GES conditions in areas deeper than 75 m can be related to spatial variability in the contaminant distributions resulting in accumulation of contaminants (other than PAHs and metals) with depth. As these contaminants were not included in PLI and chemical activity calculations, their effects were not addressed by the model. It is also possible that hydrographic conditions or any other natural characteristics of the habitat in the deeper areas are suboptimal for *Monoporeia* reproduction. If this is the case, the depth should be considered when designing effects screening studies and monitoring.

The practical implication of this outcome is to adjust the background values for reproductive aberrations in *M. affinis* by examining the long-term variability of embryo malformations in the monitoring data from these areas. Consequently, the adjusted background values will result in revising GES targets for ReproIND and increased capacity to detect pollution-induced aberrations. In parallel, more ecophysiological information for the depth-related effects on *Monoporeia* reproduction should be gathered to understand the ecological drivers involved.

In addition, salinity emerged as a significant modulating factor affecting sites with elevated metal pollution (PLI > 0.5). Specifically, sites with salinity below 6.5 were more likely to be classified as non-GES, a pattern pronounced in the northernmost study area (Bothnian Bay) and the eastern Gulf of Finland. It is well-acknowledged that low salinity conditions can exacerbate the impacts of metal pollution due to increased metal bioavailability, altered metal speciation, and heightened physiological stress (Acosta et al., 2011; Du Laing et al., 2008). As a result, coastal and estuarine ecosystems with lower salinity are particularly vulnerable, especially in areas with significant freshwater inflow or those affected by climate change (Leal Filho et al., 2022). Therefore, environmental managers should consider salinity a key factor when assessing pollution risks in the Baltic Sea and explore the validity of 6.5 as the threshold.

Finally, about 10 % of sites with low Metal-PLI (< 0.2) were classified as non-GES, implying that the only common feature for these sites was low levels of metals, particularly those with high specific toxicity (e.g., As, Hg, and Cd). However, other drivers may be influencing GES outcomes in these sites. For example, upon a closer examination, one can see that these sites are located exclusively in the Gulf of Finland and many of them are in the harbours. In the original studies, elevated levels of butyltins were reported coinciding with relatively low metal pollution and, in some cases elevated PCB levels (Kolesova et al., 2024). These are also the stations with the lowest PAH contribution to the chemical activity due to the highest PCB concentrations (Figure 2). Neither butyltins nor PCBs were included in the DT dataset due to the limited availability of these measurements or the high frequency of nondetects in other basins. However, these factors could still contribute to non-GES conditions at these sites, with the low PLI values potentially being coincidental rather than the primary drivers of the observed responses. Further investigation with more comprehensive chemical data is needed to clarify these counterintuitive results for the areas classified as non-GES but with low contaminants (metals and PAHs).

5. Conclusions and recommendations

5.1 Further developments in ReproIND implementation and GES setting

Our studies focusing on individual contaminants and composite indices (chemical activity for HOCs and PLI for metals) provide basis for revising ReproIND targets to reflect environmental realities, thereby improving its capacity to assess the contaminant impacts. Here, we present our recommendations for this revision (Figure 11):

- *Establish depth-specific GES targets for ReproIND:* Aggregate ReproIND data by bottom depth categories (> 75 m and < 75 m) and calculate 95 %-confidence intervals for the frequency of aberrant embryos and frequency of females carrying aberrant embryos. Include the new targets to the HOLAS IV assessment for more accurate assessments of biological effects based on site-specific depth conditions.
- *Determine causes of non-GES conditions that are not related to PAH and metal toxicity in the Gulf of Finland:* Conduct non-target analysis of chemical contaminants at the sites with low chemical activity and PLI yet classified as non-GES and supplement these data with bioassays to confirm biological effects.
- *Prioritize the assessment of sites where Metals-PLI exceeds 0.5 and chemical activity exceeds 0.00 001.* It is crucial to monitor and manage these metrics in concert because the combination of elevated metal and HOCs significantly increases the likelihood of non-GES outcomes.
- *Salinity consideration in pollution assessments:* Incorporate salinity as a critical factor in the evaluation of metal pollution risks, particularly in low-salinity coastal and estuarine ecosystems where metal toxicity may be enhanced. For example, revise and adjust EQS values for metals to ensure that these standards are appropriate for the low-salinity (< 6.5) areas in the Baltic Sea.

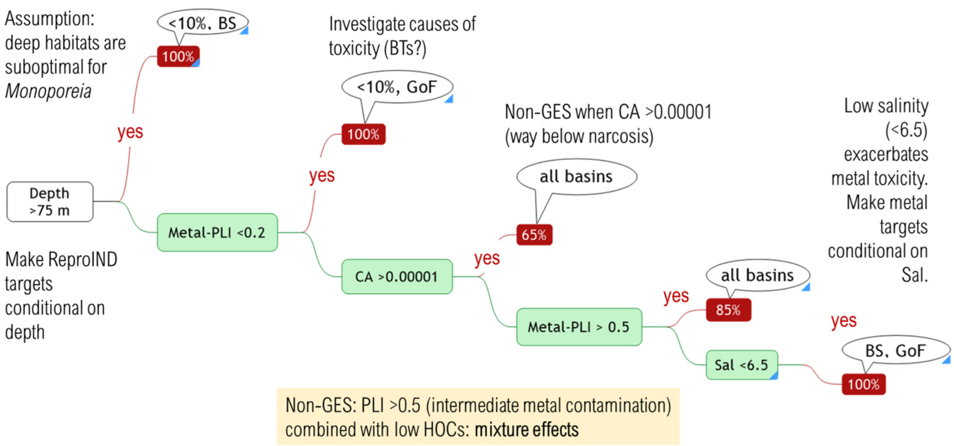


Figure 12. Summary of key decision points identified by the decision tree model (Figure 8) for assessing GES based on ReproIND outcomes. The figure highlights critical environmental thresholds, including bottom depth, Metals-PLI, and chemical activity (CA), that influence GES classification. Recommendations based on these decision points emphasize the importance of stratifying ReproIND data by site depth, prioritizing monitoring at sites with Metals-PLI > 0.5, addressing the combined effects of metal pollution and HOCs, particularly at high chemical activity levels, and accounting for salinity effects on metal toxicity.

5.2 Linking chemical and biological indicators

Our findings demonstrate significant pressure-indicator relationships for ReproIND when pressure is represented by chemical activity based on PAH concentrations in sediment (Figures 9 and 10). Based on these relationships, we propose a method for the integration of chemical and biological indicators within D8/MSFD to monitor and assess the contaminant impacts.

The two linear regressions between chemical activity and organic carbon content in sediment, differing in their intercepts (Figure 10), provide a framework for screening sites for potential pollution impacts. Since the slopes are identical, the key differentiator is the intercept, with the higher intercept representing impacted sites (GES) and the lower intercept representing non-impacted sites (non-GES).

When HOCs and organic carbon content measurements are conducted by the contaminant monitoring, the chemical activity can be calculated and the resulting data point can be statistically evaluated to determine which regression line it most likely belongs to. The estimated confidence intervals for each regression make it possible to assess the probability that a given data point is associated with either the GES or non-GES regression. This evaluation of contaminant data enables environmental managers to categorize sites based on the likelihood of biological effects and to prioritize areas for further investigation or remediation based on the evidence provided by the classification.

To ensure the relevance and accuracy of the classifications derived using this method, biological effect assessment based on ReproIND but even other biological effect indicators, can be integrated into this process as a validation. Here's how this could be done:

- *Cross-validation with ReproIND data:* After categorizing sites as potentially impacted or non-impacted based on the chemical activity-TOC regression as described above, ReproIND data are used to cross-validate these classifications. For example, sites classified as impacted should show corresponding biological effects, such as reproductive abnormalities in amphipods, as indicated by ReproIND scores. Conversely, sites classified as non-impacted should exhibit minimal or no reproductive disturbances.
- *Refining categories:* If discrepancies are observed–such as a site classified as non-impacted but displaying significant biological effects–this could indicate that the regression-based classification needs adjustment. Such inconsistencies could prompt an investigation of the pollution status of the site in question or a refinement of the regression model, potentially leading to the identification of additional influencing factors that were not previously accounted for.
- *Feedback loop for model improvement:* The integration of ReproIND data can create a feedback loop, where biological effect observations are continuously used to refine the regression model. Over time, this approach can improve the predictive power of the model, making it more sensitive to variations in chemical activity that are ecologically relevant.
- *Setting thresholds for action:* Combining the chemical activity-TOC classifications with ReproIND outcomes allows for the establishment of more precise thresholds for management action. For example, if a certain level of chemical activity consistently correlates with adverse ReproIND outcomes, this threshold can be used as a trigger point for initiating environmental interventions.

Chemical activity as an additive metric streamlines the process of setting regulatory limits and assessing environmental quality. Moreover, measuring a broader range of HOCs provides a more accurate estimate of overall chemical activity and potential toxic effects. By using this unified measure and exploring other composite indices of potential toxicity, regulatory agencies can more effectively evaluate cumulative risks from contaminant mixtures and establish appropriate safety thresholds. Integrating the chemical activity-TOC regression model with biological assessments like ReproIND ensures that pollution impact classifications are both statistically robust and ecologically meaningful. This approach strengthens the connection between chemical pressures and biological responses, supporting habitat classification systems that incorporate both chemical and biological criteria for more accurate assessments of pollution impacts on biota.

6. Acknowledgements

This work was funded by the Swedish Environmental Protection Agency (Naturvårdsverket) within the INSERT project. Additional support was provided by FORMAS (project ACTA, 2019-00863), Interreg Baltic Sea Region co-funded by the European Union (project #S007, 10/2022–10/2024), and the European Biodiversity Partnership Biodiversa+ (project D2P, proposal number 2021-473).

Technical assistance with chemical analyses and protein measurements by Z. Li and B. Liewenborg (ACES, Stockholm University) is gratefully acknowledged. We also thank Dr. Sandra Luecke-Johansson for her help with *Daphnia*-related work.

The HELCOM expert group on hazardous substances (EG-HAZ) provided valuable discussions on environmental contaminants, and the sub-group on biological effects of contaminants contributed important feedback on the indicator relevance for environmental assessment.

7. References

Abel, S., Eriksson Wiklund, A.-K., Gorokhova, E., Sobek, A., 2024. Chemical Activity-Based Loading of Artificial Sediments with Organic Pollutants for Bioassays: A Proof of Concept. Environ. Toxicol. Chem. 43, 279–287. https://doi.org/10.1002/etc.5788

Acosta, J.A., Jansen, B., Kalbitz, K., Faz, A., Martínez-Martínez, S., 2011. Salinity increases mobility of heavy metals in soils. Chemosphere 85, 1318–1324. https://doi.org/10.1016/j.chemosphere.2011.07.046

Arboleda, M.O., 2024. Sorption kinetics of 4 PAHs in different batches of peat and use of passive sampling to measure target chemical activities (MSc Thesis). Stockholm University, Stockholm, Sweden.

Beyer, J., Petersen, K., Song, Y., Ruus, A., Grung, M., Bakke, T., Tollefsen, K.E., 2014. Environmental risk assessment of combined effects in aquatic ecotoxicology: A discussion paper. Mar. Environ. Res. 96, 81–91. https://doi.org/10.1016/j.marenvres.2013.10.008

Castaldo, G., Pillet, M., Ameryckx, L., Bervoets, L., Town, R.M., Blust, R., De Boeck, G., 2021. Temperature Effects During a Sublethal Chronic Metal Mixture Exposure on Common Carp (Cyprinus carpio). Front. Physiol. 12.

Du Laing, G., De Vos, R., Vandecasteele, B., Lesage, E., Tack, F.M.G., Verloo, M.G., 2008. Effect of salinity on heavy metal mobility and availability in intertidal sediments of the Scheldt estuary. Estuar. Coast. Shelf Sci. 77, 589–602. https://doi.org/10.1016/j.ecss.2007.10.017

Förlin, L., Sundelin, B., Gorokhova, E., Magnusson, M., Bergkvist, J., Parkkonen, J., Larsson, Å., Liewenborg, B., Franzén, F., 2019. Effektscreening–Biologisk effektövervakning i förorenade områden längs Sveriges kust 2017–2018 (Nationell miljöövervakning på uppdrag av Naturvårdsverket).

Gauthier, P.T., Norwood, W.P., Prepas, E.E., Pyle, G.G., 2014. Metal–PAH mixtures in the aquatic environment: A review of co-toxic mechanisms leading to more-than-additive outcomes. Aquat. Toxicol. 154, 253–269. https://doi.org/10.1016/j.aquatox.2014.05.026

Gobas, F.A.P.C., Mayer, P., Parkerton, T.F., Burgess, R.M., Meent, D. van de, Gouin, T., 2018. A chemical activity approach to exposure and risk assessment of chemicals. Environ. Toxicol. Chem. 37, 1235–1251. https://doi.org/10.1002/etc.4091

Gorokhova, E., 2024. Integrating composite toxicity potential indices with biological indicators: Advancing contaminant status assessment under the MSFD. bioRxiv.

Gorokhova, E., Ledesma, M.G., Karlsson, A.M.L., Garbaras, A., Ketmaier, V., Sundelin, B., 2024. Understanding amphipod reproductive health drivers in the Baltic Sea: A Study Based on Structural Equation Modeling. Authorea. https://www.authorea.com/users/243058/articles/575500-understanding-amphipod-reproductive-health-drivers-in-the-baltic-sea-a-study-based-on-structural-equation-modeling

Gorokhova, E., Martella, G., Motwani, N.H., Tretyakova, N.Y., Sundelin, B., Motwani, H.V., 2020. DNA epigenetic marks are linked to embryo aberrations in amphipods. Sci. Rep. 10. https://doi.org/10.1038/s41598-020-57465-1

Gorokhova et al., 2023. Linking biological effects to contaminant levels in sediment, fish and benthic invertebrates along the Swedish coast, Report, Swedish Environmental Protection Agency, 86 pp.; https://urn.kb.se/resolve?urn=urn:nbn:se:naturvardsverket:diva-11066 (No. oai:DiVA.org:naturvardsverket-11066).

HELCOM, 2023a. Reproductive disorders: malformed embryos of amphipods. HELCOM supplementary indicator report. Online. [WWW Document]. Reprod. Disord. URL https://indicators.helcom.fi/indicator/reproductive-disorders/ (accessed 6.19.24).

HELCOM, 2023b. HELCOM Thematic assessment of hazardous substances, marine litter, underwater noise and non-indigenous species 2016-2021. Baltic Sea Environment Proceedings n°190; https://helcom.fi/post_type_publ/holas3_haz

HELCOM, 2018. HELCOM Thematic assessment of hazardous substances 2011-2016. Available at: http://www.helcom.fi/baltic-sea-trends/holistic-assessments/state-of-the-baltic-sea-2018/reports-and-materials

Holmes, C.M., Brown, C.D., Hamer, M., Jones, R., Maltby, L., Posthuma, L., Silberhorn, E., Teeter, J.S., Warne, M.S.J., Weltje, L., 2018. Prospective aquatic risk assessment for chemical mixtures in agricultural landscapes. Environ. Toxicol. Chem. 37, 674–689. https://doi.org/10.1002/etc.4049

Jahnke, A., Mayer, P., McLachlan, M.S., Wickström, H., Gilbert, D., MacLeod, M., 2014. Silicone passive equilibrium samplers as ‘chemometers’ in eels and sediments of a Swedish lake. Environ. Sci. Process. Impacts 16, 464–472. https://doi.org/10.1039/C3EM00589E

Josefsson, S., 2017. Klassning av halter av organiska föroreningar i sediment (No. 2017:12), Geological Survey of Sweden (SGU). Uppsala.

Kienzler, A., Barron, M.G., Belanger, S.E., Beasley, A., Embry, M.R., 2017. Mode of Action (MOA) Assignment Classifications for Ecotoxicology: An Evaluation of Approaches. Environ. Sci. Technol. 51, 10203–10211. https://pubs.acs.org/doi/10.1021/acs.est.7b02337

Kolesova, N., Sildever, S., Strode, E., Berezina, N., Sundelin, B., Lips, I., Kuprijanov, I., Buschmann, F., Gorokhova, E., 2024. Linking contaminant exposure to embryo aberrations in sediment-dwelling amphipods: a multi-basin field study in the Baltic Sea. Ecol. Indic. 160, 111837. https://doi.org/10.1016/j.ecolind.2024.111837

Leal Filho, W., Nagy, G.J., Martinho, F., Saroar, M., Erache, M.G., Primo, A.L., Pardal, M.A., Li, C., 2022. Influences of Climate Change and Variability on Estuarine Ecosystems: An Impact Study in Selected European, South American and Asian Countries. Int. J. Environ. Res. Public. Health 19, 585. https://www.mdpi.com/1660-4601/19/1/585

Ledesma, M., Gorokhova, E., Nybom, I., Sobek, A., Ahlström, D., Garbaras, A., Karlson, A.ML., 2024. Does pre-exposure to polluted sediment affect sub-cellular to population-level responses to contaminant exposure in a sentinel species? Environ. Pollut. 341, 122882. https://doi.org/10.1016/j.envpol.2023.122882

Löf, M., Sundelin, B., Bandh, C., Gorokhova, E., 2016. Embryo aberrations in the amphipod *Monoporeia affinis* as indicators of toxic pollutants in sediments: A field evaluation. Ecol. Indic. 60, 18–30. https://doi.org/10.1016/j.ecolind.2015.05.058

Mackay, D., Arnot, J.A., Wania, F., Bailey, R.E., 2011. Chemical activity as an integrating concept in environmental assessment and management of contaminants. Integr. Environ. Assess. Manag. 7, 248–255. https://academic.oup.com/ieam/article-abstract/7/2/248/7741324?redirectedFrom=fulltext

Martella, G., Gorokhova, E., Sousa, P.F.M., Tretyakova, N.Y., Sundelin, B., Motwani, H.V., 2023. DNA Adductomics for the Biological Effect Assessment of Contaminant Exposure in Marine Sediments. Environ. Sci. Technol. 57, 10591–10603. https://pubs.acs.org/doi/10.1021/acs.est.3c00499

Nybom, I., Horlitz, G., Gilbert, D., Berrojalbiz, N., Martens, J., Arp, H.P.H., Sobek, A., 2021. Effects of Organic Carbon Origin on Hydrophobic Organic Contaminant Fate in the Baltic Sea. Environ. Sci. Technol. 55, 13061–13071. https://pubs.acs.org/doi/10.1021/acs.est.1c04601

Oberdörster, E., Cottam, D.M., Wilmot, F.A., Milner, M.J., McLachlan, J.A., 1999. Interaction of PAHs and PCBs with Ecdysone-Dependent Gene Expression and Cell Proliferation. Toxicol. Appl. Pharmacol. 160, 101–108. https://doi.org/10.1006/taap.1999.8745

OECD, 2023. Test No. 218: Sediment-Water Chironomid Toxicity Using Spiked Sediment. Organisation for Economic Co-operation and Development, Paris.

OECD, 2007. Test No. 225: Sediment-Water Lumbriculus Toxicity Test Using Spiked Sediment. Organisation for Economic Co-operation and Development, Paris.

OECD, 2004. Test No. 202: *Daphnia* sp. Acute Immobilisation Test. https://doi.org/10.1787/9789264069947-en

Oliveira dos Anjos, T.B., Abel, S., Lindehoff, E., Bradshaw, C., Sobek, A., 2023. Assessing the effects of a mixture of hydrophobic contaminants on the algae Rhodomonas salina using the chemical activity concept. Aquat. Toxicol. 265, 106742. https://doi.org/10.1016/j.aquatox.2023.106742

Raymond, C., Gorokhova, E., Karlson, A.M., 2021. Polycylic aromatic hydrocarbons have adverse effects on benthic communities in the Baltic Sea: implications for environmental status assessment. Front. Environ. Sci. 9. https://doi.org/10.3389/fenvs.2021.624658

Redman, A.D., Parkerton, T.F., Leon Paumen, M., Butler, J.D., Letinski, D.J., den Haan, K., 2017. A re-evaluation of PETROTOX for predicting acute and chronic toxicity of petroleum substances. Environ. Toxicol. Chem. 36, 2245–2252. https://doi.org/10.1002/etc.3744

Saladin, Y., 2023. Is protein content a useful endpoint in acute ecotoxicological testing? (MSc Thesis). Stockholm University, Stockholm, Sweden.

Smith, K.E.C., Jeong, Y., 2021. Passive Sampling and Dosing of Aquatic Organic Contaminant Mixtures for Ecotoxicological Analyses. Environ. Sci. Technol. 55, 9538–9547. https://pubs.acs.org/doi/10.1021/acs.est.0c08067

Sokolov, A., Andrejev, O., Wulff, F., Rodriguez Medina, M., 1997. The Data Assimilation System for Data Analysis in the Baltic Sea, Systems Ecology Contributions, 3. Stockholm University, Stockholm, Sweden.

Steigerwald, S., Saladin, Y., Alurralde, G., Abel, S., Sobek, A., Eriksson Wiklund, A.K., Gorokhova, E., 2024. Enhanced tolerance to narcosis in starved *Daphnia magna* neonates. Environ. Toxicol. Chem. accepted.

Strode, E., Jansons, M., Purina, I., Balode, M., Berezina, N.A., 2017. Sediment quality assessment using survival and embryo malformation tests in amphipod crustaceans: The Gulf of Riga, Baltic Sea AS case study. J. Mar. Syst. 172, 93–103. https://doi.org/10.1016/j.jmarsys.2017.03.010

Sundelin, B., Eriksson, A.-K., 1998. Malformations in embryos of the deposit-feeding amphipod *Monoporeia affinis* in the Baltic Sea. Mar. Ecol. Prog. Ser. 171, 165–180.

Swain, C.K., 2024. Environmental pollution indices: a review on concentration of heavy metals in air, water, and soil near industrialization and urbanisation. Discov. Environ. 2, 5. https://doi.org/10.1007/s44274-024-00030-8

Tomlinson, D.L., Wilson, J.G., Harris, C.R., Jeffrey, D.W., 1980. Problems in the assessment of heavy-metal levels in estuaries and the formation of a pollution index. Helgoländer Meeresunters. 33, 566–575. https://doi.org/10.1007/BF02414780

Van den Berg, M., Birnbaum, L., Bosveld, A.T., Brunström, B., Cook, P., Feeley, M., Giesy, J.P., Hanberg, A., Hasegawa, R., Kennedy, S.W., Kubiak, T., Larsen, J.C., van Leeuwen, F.X., Liem, A.K., Nolt, C., Peterson, R.E., Poellinger, L., Safe, S., Schrenk, D., Tillitt, D., Tysklind, M., Younes, M., Waern, F., Zacharewski, T., 1998. Toxic equivalency factors (TEFs) for PCBs, PCDDs, PCDFs for humans and wildlife. Environ. Health Perspect. 106, 775–792.

Warne, M.S.J., 2003. A Review of the Ecotoxicity of Mixtures, Approaches to, and Recommendations for, their Management. Proc. Fifth Natl. Workshop Assess. Site Contam.

8. Publications and data availability

8.1 Peer-reviewed publications

Abel, S., Eriksson Wiklund, A.-K., Gorokhova, E., Sobek, A., 2024. Chemical Activity-Based Loading of Artificial Sediments with Organic Pollutants for Bioassays: A Proof of Concept. Environ. Toxicol. Chem. 43, 279–287. https://academic.oup.com/etc/article/43/2/279/7728709 (Abel et al., 2024)

Kolesova, N., Sildever, S., Strode, E., Berezina, N., Sundelin, B., Lips, I., Kuprijanov, I., Buschmann, F., Gorokhova, E., 2024. Linking contaminant exposure to embryo aberrations in sediment-dwelling amphipods: a multi-basin field study in the Baltic Sea. Ecol. Indicators. 160, 111837. https://www.sciencedirect.com/science/article/pii/S1470160X24002942?via%3Dihub (Kolesova et al., 2024)

Raymond, C., Gorokhova, E., Karlson, A.M., 2021. Polycylic aromatic hydrocarbons have adverse effects on benthic communities in the Baltic Sea: implications for environmental status assessment. Front. Environ. Sci. 9. https://www.frontiersin.org/journals/environmental-science/articles/10.3389/fenvs.2021.624658/full (Raymond et al., 2021)

Ledesma, M., Gorokhova, E., Nybom, I., Sobek, A., Ahlström, D., Garbaras, A., Karlson, A.ML., 2024. Does pre-exposure to polluted sediment affect sub-cellular to population-level responses to contaminant exposure in a sentinel species? Environ. Pollut. 341, 122882. (Ledesma et al., 2024)

Steigerwald, S., Saladin, Y., Alurralde, G., Abel, S., Sobek, A., Eriksson Wiklund, A.K., Gorokhova, E., 2024. Enhanced tolerance to narcosis in starved *Daphnia magna* neonates. Environ. Toxicol. Chem. accepted. (Steigerwald et al., 2024)

8.2 Reports

Gorokhova et al., 2023. Linking biological effects to contaminant levels in sediment, fish and benthic invertebrates along the Swedish coast, Report, Swedish Environmental Protection Agency, 86 pp.; https://urn.kb.se/resolve?urn=urn:nbn:se:naturvardsverket:diva-11066 (No. oai:DiVA.org:naturvardsverket-11066). (Gorokhova et al., 2023)

8.3 Preprints

Gorokhova, E., 2024. Integrating composite toxicity potential indices with biological indicators: Advancing contaminant status assessment under the MSFD. bioRxiv. (Gorokhova, 2024)

Gorokhova, E., Ledesma, M.G., Karlsson, A.M.L., Garbaras, A., Ketmaier, V., Sundelin, B., 2024. Understanding amphipod reproductive health drivers in the Baltic Sea: A Study Based on Structural Equation Modeling. Authorea. https://www.authorea.com/users/243058/articles/575500-understanding-amphipod-reproductive-health-drivers-in-the-baltic-sea-a-study-based-on-structural-equation-modeling. (Gorokhova et al., 2024)

8.4 Master theses

Yves Saladin. 2023. The Use of Protein Content as a New Endpoint in Acute Ecotoxicological Testing, Stockholm University. (Saladin, 2023)

Manuela Ospina Arboleda. 2023. Sorption kinetics of 4 PAHs in different batches of peat and use of passive sampling to measure target chemical activities. Stockholm University. (Arboleda, 2024)

8.5 Scientific conference presentations

Abel S., Akkanen J., Eriksson Wiklund A.-K., Gorokhova E., Sobek A. 2023. Assessing environmental risks of hydrophobic organic contaminants in sediments using chemical activity. Conference on Sustainable Management of Contaminated Sediments: ongoing research in Sweden (Swedish EPA). Platform presentation. Stockholm, Sweden.

Abel S., Akkanen J., Chaumet B., Eriksson Wiklund A.-K., Gorokhova E., Sobek A. 2023. Integrating Chemical Activity into Sediment-Water Bioassays with Benthic Invertebrates. SETAC North America 44th annual meeting. Poster presentation. Louisville, KY, USA.

Abel S., Ospina Arboleda M., Eriksson Wiklund A.-K., Gorokhova E., Sobek A. 2023. How to set up sediment-water bioassays for hydrophobic organic contaminants using their chemical activity as a dose metric. SETAC North America 44th annual meeting. Poster presentation. Louisville, KY, USA.

Abel S., Chaumet B., Steigerwald S., Eriksson Wiklund A.-K., Gorokhova E., Sobek A. 2023. The (sometimes overlooked) role of temperature in passive dosing experiments and chemical activity-based exposures. SETAC North America 44th annual meeting. Poster presentation. Louisville, KY, USA.

Abel S., Steigerwald S., Alurralde G., Eriksson Wiklund A.-K., Gorokhova E., Sobek A. 2022 Moving Towards a Chemical Activity-Based Risk Assessment of Sediments. SETAC North America 43rd annual meeting. Platform presentation. Pittsburgh, PA, USA.

Abel S., Akkanen J., Steigerwald S., Alurralde G., Eriksson Wiklund A.-K., Gorokhova E., Sobek A. 2022. Validation of a Novel Test System for Exposure Assessment with Benthic Invertebrates and Chemical Activity as a Dose Metric. SETAC North America 43rd annual meeting. Poster presentation. Pittsburgh, PA, USA.

Abel S, Eriksson Wiklund A.-K., Gorokhova E., Sobek A. 2022 Moving Towards a Chemical Activity-Based Risk Assessment of Sediments. SETAC Europe 32nd annual meeting. Poster presentation. Copenhagen, Denmark

SETAC Europe 33rd Annual Meeting (30 April–4 May 2023)–Ireland. Isotopic variation in response to the combined chemical exposure and temperature. Alurralde G, Steigerwald S, Garbaras A, Abel S, Eriksson-Wiklund A-K, Sobek A, Gorokhova E.

Alurralde G, Steigerwald S, Gineitytė M, Eriksson-Wiklund A-K, Sobek A, Abele S, Garabas A, Gorokhova E. International Symposium PRIMO 22 (26-29 May 2024)–France. Expanding the scope of stable isotope analysis to unravel chemical exposure effects.

S. Steigerwald, G. Alurralde, S. Abel, AK. E.-Wiklund, A. Sobek, E. Gorokhova. Incorporating nonchemical stressors into ReproIND indicator. SETAC North America; November 2022; POSTER.

S. Steigerwald, G. Alurralde, S. Abel, AK. E.-Wiklund, A. Sobek, E. Gorokhova Combined effects of chemical mixtures and temperature on *Daphnia magna*. SETAC Young Environmental Scientists; August 2023; POSTER

S. Steigerwald, G. Alurralde, S. Abel, AK. E.-Wiklund, A. Sobek, E. Gorokhova. Evaluating combined effects of temperature and chemical exposure in *Daphnia magna*. SETAC Europe; May 2024; POSTER

S. Steigerwald, Y. Saladin, G. Alurralde, S. Abel, AK. E.-Wiklund, A. Sobek, E. Gorokhova. Enhanced tolerance to narcosis in starved *Daphnia magna* neonates. SETAC Europe; May 2024; PLATFORM

8.6 Stakeholder presentations

Alurralde Gastón, Gorokhova Elena. Second Informal Consultation Session of the Expert Group on Hazardous Substances (IC EG HAZ 2-2023). Online meeting–February 7-8, 2023. Presentation of INSERT and discussion on chemical activity applications.

Alurralde Gastón, Gineitytė Monika, Garbaras Andrius, Gorokhova Elena. Baltic Sea Day 2021. Baltic Sea Center, Stockholm University–Sweden. Reproductive disorders and contaminant exposure drive isotopic variability in *Monoporeia affinis*.

8.7 Open access data

Integrated dataset on contaminant exposure, reproductive health, and environmental variables in the Baltic Sea (INSERT project): DOI: 10.5281/zenodo.17357254

REPORT 7208 | DECEMBER 2025 ISBN 978-91-620-7208-7 ISSN 0282-7298

Cumulative effects: Incorporating Nonchemical StrEssors into ReproIND indicator

Pollution in marine environments is a pressing global concern, with the Baltic Sea being particularly vulnerable due to its semi-enclosed nature and heavy human activity. The Marine Strategy Framework Directive (MSFD) aims to protect and preserve the marine environment, promoting sustainable use of marine resources. In this context, the project INSERT united experts in ecotoxicology, environmental chemistry, ecology, and biochemistry to enhance biological effect assessment under the MSFD. The project incorporated nonchemical stressors and utilized recent advancements in mixture toxicity assessment and pollutant monitoring in the Baltic Sea, addressing gaps in traditional risk assessments that overlook cumulative environmental effects. The study resulted in recommendations to enhance the practical value of contaminant screening surveys, monitoring and status assessment in the Baltic Sea. The study also demonstrated the value of using long-term monitoring data on the health status of sentinel species to integrate chemical and biological criteria for MSFD. The integrated approach supports ongoing revisions of monitoring and assessment strategies for biological effects. The project outcomes are communicated to HELCOM and national environmental agencies to ensure high impact, advancing environmental assessment and management of the Baltic Sea, both nationally and internationally.

The research project has been funded by the Swedish EPA's environmental research grant, which funds research in support of the Swedish EPA and the Swedish Marine and Water Authority's knowledge needs.

NATUR VÅRDS VERKET

SWEDISH ENVIRONMENTAL PROTECTION AGENCY

Swedish Agency for Marine and Water Management

Naturvårdsverket, SE-106 48 Stockholm. E-mail: registrator@naturvardsverket.se, www.swedishepa.se
Visiting address Stockholm: Virkesvägen 2. Visiting address Östersund: Forskarens väg 5, hus Ub.