Ecological Indicators 160 (2024) 111837

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind

Linking contaminant exposure to embryo aberrations in sediment-dwelling amphipods: a multi-basin field study in the Baltic Sea

N. Kolesova [a,*], S. Sildever [a], E. Strode [b], N. Berezina [c], B. Sundelin [d], I. Lips [a,e], I. Kuprijanov [a], F. Buschmann [a], E. Gorokhova [d,*]

[a] *Department of Marine Systems, Tallinn University of Technology, Estonia*
[b] *Latvian Institute of Aquatic Ecology, Latvia*
[c] *Zoological Institute, Russian Academy of Sciences, Russia*
[d] *Department of Environmental Science, Stockholm University, Sweden*
[e] *EuroGOOS AISBL, Belgium*

ARTICLE INFO

Keywords:
Monoporeia affinis
Reproductive disorders
Embryo aberrations
Baltic Sea
Contaminants

ABSTRACT

Embryo development of sediment-dwelling amphipod *Monoporeia affinis* is sensitive to contaminant exposure. Therefore, embryo aberrations in gravid females are used to detect the biological effects of contaminant exposure in the Baltic Sea benthic habitats. The indicator based on the aberration frequencies in wild populations (ReproIND) is currently used for environmental status assessment within the Marine Strategy Framework Directive, Descriptor 8.2. However, so far, it has mainly been applied in the Bothnian Sea (BoS) and the Western Gotland Basin (WGB), where it was found to respond to contaminant pressure and non-chemical environmental factors, such as temperature.

To expand the applicability of the indicator to other Baltic Sea basins, we used field data from the gulfs of Finland and Riga, BoS, and WGB to investigate the relationships between reproductive disorders and contaminants and environmental factors, thus evaluating the indicator suitability in these areas. Despite the natural variability of the environments and contaminant distribution across and within the basins, we found that high concentrations of contaminants, e.g. metals, PAHs, and PCBs, contribute significantly to the embryo aberrations in *M. affinis*. These findings support ReproIND applicability in the Baltic Sea and, perhaps, in other marine areas.

1. Introduction

The Baltic Sea is a semi-enclosed coastal sea with a limited water exchange, which makes it vulnerable to chemical pollution (Schneider et al., 2000, Jędruch et al., 2017, HELCOM, 2021). Despite massive efforts to reduce the input of hazardous substances, the contamination levels are still above acceptable levels (Löf et al., 2016a, HELCOM, 2018c, Kuprijanov et al., 2021). Intensive anthropogenic activities in the region are responsible for a broad spectrum of pollutants that reach the Baltic Sea, mainly through river runoff and atmospheric deposition (HELCOM, 2018a). Due to the slow water exchange and the long persistence of some contaminants, hazardous substances can be trapped in the sediments for long periods, leading to chronic exposure and adverse effects in biota (Strand & Asmund, 2003, Viglino et al., 2004, Cornelissen et al., 2008, Tansel et al., 2011). Effects caused by contaminants can be detected across different levels of biological organization, with measurable responses in animal biochemistry, physiology, reproduction, and behavior (Sundelin et al., 2000, Turja et al., 2014b, Kholodkevich et al., 2017, Pérez and Hoang, 2017, Podlesińska and Dąbrowska, 2019, Gorokhova et al., 2010, 2020, Berezina et al., 2022).

In the European Union, marine environment protection is coordinated by the Marine Strategy Framework Directive (MSFD), which aims to establish and maintain the good environmental status (GES) of marine ecosystems (European Commission, 2008). The Helsinki Commission (HELCOM) is responsible for MSFD implementation in the Baltic Sea region. It also serves as a platform for assessing and monitoring the status of the marine environment for establishing and maintaining GES. The evaluation of the marine environmental status involves eleven descriptors, including Descriptor 8 (D8), which focuses on the levels and impacts of the environmental pollution load (European Commission, 2008). The goal of the D8 assessment is to maintain contaminant

* Corresponding authors.
E-mail addresses: natalja.kolesova@taltech.ee (N. Kolesova), elena.gorokhova@aces.su.se (E. Gorokhova).

https://doi.org/10.1016/j.ecolind.2024.111837
Received 7 July 2023; Received in revised form 28 February 2024; Accepted 29 February 2024
Available online 13 March 2024

concentrations below levels that cause pollution effects. Biological effects caused by contaminants are assessed through Criteria D8C2, which considers population abundance and health, community composition, and habitat condition. Currently, the assessment of the Baltic Sea environmental status primarily relies on Criteria D8C1, which sets concentration thresholds for selected contaminants in water, sediment, and biota matrices, whereas only a few biological effect indicators are included in the assessments (HELCOM, 2010, 2018c, 2023).

The biological effect indicators allow quantifying the influence of contaminants on biota (Lam & Gray, 2003, Martín-Díaz et al., 2004). Therefore, establishing relationships between the pressures (i.e., concentrations of environmentally relevant contaminants and their mixtures) and effect indicators is a pre-requisite for the design, selection, and quality assessment under Criteria D8C2 and subsequent integration of the chemical and biological indicators for the overall D8 assessment (Queirós et al., 2016, Lyons et al., 2017). Moreover, the confounding factors need to be recognized and, if possible, accounted for in the indicator-based assessment to ensure that the regional threshold values and assessment criteria are ecologically meaningful (Gorokhova et al., 2010, Strode et al., 2023). Considering that environmental variables, e.g., the sediment grain size, temperature, oxygenation, and total organic carbon (TOC) content, are well-recognized drivers of the distribution and bioavailability of organic contaminants and metals in sediments (Tanner et al., 1993, Li et al., 2022), these factors should be evaluated for their potential contribution for the indicator variability in the target area.

Different species of amphipods are widely used in sediment bioassay tests (Podlesińska and Dąbrowska, 2019). The indicator ReproIND, *Reproductive disorders: malformed embryos of amphipods*, is one of the few biological effect indicators used for assessing the contaminant impacts *in situ* (HELCOM, 2023; HELCOM, 2018b). The indicator was developed for amphipod species *Monoporeia affinis* and *Pontoporeia femorata* and since 1994 used in the Swedish Marine Monitoring Programme (SNMMP) for the assessment of the contaminant effects in the Swedish coastal waters, the Bothnian Sea (BoS), the Quarken, the Northern Baltic Proper, and the Western Gotland Basin (WGB) (HELCOM, 2018b). ReproIND can be applied to other amphipod species that share similar reproduction biology, even if their habitat preferences differ (Sundelin et al., 2008b; HELCOM, 2018b).

The embryo aberration frequency is informative as it reflects the sensitivity of amphipod embryos to contaminant exposure during their development within the brood pouch, which lasts for weeks to months. Furthermore, it provides insights into female exposure over an extended period of up to two years, encompassing growth, maturation, and oogenesis (Löf et al., 2016a) . Several laboratory studies have reported the increased frequency of aberrant embryos in amphipods (Eriksson et al., 1996, Sundelin et al., 2008b, Berezina et al., 2019, 2022). However, demonstrating the linkage between pollution and biological effects indicators in the field is more challenging due to the simultaneous influence of several factors (McCarty & Mackay, 1993, Martín-Díaz et al., 2004). In the Baltic Sea, a higher frequency of aberrant embryos was found in contaminated than reference sites and decreased with a distance from the point sources (Sundelin & Eriksson, 1998, Reutgard et al., 2014). Also, in *M. affinis* population, the frequency of females with aberrant embryos showed a positive correlation with sediment concentrations of cadmium (Cd), polychlorinated biphenyls (PCBs), and chlorinated hydrocarbons (PAHs) (Löf et al., 2016a). However, the field evidence linking amphipod embryo aberrations to contaminants has been predominantly limited to regional studies, with a restricted range of environmental factors and contaminant distribution considered (Reutgard et al., 2014, Löf et al., 2016a, Berezina et al., 2017, Strode et al., 2017). To effectively apply ReproIND in the environmental assessment of different Baltic subbasins, obtaining a broader geographic coverage of embryo aberration occurrence across diverse environmental settings is crucial.

Here, we established relationships between embryo aberrations in *M. affinis* and hazardous substances across several subbasins in the Baltic Sea, where this species is abundant and used as a sentinel species to monitor the biological effects of contaminants (e.g., Löf et al., 2016a) and macrobenthic diversity (e.g., Raymond et al., 2021). Moreover, we also addressed the effects of the ecologically relevant confounding factors, i.e., bottom depth, temperature, oxygen, sediment grain size, and TOC, on the contaminants and the reproductive responses. Our findings validate the ReproIND indicator based on field data and expand its operationalization in the Baltic Sea, contributing to the comprehensive biological effect monitoring tools and supporting the Baltic Sea Action Plan (BSAP) (HELCOM, 2021).

2. Materials and methods

2.1. Sampling

Gravid *M. affinis* females were collected in four subbasins in the Baltic Sea (Fig. 1). Samples from GoF (14 sites) and GoR (3 sites) were a part of pilot projects between 2016 and 2021, whereas BoS (5 sites) and WGB (3 sites) data were collected within the Effect Screening Study (ESS; 2017–2018) coordinated by the Swedish Environmental Protection Agency. The sampling sites (25 in total, visited from December to March) represented both contaminated and relatively uncontaminated (i.e., without known point sources) sediments (Fig. 1).

Sediments and amphipods were collected with a bottom sled collecting the uppermost sediment layer (Blomqvist and Lundgren, 1996) in BoS and WGB, and van Veen sediment grab (sample area of 0.1 m^2) in GoF and GoR. The surface sediment layer (up to 5 cm) from the van Veen grab and mixed sediment from the sled samples were used for contaminant analysis and granulometry. After that, the sediment was sieved through 0.5–1 mm mesh size; gravid females were gently separated and kept alive in the seawater at the same temperature as the ambient environment ($< +4$ °C) until laboratory analyses. Salinity, temperature, and dissolved oxygen concentrations were measured from the near-bottom water layer (Table S1). For BoS and WGB sites, the temperatures (°C) within two weeks of the sampling occasion were extracted from the Swedish (Swedish National Oceanographic Data Centre at the Swedish Hydrological and Meteorological Institute, data available at: https://sharkweb.smhi.se).

2.2. Chemical analysis

Collected sediments were analyzed for heavy metals (As, Pb, Cd, Cr, Cu, Ni, Hg, Zn; mg kg^{-1}DW), polycyclic aromatic hydrocarbons (PAHs; ng g^{-1}DW), polychlorinated biphenyls (PCBs; ng g^{-1}DW), and butyltins (BTs; ng g^{-1}DW) (Table S2). In addition, the sediment grain size (%) and total organic carbon (TOC, %) were also analyzed (Table S1). Analysis of contaminants from sediment samples was outsourced and conducted in chemical laboratories in Estonia, Russia, Latvia, Germany, and Sweden (Table S3).

2.3. Embryo analysis

The embryo analysis was conducted using a stereomicroscope and following the methodology established by Sundelin et al. (2008b). The number of analyzed gravid females was recorded for each sampling site (Table S4). In each female, fecundity (eggs $female^{-1}$), embryo development stage (from 1 to 9; Stage) (Sundelin et al., 2008b), and embryo aberrations following the classification summarized by Löf et al. (2016a) were recorded. Different aberration types (Malf: malformed embryos, Membr: embryos with membrane damage, AD: embryos with arrested development, DE: dead eggs, and DB: dead brood) were used for calculating the ReproIND (HELCOM, 2018b). The indicator consists of two components calculated for each station: the proportion of aberrant embryos (%AbEmb; a sum of Malf, Membr, AD, and DE divided by the total number of the examined embryos) and the proportion of females

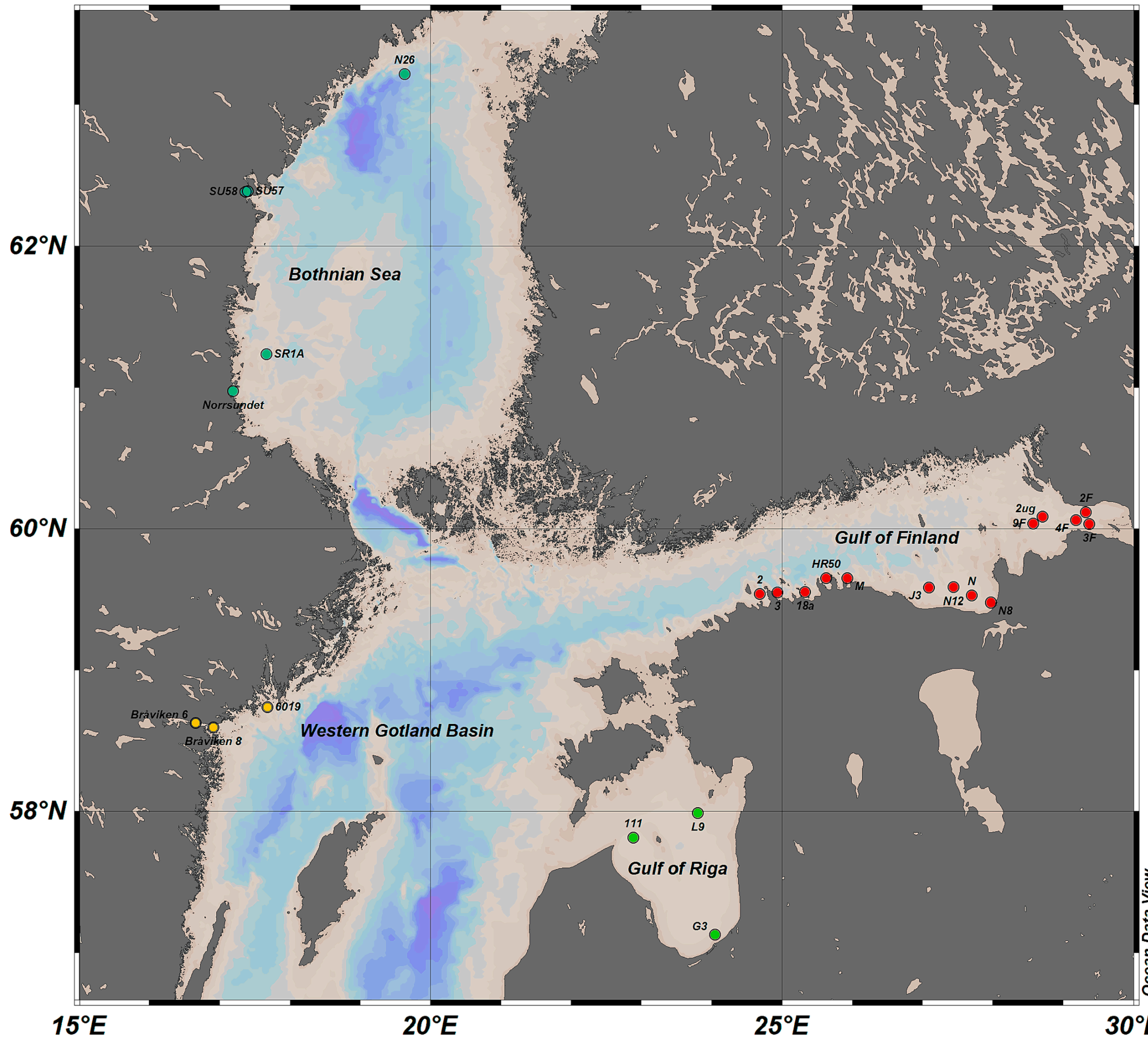

Fig. 1. Sampling locations for *M. affinis* and sediment samples in the Gulf of Finland (GoF): filled red circles, the Gulf of Riga (GoR): green, Bothnian Sea (BoS): blue, and the Western Gotland Basin (WGB): yellow.

with more than one aberrant embryo (%Fem > 1; the sum of gravid females carrying more than one aberrant embryo divided by the total number of the examined females).

2.4. Data analysis

2.4.1. Chemical data

For the total assessment of the metal contamination, we used the Tomlinson Pollution Load Index (PLI; Tomlinson et al., 1980) calculated as a geometric average of contamination factors for the nine metals (As, Cd, Co, Cr, Cu, Hg, Ni, Pb and Zn) measured in all data sets and used Swedish regional reference values for these metals in sediment (Havs- och vattenmyndigheten, 2018). The total concentrations of the 15 PAHs calculated as the sum of individual compounds (PAHs), the sum of four low-molecular-weight PAHs (LmPAHs: Phe, Ant, Flu, Pyr), and the sum of eight high-molecular-weight carcinogenic PAHs (HmPAHs: BaA, Chr, BbF, BkF, BaP, Inp, BgP, DBA) were used as descriptors of PAH contamination. For PCBs and BTs concentrations, the sum of congeners within each group was calculated. The concentrations below the limit of quantification (LOQ) were included in the univariate analysis as if they were true observations with a zero value. Due to the differences in LOQ among the analytical laboratories, we avoided replacing the non-detects with a fractional value of LOQ. The variables with > 60 % of the non-detects were omitted from the regression analysis.

2.4.2. Statistical evaluation

All *p* values presented are two-tailed, with $p < 0.05$ considered significant for all statistical tests. When relevant, the numerical variables were assessed for normality using the Shapiro-Wilk test. The numerical variables were presented as the mean ± standard deviation if the distribution was normal or as a median and interquartile range if the distribution was skewed. Reproductive aberrations were presented as frequencies (proportions), and the differences across the basins were evaluated using Fisher's exact test. Differences in sediment grain size (clay, sand, silt), TOC concentrations, and other environmental parameters (oxygen, salinity, and temperature) across the subbasins were evaluated using Kruskal-Wallis tests in R (R Core Team, 2024); here, we used non-parametric tests due to the lack of homogeneity in the variances and different number of sites across the subbasins.

2.4.3. Multivariate analyses

Multivariate ordination techniques were used for grouping the sampling sites according to their contaminant profiles and identification of the associated variables. The missing data for contaminants was replaced using EM (maximum expectation likelihood) algorithm, which assumes a multi-normal distribution model for the data (Dempster et al., 1977) with 1000 permutations as implemented in Primer 7 v.7.0.21 software (Anderson, 2017). The environmental and contaminant data were log-transformed and normalized to avoid misclassification due to the differences in data dimensionality (Anderson et al., 2008). As a pre-treatment transformation, the environmental and contaminant variables were log-transformed of log $(x + n)$, where n is a small number compared to the measured value. The transformed data were subjected to zero mean and unit variance normalization (z-scores), which were used in correlation and regression analyses and for the resemblance matrices in the multivariate analysis.

For visual data exploration, we used FreeViz, a free software including an optimization method that finds linear projection and associated scatterplots based on gradient descent modeling (Demšar et al., 2007). FreeViz separates instances of a different class evaluated through mean scores and graphical optimization for compaction and separation between instances of the same class. Also, a hierarchical cluster analysis (HCA) was performed on the contaminant data to identify groups of sites that had similar composition and concentrations of the contaminants. A classification scheme using the Euclidean distance for similarity measures between sites was performed. Ward's method was used for the establishment of the links between the sites to improve the distinctive power of the classification (Güler et al., 2002). Once the clusters were identified, we conducted a comparison of reproductive aberration frequencies across these clusters using an unpaired t-test following Box-Cox transformation of the frequency data to stabilize variances. This statistical analysis aimed to assess whether a general response to the contaminants present in the dataset exists, independent of other environmental variables.

Further, Distance-based linear modeling (DistLM) was used to assess the relative importance of (1) environmental factors (sediment texture, temperature, oxygen, and TOC) in shaping the distribution of contaminants (concentrations and indices) and (2) environmental factors and contaminants in explaining the observed reproductive responses in amphipods. DistLM is a multivariate linear model that uses explanatory information provided as a distance matrix to generate the most parsimonious combination of predictors to explain variation in the response variables while accounting for the potential overlap of the predictors. In DistLM, pseudo-F statistic for testing the general null hypothesis of no relationship is used and p values for individual predictors are obtained through permutations (in this study 9999). Thus, the approach is robust to non-normal data, and errors do not need to be normally distributed (McArdle & Anderson, 2001).

To avoid collinearity between predictors, variables that were strongly correlated with each other ($r > 0.8$; Table S5) were omitted, and no interactions were considered due to the relatively low number of observations. The subsets of variables were established using forward selection based on the multivariate analogue to the small-sample-size corrected version of the Akaike Information Criterion (AICc). Relationships between environmental parameters and reproductive attributes were initially examined by analyzing each predictor separately (marginal tests). Then, partial regressions were used to characterize the relationships accounting for the effect of the remaining variables by sequential tests with step-wise selection procedures and AICc as the selection criterion. The models were visualized using a distance-based Redundancy Analysis plot (dbRDA).

3. Results

3.1. Environmental data

The bottom depth ranged from 12 to 134 m in BoS, 19 to 42 m in WGB, 16 to 52 m in GoF, and 24 to 38 m in GoR. TOC in sediments ranged from 0.3 to 5.9 % in GoF, 0.5–1.1 % in GoR, 0.4–3.8 % in BoS, and 0.7–2.3 % in WGB. The other environmental parameters, including salinity (1.4–7.2), near-bottom temperature (2.1–5.7℃), oxygen concentration (6.5–12.5 mg l^{-1}), and sediment texture (percentage of clay, silt, and sand) were highly variable across the subbasins (Table S1). The salinity was the main parameter associated with the variability between the sites and the only parameter that displayed a statistically significant difference ($p < 0.05$) between the subbasins (Fig. S1). Compared to other subbasins, lower salinity, and higher temperature were recorded for GoF, with the opposite pattern for WGB (Fig. S1; Table S1). Silt prevailed in the GoF and BoS sediments, 59 % and 56 %, respectively, whereas sand dominated in the GoR and the WGB sediments (Fig. 2; Table S1). WGB and BoS sites generally displayed higher salinity and more sandy sediments. In addition, GoF and GoR sites were characterized by better oxygenation (Fig. 2), although GoF sites were more variable in terms of the sediment grain size and oxygenation, whereas GoR sites had more uniform oxygenation and sandy sediments (Fig. 2). This variability in GoF was mostly related to the sites in the Neva Estuary, the eastern part of GoF, that had lower salinity and higher proportion of silt in the sediment (Table S1).

3.2. Contaminant data

The highest variability in the contaminant levels was observed for BoS, mostly due to the highly contaminated sites SU57 and SU58 (Sundsvall), and GoF (near-harbour areas in the Neva Estuary) subbasins. The BoS sediments had the highest concentrations of metals (mean 283 mg kg^{-1} DW with a PLI value of 13.8), PAHs (mean 30.3 μg g^{-1} DW), and PCBs (mean 23.7 ng g^{-1}DW; Table S2). The GoF sediments had the highest BT loading (mean 61.7 ng g^{-1}DW), but also high variability of the PCBs, Pb, and PLI values (Fig. S2).

The total metal concentration in sediments was 37 – 364 mg kg^{-1} DW, with the lowest and the highest values at the G3 site (GoR) and Bråviken 8 (WGB), respectively. Moreover, high As levels were recorded at SU57 and SU58 sites (BoS) and high Cd in the Neva Estuary at 3F, 4F, and 9F sites (GoF; Fig. 3). The total PAH concentrations ranged from non-detectable levels (G3 site, GoR) to 81 μg g^{-1} DW (SU58 site, BoS; Table S2). The total PCB concentations ranged from non-detectable levels observed at several sites in different subbasins to 81 ng g^{-1} DW (SU58 site, BoS). Finally, the total BT concentrations were frequently non-detectable, but relatively high in GoF, sites 2F, 3F, 4F, 9F and 2ug (up to 335 ng g^{-1} DW; Table S2).

3.3. Associations between environmental parameters and contaminants

The environmental parameters explained 90.1 % of the fitted and 52.5 % of the total variation in the contaminant load across the stations (DistLM; Table 1, Table S5, Fig. 4). Based on the environmental variability and contaminants, the sampling sites formed two groups, where the first consisted of the stations located in the Neva Estuary in GoF, which are characterized as stations with high TOC, BTs and low PCBs concentrations, and the second comprised the remaining sites, with higher PCB concentrations than the first group. Variability in some GoF stations was influenced by salinity and/or temperature, while higher clay content in sediments played a significant role in the differences in chemical load observed in the majority of other stations (Fig. 4). Overall, higher levels of contaminants were found at lower temperatures, higher salinities and in fine-grain sediments.

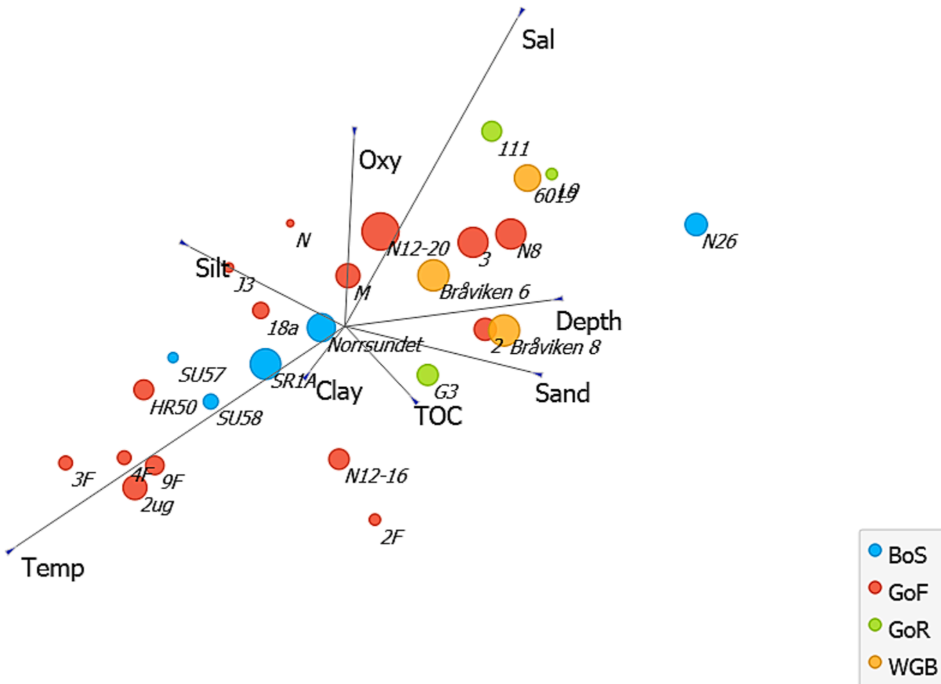

Fig. 2. Association between the environmental variables and stations across the subbasins. The vectors show the variable variability (length) and orientation; the bubble size corresponds to the number of amphipod females collected for the analysis from each station; the subbasins are color-coded according to the legend. The station names are indicated on the circles, and arrows carry the environmental variable names; please see Table S1 for the complete list of the stations and environmental variables.

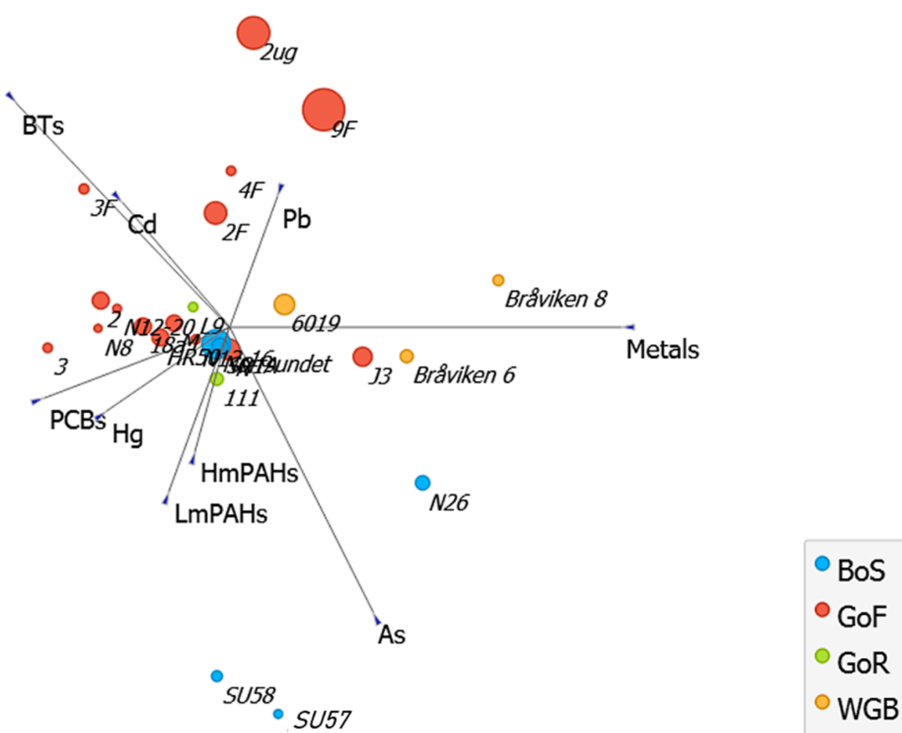

Fig. 3. Associations between the contaminants and stations across the subbasins. The vectors show the chemical contaminant variability (length) and orientation; see Table S2 for the complete list. The circle size corresponds to the sediment TOC value at the station, with the color coding indicating the subbasin. The circles are labeled with station names as in Fig. 2.

3.4. Biological parameters

Due to natural phenological variability across the subbasins and differences in the time of mating (December – January), the embryos used for analysis were in different developmental stages, with the average stage ranging from ~3 to 7, with the least developed embryos found in GoF (Table S4; Fig. 5). Fecundity between stations also varied (13.6–42.2 eggs female^{-1}), and the ranges for the proportion of the

Table 1
DistLM output for marginal and sequential tests linking environmental factors to the main contaminant groups at the sampling stations. The best model was selected based on Akaike information criteria (AICc) for all sites. Pseudo-*F* value, a multivariate analogue of F-value for linear regression; Adj. R^2, adjusted proportion of explained variation attributable to variables in a model; P, probability; R^2, the proportion of explained variation attributable to each variable; R^2 (cum), the cumulative proportion of variation; rs.df, residual degrees of freedom.

Marginal tests				
Variables	SS (trace)	Pseudo-*F*	P	R^2
Clay/silt	21.79	6.686	0.004	0.218
Temperature	9.936	2.648	0.06	0.099
Salinity	18.71	5.524	0.005	0.187
TOC	4.617	1.646	0.17	0.064
Oxygen	8.866	2.335	0.08	0.089
Depth	2.27	0.557	0.64	0.023
Sand	16.919	4.888	0.009	0.169
Sequential tests				
Variables	P	R^2	R^2 cum	rs.df
Clay/silt	0.001	0.18	0.2179	24
Temperature	0.009	0.32	0.37713	23
Salinity	0.02	0.39	0.46281	22
TOC	0.03	0.45	0.54072	21
Oxygen	0.12	0.48	0.58348	20
Overall best solution				
Adj. R^2	R^2	RSS	No. Vars	Selections
0.479	0.583	41.652	5	2–6

aberrant embryos (%AbEmb) and females carrying more than 1 aberrant embryo (%Fem > 1) were 2–22 % and 1–80 %, respectively. The lowest mean %AbEmb and %Fem > 1 values were recorded in GoR (6 % and 33 %, respectively) and the highest in BoS (12 and 56 %; Table S4).

Ordination of the reproductive variables (Fecundity, Stage, %AbEmb and %Fem > 1) revealed substantial variability between the subbasins, with a significant correlation between %AbEmb and %Fem > 1 (Spearman *rho*: 0.6, $p < 0.004$; Fig. 5). The highest variability for the reproductive variables was detected in GoF and the lowest in WGB (Fig. S3). Most of the GoF sites were characterized by relatively high levels of both %Fem > 1 and %AbEmb, albeit a lower proportion of aberrant embryos was recorded in the Neva Estuary.

3.5. *Linking reproductive variables to environmental factors and contaminants*

The classification of the sampling sites by cluster analysis revealed four clusters differing by the loads and compositions of contaminants (Fig. 6). Cluster 1 consists of only two highly contaminated sites SU57 and SU58 (Sundsvall, BoS), with very high levels of metals (As, Hg, Cu, and Ni), PAHs and PCBs. Cluster 2 consists of five sites located in GoF and having relatively high levels of BTs and some metals (Cd and Pb), whereas the other two clusters combine sites located in multiple subbasins and characterized by diverse mixtures of PAHs, PCBs, and metals. Cluster 3 includes sites with low levels of metals and intermediate levels of PAHs, and Cluster 4 includes sites with consistently high levels of PCBs, and intermediate-to-high levels of some PAHs and metals (Cr and Ni). Thus, none of the clusters represented non-polluted conditions. When %Fem > 1 was compared across the four clusters, significantly higher values were found for Cluster 4 compared to Clusters 3 (Unpaired t-Test: t-value = −2.27, df = 17, $p < 0.037$; Fig. 6) and to Cluster 2 (t-value = −2.52, df = 14; $p < 0.025$). For %AbEmb, the differences between the clusters were not significant, albeit higher values were observed for C4 compared to C3 and C2 (0.11 vs 0.082 and 0.077, respectively).

When both chemical and non-chemical predictors were considered, reproductive aberrations were significantly associated with environmental factors (salinity, temperature, TOC, and proportion of clay), PAH concentrations (naphthalene; NAP and acenaphthene; DBAHA), and PLI as a proxy for metals (DistLM; Table 2, Fig. 7). The individual

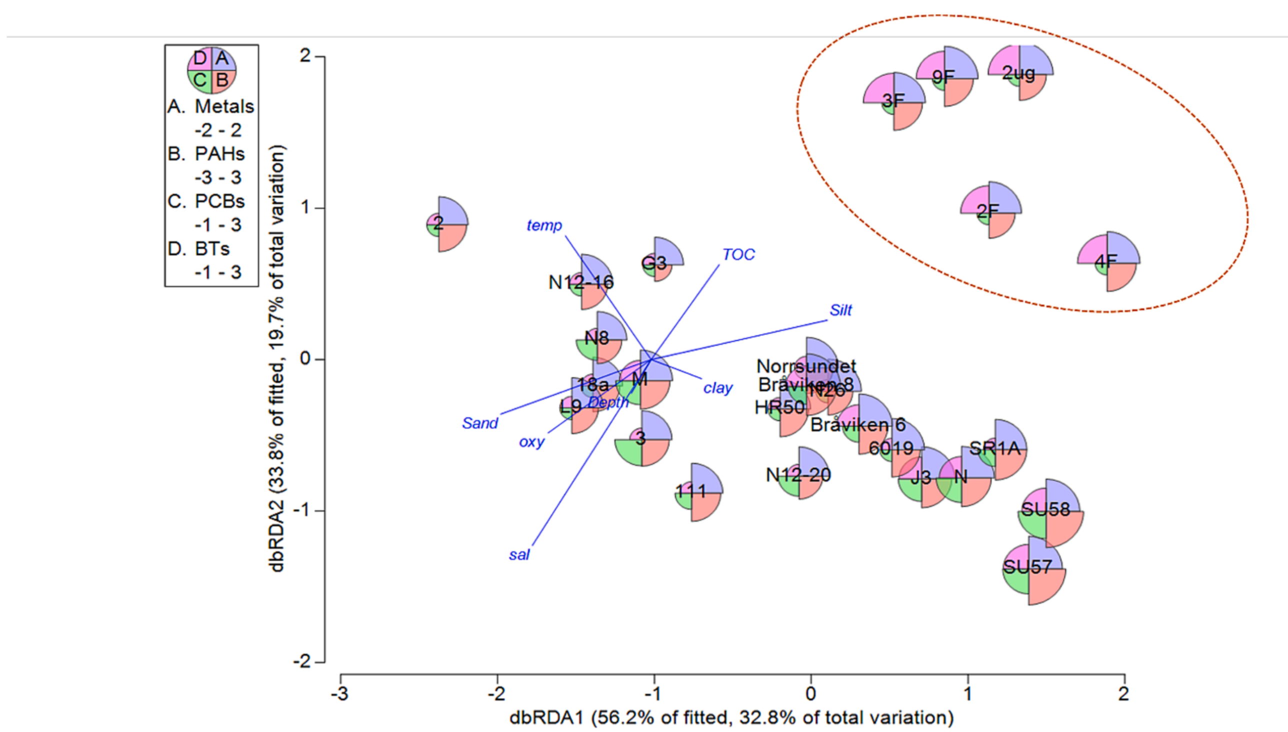

Fig. 4. dbRDA biplot for DistLM relating chemical load (by contaminant group, metals, PAHs, PCBs, and BTs) in sediment as multivariate response variables to non-chemical predictors: depth, temperature, oxygen, TOC, and sediment composition (clay, silt, and sand). The orange ellipse indicates stations located in the Neva Estuary (GoF) and characterized by high TOC, BTs and low PCBs concentrations.

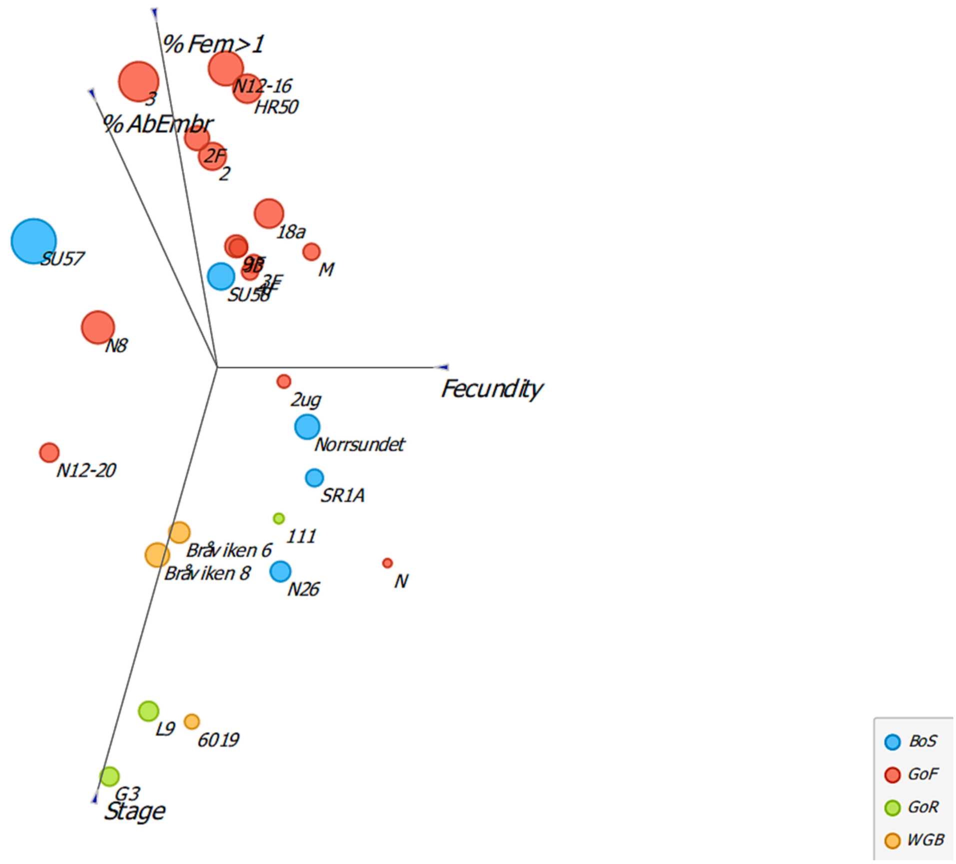

Fig. 5. Associations between the reproductive variables (fecundity, %AbEmb, %Fem > 1, and stage) and stations across the subbasins. The vectors show the variable variability (length) and orientation; the bubble size corresponds to the number of individuals analyzed from each station, and the color code indicates the subbasin. Red color represents GoF, blue: BoS, green: GoR and yellow: WGB.

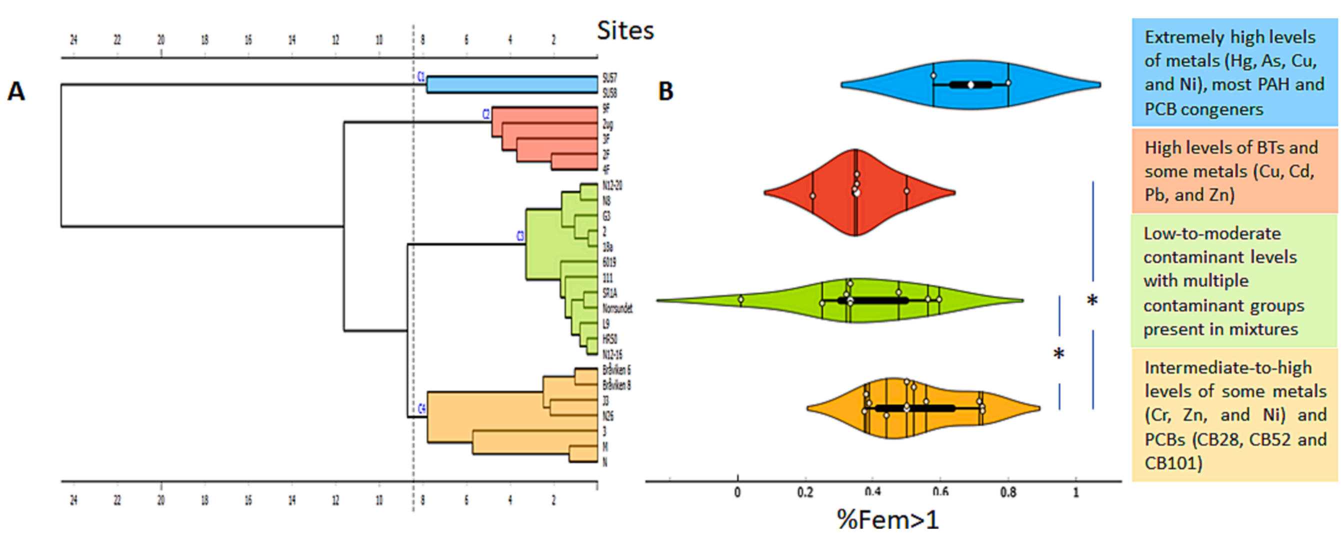

Fig. 6. Variability of the reproductive aberrations across sites grouped according to their contaminant load. **(A)** Hierarchical clustering of the sampling sites according to the contaminant load (Ward linkage), yielding four clusters (C1 to C4), and **(B)** Violin plots for the proportion of the females carrying more than 1 aberrant embryo (%Fem > 1). 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 was not included in the statistical comparison because it was comprised of only 2 stations.

contributions of the environmental variables ranged from 14 % (silt) to 34 % (salinity) of the total variance explained, whereas contaminants contributed from 13 % (NAP) to 18 % (PLI) (Table 2). The first two axes of the dbRDA plot of reproductive aberrations explained 80 % of the total and 99 % of the fitted variation, indicating that most of the salient patterns in the fitted model were captured (Fig. 7). Based on the model, the aberration frequencies increased with the contaminant concentrations, temperature, and salinity and decreased in organic-rich sediments

Table 2
DistLM output for marginal and sequential tests showing relationships between environmental factors, chemical variables, and reproductive aberrations in *Monoporeia affinis*.

Variable			Variable				
Marginal tests	P	R^2	Sequential tests	P	R^2	R^2 cum	rs.df
Physical properties							
(−) depth	0.945	<0.01					
(+) sal	**0.001**	**0.34**	sal	0.001	0.33957	0.33957	23
(−) oxy	0.188	0.07					
(+) temp	**0.001**	**0.33**	temp	0.001	0.17665	0.51622	22
(−) TOC	**0.004**	**0.20**	TOC	0.007	0.11002	0.62624	21
(−) clay	0.191	0.07	clay	0.036	0.045144	0.80329	18
(−) silt	**0.029**	**0.14**					
Contaminants							
(+) PLI	**0.042**	**0.18**					
(+) As	0.615	0.02					
(+) Cu	0.930	<0.01					
(+) Hg	0.180	0.07	Hg	0.045	0.0599	0.68614	20
(−) Ni	0.701	0.02					
(+) PAHs	0.218	0.07					
(+) NAP	**0.028**	**0.13**					
(+) DBAHA	**0.031**	**0.14**					
(+) FLE	0.124	0.09					
(+) PCBs	0.19	0.07					
(+) PCB-28	0.458	0.03					
(+) PCB-118	0.241	0.06					
(+) PCB-180	0.168	0.08	PCB-180	0.019	0.072004	0.75815	19

P, probability; R^2, the proportion of explained variation attributable to each variable; R^2 (cum), the cumulative proportion of variation; rs.df, residual degrees of freedom; (+/−), positive/negative relationship with %AbEmb and %Fem > 1.

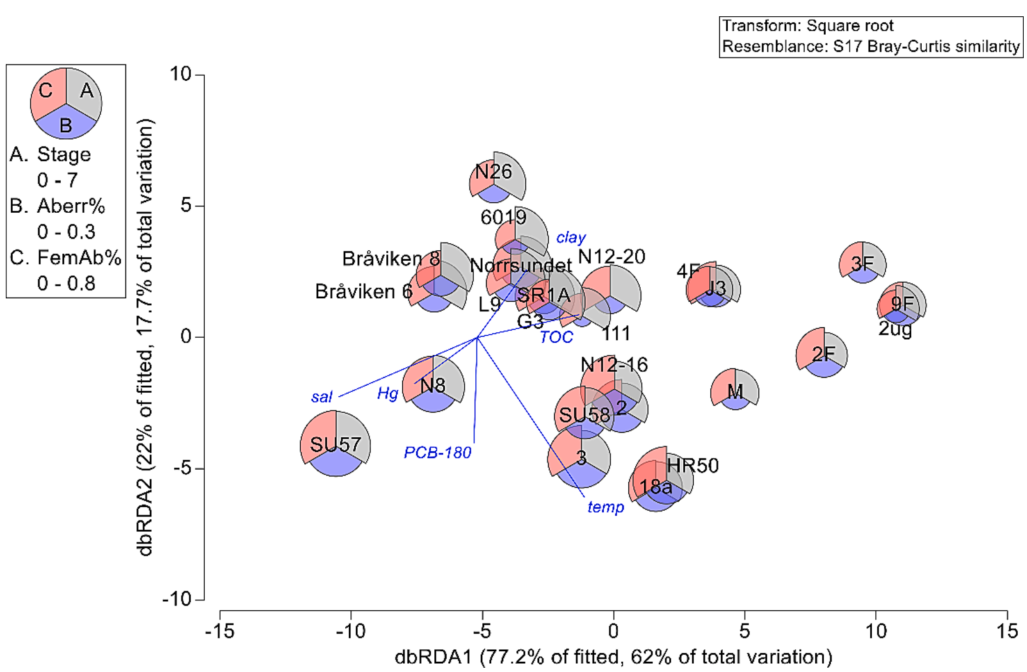

Fig. 7. Constrained ordination (dbRDA biplot) of the fitted values of the most parsimonious DistLM for reproductive aberration frequencies in amphipods (%AbEmb and %Fem > 1) as a multivariate response to the predictors identified by the model: contaminants (PCB-180 and Hg) and non-chemical environmental factors (temperature, TOC, salinity, and clay). In addition to the aberration frequencies, a developmental stage of the embryos is also shown.

with a high proportion of fine (clay and silt) particles.

There was a significant increase of aberrations with increased contaminant concentrations, i.e., metals (PLI, Hg), PAHs (NAP and DBAHA), and PCBs (PCB180). Given the high levels of cross-correlations between different metals and PLI values as well as across congeners of PAHs and PCBs, those were excluded from the list of candidate predictors in DistLM to avoid multicollinearity. Thus, it is likely that the detected effects are not specific to these compounds but reflected toxicity exerted by other congeners in the contaminant mixture. Therefore, the percentage of variance explained by a single component of a mixture of covarying compounds does not necessarily reflect the contribution of the overall chemical exposure to the observed effects.

4. Discussion

Using a geographically broad dataset on embryo aberrations in common amphipod *Monoporeia affinis*, environmental factors and contaminant levels in sediment, we found that aberration frequency is driven by both environmental factors (salinity, temperature, TOC and silt proportion) and contaminants (metals [PLI and Hg], PAHs [NAP and DBAHA] and PCBs [PCB180]). These findings confirm earlier reports that embryo aberrations in this species respond to chemical pollution (Sundelin et al., 2008a, Löf et al., 2016a; Löf et al., 2016b), and that this response is detectable *in situ*. Both components of the ReproIND indicator (i.e., percentage of aberrant embryos in a population and percentage of females carrying more than 1 aberrant embryo) increased at higher contaminant load (Fig. 7). Therefore, the proposed indicator demonstrates its applicability in all the investigated Baltic Sea subbasins for assessing contaminant impacts, even in the face of high environmental variability and substantial differences in contaminant load across the study areas.

4.1. Environmental variability and contaminants

The differences between the amphipod habitats in the western (BoS and WGB) and the eastern (GoR and GoF) subbasins were substantial, both in relation to the environmental parameters (Fig. S1) and the contaminant load (Fig. S2), with no noticeable differences in the TOC values. The eastern areas were characterized by significantly higher temperatures, better oxygenation, and mostly fine-grain sediments, albeit the grain size distribution varied across the GoF sites, with the finest sediment in the Neva Estuary, coinciding with the lowest salinity. Conversely, higher salinities and often more sandy sediments were common for the western sites.

The elevated contaminant levels were associated with high salinity, low temperature, and fine-grained sediments, consistent with other reports (Höglund and Jonsson, 2008, Löf et al., 2016a; Löf et al., 2016b, Erm et al., 2021). These observations were primarily driven by the sites with high contaminant loads in WGB and BoS and relatively low levels of all contaminants in GoR. Among the contaminant groups, trace metals were the most widespread contaminants in all subbasins, with particularly high concentrations of As in the Bråviken Bay (WGB) and Hg in the Bothnian Sea, whereas extremely high PAH and PCB levels were recorded in Sundsvall (SU57 and SU58; BoS). The latter is related to the historical contamination by pulp and paper mills leading to elevated levels of metals and organic pollutants in the sediments, along with high fiber concentrations (Höglund and Jonsson, 2008; Apler et al., 2014, 2019). Notably, the lowest contaminant levels were consistently found in the GoR sediments, with BTs being common in the Neva Estuary, particularly, in the proximity to shipping lanes (e.g., sites 3F, 4F, 9F, and 2ug; GoF).

4.2. Linkage between the embryo aberrations and contaminants

The environmental factors accounting for the differences between the subbasins explained most of the variability in the embryo aberrations, which was expected, whereas individual contaminants explained up to 18 % of the captured variability. Both parameters describing reproductive aberration frequency in a population (%AbEmb and %Fem $>$ 1) were positively correlated with salinity and temperature and negatively with TOC and silt. An increase in embryo aberrations at higher temperatures agrees with the current knowledge on *M. affinis* ecology (Wiklund & Sundelin, 2001, Eriksson Wiklund and Sundelin, 2004). Furthermore, similar negative effects of temperature on embryo development have been observed in other amphipod species in the region (Berezina et al., 2017).

In the Neva Estuary, aberration frequency was relatively low, despite very high BT concentrations (on a global scale; Meador et al., 1997), coinciding with high TOC levels and clay/silt in the sediments. High organic carbon content and fine-grained sediments have been reported to convey reduced contaminant bioavailability (Kreitinger et al., 2007; Baran et al., 2019). As hydrophobic BTs are sorbed to organic matter, their bioavailability is reduced (Rüdel, 2003, Cornelissen et al., 2005), leading to limited uptake and bioaccumulation as reported for tributyltin in amphipod *Rhepoxynius abronius* inhabiting high-TOC sediments (Meador et al., 1993,1997). Further, interspecific variability in bioaccumulation and capacity to metabolize contaminants, such as TBT, as well as differences in the depth and rate of feeding have been reported in amphipods (Byrén et al., 2002, Ohji et al., 2002). A further study investigating intraspecies variability in *M. affinis*, might offer an explanation for the low embryo aberration rates from the Neva Estuary.

Among the analysed contaminants, the metals (PLI, Hg), PAHs (NAP and DBAHA), and PCBs (PCB180 but also PCB28 and PCB118; Table 2) explained the most variability in the reproductive aberrations (% AbEmb, %Fem $>$ 1). These findings corroborate those by Löf et al., (2016a) from two localities in the Gulf of Bothnia, where the %AbEmb variability was best explained by PCBs (with PCB180 being the primary driver), PAHs (Phenanthrene, 1-Methylphenanthrene, benzo[ghi]perylene) and Cd. A higher aberration rate near known pollution sources has been reported as a general tendency in *M. affinis* and other amphipod species (Sundelin & Eriksson, 1998, Bach et al., 2010, Reutgard et al., 2014, Tairova and Strand, 2022). However, as contaminants occur in mixtures and under heterogeneous conditions, the univariate relationships between specific pollutants and reproductive responses should not be expected, which further motivates the development of biological indicators of contaminant exposure and effects applicable across ecosystems.

The significant contribution of contaminants to the embryo aberration variability reported here supports the applicability of the ReproIND indicator in all Baltic Sea areas where *M. affinis* and several other amphipod species are common. Previous studies exploring environmental and contaminant drivers on the reproductive aberrations in *M. affinis* were limited to the Gulf of Bothnia (Reutgard et al., 2014, Löf et al., 2016a; Löf et al., 2016b). Despite extensive studies on contaminants and their biological effects in the Baltic Sea region (Kankaanpää et al., 2022), the use of biological effect indicators in monitoring programs is currently limited. One reason for this is a lack of standardization efforts and evaluations based on data from different areas and populations existing under contrasting conditions, especially concerning the temperature and salinity gradients known to shape the physiology and phenology of the Baltic biota (Eriksson Wiklund and Sundelin, 2004, Nohrén et al., 2009, Larsson et al., 2017). Therefore, our findings provide a strong background for the future development of the ReproIND indicator, particularly for setting regional thresholds to account for the confounding reproductive responses to the temperature, salinity, TOC, and sediment grain size distribution in different subbasins and sampling locations.

Moreover, to expand the biological effect assessment and include non-reproductive responses, ReproIND can be combined with other biomarkers, e.g., antioxidant defense, geno- and cytotoxicity (Turja et al., 2020), metabolic activity, oxidative balance, neurotoxicity (Löf et al., 2016a; Löf et al., 2016b), and DNA adductome (Gorokhova et al., 2020). An integrated approach employing a battery of biological effects would provide a better understanding of the broad impact of multiple mixed effects of contaminants (Lehtonen et al., 2014,2016, Turja et al., 2014a, Berezina et al., 2019) and support monitoring and management of environmental contaminants at the national, regional, and local levels.

5. Conclusion

Using data on reproductive aberrations in *M. affinis*, chemical contaminants and environmental factors from four subbasins of the Baltic Sea, we found that both environmental factors and contaminants contributed significantly to the variability in the embryo aberrations. The most influential natural drivers associated with high aberration frequency were sandy sediments, high temperature and salinity, and low TOC. Among the chemical contaminants, the significant contributors were metal mixtures (assayed as Tomlinson Pollution Level Index for metals) and total mercury concentrations, low-molecular PAHs (NAP and DBAHA) and high-molecular PCBs (PCB-180). However, high cross-correlations for PAHs and PCBs imply that other exposure metrics, such as chemical activity (Gobas et al., 2018) should be explored as a dose metric to assess the mixture effects of these hydrophobic substances. Nevertheless, our findings justify the application of reproductive aberrations as a biological effect indicator in the Baltic Sea, and, perhaps, other areas where pollution levels are of concern.

CRediT authorship contribution statement

N. Kolesova: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **S. Sildever:** Writing – review & editing, Writing – original draft, Supervision, Funding acquisition. **E. Strode:** Writing – review & editing, Methodology, Investigation, Funding acquisition, Data curation. **N. Berezina:** Writing – review & editing, Funding acquisition, Data curation. **B. Sundelin:** Writing – review & editing, Methodology, Data curation. **I. Lips:** Writing – review & editing, Funding acquisition. **I. Kuprijanov:** Writing – review & editing, Funding acquisition, Data curation. **F. Buschmann:** Writing – review & editing, Data curation. **E. Gorokhova:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All data has been made available as supplementary material.

Acknowledgments

The crew of r/v Salme is thanked for their assistance during the sampling campaigns. Oliver Samlas and Ieva Barda are thanked for assisting in sampling M. affinis and sediment collection. This study was supported by funding from Environmental Investment Centre (KIK 10313, 16300, 17253, RE.4.07.22-0014) [NK, IK], Estonia-Russia Cross-Border Cooperation Program 2014-2020 (HAZLESS project ER90) [IL, IK, NK, NB], European Regional Development Fund Postdoctoral Research Grant 1.1.1.2/16/l/001 (application No 1.1.1.2/VIAA/3/19/465) [ES], Swedish National Marine Monitoring Program (SNMMP) [BS, EG], Interreg Baltic Sea Region co-funded by the European Union (#S007 10/2022-10/2024) [NK, IK, EG, BS, ES], Estonian Research Council grant (PSG735) [SS], European Biodiversity Partnership Biodiversa + (D2P, Proposal number: 2021-473) [IK, NK, EG, BS, ES].

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecolind.2024.111837.

References

Anderson, M., Gorley, R.N., Clarke, K., 2008. PERMANOVA+ for primer: Guide to software and statistical methods.

Anderson, 2017. Permutational Multivariate Analysis of Variance (PERMANOVA). Wiley StatsRef Stat Ref Online:1–15.

Apler, A., Nyberg, J., Jönsson, K., Hedlund, I., Heinemo, S.-Å., Kjellin, B., 2014. Mapping of fiber in the sediments along the Västernorrland coast. Geological Survey of Sweden. [in Swedish] http://resource.sgu.se/produkter/sgurapp/s1416-1-rapport.pdf [accessed 07.07.2023].

Apler, A., Snowball, I., Frogner-Kockum, P., Josefsson, S., 2019. Distribution and dispersal of metals in contaminated fibrous sediments of industrial origin. Chemosphere 215, 470–481.

Bach, L., Fischer, A., Strand, J., 2010. Local anthropogenic contamination affects the fecundity and reproductive success of an Arctic amphipod. Mar. Ecol. Prog. Ser. 419, 121–128.

Baran, A., Mierzwa-Hersztek, M., Gondek, K., Tarnawski, M., Szara, M., Gorczyca, O., Koniarz, T., 2019. The influence of the quantity and quality of sediment organic matter on the potential mobility and toxicity of trace elements in bottom sediment. Environ. Geochem. Health 41, 2893–2910.

Berezina, N.A., Gubelit, Y.I., Polyak, Y.M., Sharov, A.N., Kudryavtseva, V.A., Lubimtsev, V.A., Petukhov, V.A., Shigaeva, T.D., 2017. An integrated approach to the assessment of the eastern gulf of Finland health: a case study of coastal habitats. J. Mar. Syst. 171, 159–171.

Berezina, N.A., Lehtonen, K.K., Ahvo, A., 2019. Coupled application of antioxidant defense response and embryo development in amphipod crustaceans in the assessment of sediment toxicity. Environ. Toxicol. Chem. 38, 2020–2031.

Berezina, N.A., Sharov, A.N., Chernova, E.N., Malysheva, O.A., 2022. Effects of diclofenac on the reproductive health, respiratory rate, cardiac activity, and heat tolerance of aquatic animals. Environ. Toxicol. Chem 41, 677–686.

Blomqvist, S., Lundgren, L., 1996. A benthic sled for sampling soft bottoms. Helgoländer Meeresunters. 50, 453–456.

Byrén, L., Ejdung, G., Elmgren, R., 2002. Comparing rate and depth of feeding in benthic deposit-feeders: a test on two amphipods, Monoporeia affinis (lindström) and Pontoporeia femorata kröyer. J. Exp. Mar. Bio. Ecol. 281, 109–121.

Cornelissen, G., Gustafsson, Ö., Bucheli, T.D., Jonker, M.T.O., Koelmans, A.A., Van Noort, P.C.M., 2005. Extensive sorption of organic compounds to black carbon, coal, and kerogen in sediments and soils: mechanisms and consequences for distribution, bioaccumulation, and biodegradation. Environ. Sci. Technol. 39, 6881–6895.

Cornelissen, G., Pettersen, A., Nesse, E., Eek, E., Helland, A., Breedveld, G.D., 2008. The contribution of urban runoff to organic contaminant levels in harbour sediments near two norwegian cities. Mar. Pollut. Bull. 56, 565–573.

Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm. J. R. Stat. Soc. Ser. B 39, 1–22.

Demšar, J., Leban, G., Zupan, B., 2007. FreeViz-an intelligent multivariate visualization approach to explorative analysis of biomedical data. J. Biomed. Inform. 40, 661–671.

Eriksson, A.K., Sundelin, B., Broman, N. C., 1996. Effects on Monoporeia affinis of HPLC-fractionated extracts of bottom sediments from a pulp mill recipient. St Lucie Press, Florida.

Eriksson Wiklund and Sundelin, 2004. Biomarker sensitivity to temperature and hypoxia – a seven year field study. Mar. Ecol. Prog. Ser. 274, 209–214.

Erm, A., Buschmann, F., Aan, A., 2021. Prioriteetsete ainete ja toiteainete vertikaalsest jaotusest Väinameres ja Liivi lahes. Tallinna Tehnikaülikooli Meresüsteemide instituut. KIK nr 16300.

European Commission, 2008. Directive 2008/56/EC of the European Parliament and of the Council of 17 June 2008 establishing a framework for community action in the field of marine environmental policy (Marine Strategy Framework Directive).

Gobas, F.A.P.C, Mayer, P., Parkerton, T.F., Burgess, R.M., van de Meent, D., Gouin, T., 2018. A chemical activity approach to exposure and risk assessment of chemicals. EPA Public Access. Environ. Toxicol. Chem. 37, 1235–1251.

Gorokhova, E., Löf, M., Halldórsson, H.P., Tjärnlund, U., Lindström, M., Elfwing, T., Sundelin, B., 2010. Single and combined effects of hypoxia and contaminated sediments on the amphipod Monoporeia affinis in laboratory toxicity bioassays based on multiple biomarkers. Aquat. Toxicol. 99, 263–274.

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, 1–11.

Güler, C., Thyne, G.D., McCray, J.E., Turner, K.A., 2002. Evaluation of graphical and multivariate statistical methods for classification of water chemistry data. Hydrogeol. J. 10, 455–474.

HELCOM, 2010. Ecosystem health of the Baltic Sea 2003-2007: HELCOM initial holistic assessment. Balt Sea Environ. Proc. 122.

Havs- och vattenmyndigheten, 2018. Metals and environmental toxins - base and indicative values for sediments for effect-based assesment. The Sea and Water Authority report nr. 31. [in Swedish].

HELCOM, 2018a. Inputs of hazardous substances to the Baltic Sea. Balt. Sea Environ. Proc. 161.

HELCOM, 2018b. Reproductive disorders: malformed embryos of amphipods. HELCOM supplementary indicator report.

HELCOM, 2018c. State of the Baltic Sea- Second HELCOM holistic assessment, 2011-2016. Balt Sea Environ Proc 155:4–7. https://helcom.fi/wp-content/uploads/2019/06/BSEP155.pdf [accessed 07.07.2023].

HELCOM, 2021. HELCOM Baltic Sea Action Plan - 2021 update.

HELCOM ACTION, 2021. Conditions that influence Good Environmental Status (GES) in the Baltic Sea.

HELCOM, 2023. HELCOM Thematic assessment of hazardous substances, marine litter, underwater noise and non-indigenous species 2016-2021. Balt. Sea Environ Proceeding 190.

Höglund, L.O. and Jonsson, K., 2008. Utredning rörande kvicksilvertunnor i Sundsvallsbukten. Kemakata Konsult AB.Dnr 577-8179-07.

Jędruch, A., Kwasigroch, U., Bełdowska, M., Kuliński, K., 2017. Mercury in suspended matter of the Gulf of Gdańsk: origin, distribution and transport at the land–sea interface. Mar. Pollut. Bull. 118, 354–367.

Kankaanpää, H.T., Turja, R., Lehtonen, K.K., 2.22) Advanced monitoring of harmful substances and their effects in the Baltic Sea is desired: A comment on Kanwischer et al. (2021). Ambio 51:1611–1613.

Kholodkevich, S.V., Kuznetsova, T.V., Sharov, A.N., Kurakin, A.S., Lips, U., Kolesova, N., Lehtonen, K.K., 2017. Applicability of a bioelectronic cardiac monitoring system for the detection of biological effects of pollution in bioindicator species in the Gulf of Finland. J. Mar. Syst. 171, 151–158.

Kreitinger, J.P., Neuhauser, E.F., Doherty, F.G., Hawthorne, S.B., 2007. Greately reduced bioavailability and toxicity of polycyclic aromatic hydrocarbons to *Hyalella azteca* in sediments from manufactured-gas plant sites. Environ. Toxicol. Chem. 26, 1146–1157.

Kuprijanov, I., Väli, G., Sharov, A., Berezina, N., Liblik, T., Lips, U., Kolesova, N., Maanio, J., Junttila, V., Lips, I., 2.21) Hazardous substances in the sediments and their pathways from potential sources in the eastern Gulf of Finland. Mar Pollut Bull 170.

Lam, P.K.S., Gray, J.S., 2003. The use of biomarkers in environmental monitoring programmes. Mar. Pollut. Bull. 46, 182–186.

Larsson, J., Lind, E.E., Corell, H., Grahn, M., Smolarz, K., Lönn, M., 2017. Regional genetic differentiation in the blue mussel from the Baltic Sea area. Estuar. Coast Shelf Sci. 195, 98–109.

Lehtonen, K.K., Sundelin, B., Lang, T., Strand, J., 2014. Development of tools for integrated monitoring and assessment of hazardous substances and their biological effects in the Baltic Sea. Ambio 43, 69–81.

Lehtonen, K.K., Turja, R., Budzinski, H., Devier, M.H., 2016. An integrated chemical-biological study using caged mussels (*Mytilus trossulus*) along a pollution gradient in the Archipelago Sea (SW Finland, Baltic Sea). Mar. Environ. Res. 119, 207–221.

Li, Z., Zhang, W., Shan, B., 2022. Effects of organic matter on polycyclic aromatic hydrocarbons in riverine sediments affected by human activities. Sci. Total Environ. 815, 152570.

Löf, M., Sundelin, B., Bandh, C., Gorokhova, E., 2016a. Embryo aberrations in the amphipod Monoporeia affinis as indicators of toxic pollutants in sediments: a field evaluation. Ecol. Indic. 60, 18–30.

Löf, M., Sundelin, B., Liewenborg, B., Bandh, C., Broeg, K., Schatz, S., Gorokhova, E., 2016b. Biomarker-enhanced assessment of reproductive disorders in Monoporeia affinis exposed to contaminated sediment in the Baltic Sea. Ecol .Indic. 63, 187–195.

Lyons, B.P., Bignell, J.P., Stentiford, G.D., Bolam, T.P.C., Rumney, H.S., Bersuder, P., Barber, J.L., Askem, C.E., Nicolaus, M.E.E., Maes, T., 2017. Determining good environmental status under the marine strategy framework directive: case study for descriptor 8 (chemical contaminants). Mar. Environ. Res. 124, 118–129.

Martín-Díaz, M.L., Blasco, J., Sales, D., DelValls, T.A., 2004. Biomarkers as tools to assess sediment quality: laboratory and field surveys. TrAC Trends Anal. Chem. 23, 807–818.

McArdle, B.H., Anderson, M.J., 2001. Fitting multivariate models to community data: a comment on distance-based redundancy analysis. Ecology 82, 290–297.

McCarty, L.S., Mackay, D., 1993. Enhancing ecotoxicological modeling and assessment. Environ. Sci. Technol. 27, 1719–1728.

Meador, J.P., Varanasi, U., Krone, C.A., 1993. Differential sensitivity of marine infaunal amphipods to tributyltin. Mar. Biol. Int. J. Life Ocean Coast Waters 116, 231–239.

Meador, J.P., Krone, C.A., Wayne Dyer, D., Varanasi, U., 1997. Toxicity of sediment-associated tributyltin to infaunal invertebrates: species comparison and the role of organic carbon. Mar. Environ. Res. 43, 219–241.

Nohrén, E., Pihl, L., Wennhage, H., 2009. Spatial patterns in community structure of motile epibenthic fauna in coastal habitats along the Skagerrak - Baltic salinity gradient. Estuar. Coast Shelf Sci. 84, 1–10.

Ohji, M., Takeuchi, I., Takahashi, S., Tanabe, S., Miyazaki, N., 2002. Differences in the acute toxicities of tributyltin between the caprellidea and the gammaridea (crustacea: amphipoda). Mar. Pollut. Bull. 44, 16–24.

Pérez, E., Hoang, T.C., 2017. Chronic toxicity of binary-metal mixtures of cadmium and zinc to Daphnia magna. Environ. Toxicol. Chem. 36, 2739–2749.

Podlesińska, W., Dąbrowska, H., 2019. Amphipods in estuarine and marine quality assessment – a review. Oceanologia 61, 179–196.

Queirós, A.M., Strong, J.A., Mazik, K., Carstensen, J., Bruun, J., Somerfield, P.J., Bruhn, A., Ciavatta, S., Flo, E., Bizsel, N., Özaydinli, M., Chuševe, R., Muxika, I., Nygård, H., Papadopoulou, N., Pantazi, M., Krause-Jensen, D., 2016. An objective framework to test the quality of candidate indicators of good environmental status. Front. Mar. Sci. 3.

R Core Team, 2024. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing.

Raymond, C., Gorokhova, E., Karlson, A.M.L., 2021. Polycyclic aromatic hydrocarbons have adverse effects on benthic communities in the Baltic Sea: implications for environmental status assessment. Front. Environ. Sci. 9, 1–12.

Reutgard, M., Eriksson Wiklund, A.K., Breitholtz, M., Sundelin, B., 2014. Embryo development of the benthic amphipod Monoporeia affinis as a tool for monitoring and assessment of biological effects of contaminants in the field: a meta-analysis. Ecol. Indic. 36, 483–490.

Rüdel, 2003. Case study: bioavailability of tin and tin compounds. Ecotoxicol. Environ. Saf. 56, 180–189.

Schneider, B., Ceburnis, D., Marks, R., Munthe, J., Petersen, G., Sofiev, M., 2000. Atmospheric Pb and Cd input into the Baltic Sea: a new estimate based on measurements. Mar. Chem. 71, 297–307.

Strand, J., Asmund, G., 2003. Tributyltin accumulation and effects in marine molluscs from West Greenland. Environ. Pollut. 123, 31–37.

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.

Strode, E., Barda, I., Suhareva, N., Kolesova, N., Turja, R., Lehtonen, K.K., 2023. Influence of environmental variables on biochemical biomarkers in the amphipod Monoporeia affinis from the Gulf of Riga (Baltic Sea). Water 15, 248.

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.

Sundelin, B., Ryk, C., Malmberg, G., 2000. Effects on the sexual maturation of the sediment-living amphipod Monoporeia affinis. Environ. Toxicol. 15, 518–526.

Sundelin, B., Rosa, R., Wiklund, A.K.E., 2008a. Reproduction disorders in the benthic amphipod Monoporeia affinis: an effect of low food resources. Aquat. Biol. 2, 179–190.

Sundelin, B., Wiklund, A.K.E., Ford, A.T., 2008b. Biological effects of contaminants : the use of embryo aberrations in amphipod crustaceans for measuring effects of environmental stressors. ICES Tech. Ma.r Environ. Sci. 41, 1–23.

Tairova, Z. and Strand, J., 2022. Biological effect measurments in *Gammarus* spp. and *Corophium volutator* as indicators of toxic effects of hazardous substances in Danish coastal waters. Aarhus University. Technical Report from DCE - Danish Centre for Environment and Energy 237.

Tanner, P.A., James, J., Chan, K., Leong, L.S., 1993. Variations in trace metal and total organic carbon concentrations in marine sediments from Hong Kong. Environ. Technol. (United Kingdom) 14, 501–516.

Tansel, B., Fuentes, C., Sanchez, M., Predoi, K., Acevedo, M., 2011. Persistence profile of polyaromatic hydrocarbons in shallow and deep Gulf waters and sediments: effect of water temperature and sediment-water partitioning characteristics. Mar. Pollut. Bull. 62, 2659–2665.

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. Hegoländer Meeresunters 33, 566–575.

Turja, R., Höher, N., Snoeijs, P., Baršiene, J., Butrimavičiene, L., Kuznetsova, T., Kholodkevich, S.V., Devier, M.H., Budzinski, H., Lehtonen, K.K., 2014a. A multibiomarker approach to the assessment of pollution impacts in two Baltic Sea coastal areas in Sweden using caged mussels (Mytilus trossulus). Sci. Total Environ. 473–474, 398–409.

Turja, R., Lehtonen, K.K., Meierjohann, A., Brozinski, J.M., Vahtera, E., Soirinsuo, A., Sokolov, A., Snoeijs, P., Budzinski, H., Devier, M.H., Peluhet, L., Pääkkönen, J.P., Viitasalo, M., Kronberg, L., 2014b. The mussel caging approach in assessing biological effects of wastewater treatment plant discharges in the Gulf of Finland (Baltic Sea). Mar. Pollut. Bull. 97, 135–149.

Turja, R., Sanni, S., Stankevičiūtė, M., Butrimavičienė, L., Devier, M.H., Budzinski, H., Lehtonen, K.K., 2020. Biomarker responses and accumulation of polycyclic aromatic hydrocarbons in Mytilus trossulus and Gammarus oceanicus during exposure to crude oil. Environ. Sci. Pollut. Res. 27, 15498–15514.

Viglino, L., Pelletier, É., St. Louis, R., 2004. Highly persistent butyltins in northern marine sediments: a long-term threat for the Saguenay Fjord (Canada). Environ. Toxicol. Chem. 23, 2673–2681.

Wiklund, A.K.E., Sundelin, B., 2001. Impaired reproduction in the amphipods Monoporeia affinis and Pontoporeia femorata as a result of moderate hypoxia and increased temperature. Mar. Ecol. Prog. Ser. 222, 131–141.