December 2020

Development of a BEBA-based-classification system for sediments

Susanne Heise & Sonja Faetsch (HAW-Hamburg)
Hanne Hetjens, Johannes Teuchies & Lieven Bervoets (University of Antwerp)
Jeanette Rotchell & Will Mayes (University of Hull), Sebastian Höss (Ecossa)
Raf Elst, Els Ryken & Ward de Cooman (VMM)
and the various partners from IDN and Socotec for providing chemical and ecotoxicological data and valuable input to the discussion.

Contents

Sullied Sediments

Sediment Assessment and Clean Up Pilots in Inland Waterways in the North Sea Region

Many of the inland waterways in Europe are under threat due to the introduction of Watch List chemicals that are not currently regulated under the European Water Framework Directive. These chemicals enter our waterways as a result of our day-to-day activities and through industry, and many have been shown to be harmful to wildlife and the wider aquatic environment. Regardless of their source, these pollutants accumulate in the sediments in our rivers and canals over time.

Water regulators and managing authorities do not always know the levels, locations or impacts of these pollutants. Nor do they have the tools to assess sediments confidently and make informed environmental management decisions. To address these issues, the Sullied Sediment project partnership of scientific experts, regulators and water managers is developing and testing new tools that will enable stakeholders to better assess, treat and prevent contamination from these chemicals. This work is being carried out at selected sites in the Elbe, Humber and Scheldt river catchments.

The intention of the Sullied Sediments project is therefore to help regulators and water managers make better decisions with regard to the management, removal and disposal of sediments, thereby reducing economic costs to private and public sector organisations, and the impact of these pollutants on the environment.

The partnership is also working to reduce the extent of chemicals entering the water system by raising awareness about what we, as consumers, are releasing into the environment through the use of common drugs and household products. This includes the involvement of volunteers in a sediment sampling initiative across the North Sea Region, which will inform and empower them as water champions in their local communities.

The Sullied Sediments project has been co-funded by the European Regional Development Fund through the Interreg VB North Sea Region Programme with match funding from the 13 partners involved. The project partnership includes public, private, community and voluntary sector organisations based in the United Kingdom, Germany, Belgium and the Netherlands.

The project has been supported under the Interreg VB North Sea Region Programme's third priority, which is focused on a Sustainable North Sea Region, and is led by the University of Hull (UK).

Website: northsearegion.eu/sullied-sediments

Blog: sulliedsediments.wordpress.com

Twitter:@SulliedSediment

Abstract

Introduction: One of the objectives of the Sullied Sediments Project was to develop a "better assessment" framework for sediments, that is environmentally safe but does not increase costs for managers. Work package 3 investigated whether such a framework would need to be biological effect based and whether chemical data should be complemented with biological community evaluation and ecotoxicological information.

Approach: In order to provide the data to answer this research question, 6 sediment sampling surveys were carried out in each of 3 different catchments at 3 sites with very different historical background and chemical-physical characteristics. For each sediment, about 130 chemical contaminants were analysed, up to 10 bioassays performed and 2 biological community indices determined. This report focusses on the added value of ecotoxicological testing for assessing the hazardous properties of sediments, in order to derive a more reliable sediment classification.

Results: It was shown that the information from ecotoxicological testing was not correlated to either chemical data or biological community data. Therefore, biotest data were considered not to be redundant in any assessment framework but of complementary value. From the 10 bioassays that were performed during this project, 6 were selected for a potential biotest battery which fulfilled the following conditions:

- Data were available for at least 50% of the sediments.
- They were not correlated to each other, thus no test was redundant
- They showed a wide response range, indicating their sensitivity to the pollution at the sites

Conclusions: The suggested biotest battery comprised:

- Algae inhibition test (elutriates) with *Rhaphidocelis subcapitata* (endpoint: growth)
- *Daphnia magna* inhibition test (elutriates) (endpoint: swimming mobility)
- Luminescence bacteria test (*Aliivibrio fischeri*) (elutriates) (endpoint: energy metabolism)
- Direct sediment bacteria contact test with *Arthrobacter globiformis* (endpoint: respiratory enzyme activity).
- Direct sediment contact test with the nematode *Caenorhabditis elegans* (endpoint: reproduction)
- Direct sediment contact test with the ostracod *Heterocypris incongruens* (endpoint: growth)

Considering the biotests' inherent variability, an integrative evaluation of the suggested biotest battery responses was carried out, featuring a weight of evidence approach. Four hazard classes were derived: Class 1 - no hazard; Class 2 – potential hazard; Class 3 – moderate hazard with high certainty; Class 4 – severe hazard with high certainty. When all Sullied Sediment samples were classified accordingly, no clear distinction of sites or seasons became obvious. Trends over the life of the project, however, were visible, and are discussed in relation to the river water discharge for the example of the Elbe estuary.

Future perspectives: The advantage of an integrated classification system becomes clear, if compared to a worst-case assessment, which is used by some national regulations or guidelines. Assessment frameworks that only use the bioassay with the most negative result for management decisions are potentially overprotective. This is especially the case in the light of remaining uncertainties as they have been described in the **Sullied Sediments report on Reproducibility of Biotests**.

Even though variability of biotests is in the same range as that of chemical analysis (Heise et al. 2020), assigning all weight of a decision on one test system may lead to valid resistance and opposition from stakeholders regarding ecotoxicological testing. By combining biotests into a test battery and interpreting the results in an integrative way, the issue of assigning unknowingly too much weight to outliers can be overcome. The overall assessment then reflects an impact on the environment with higher certainty, as a potential hazard is concluded from more than the results of one single test species. Therewith, it is not only a question of how high a hazard is assessed to be, but also how certain this assessment is.

Introduction for stakeholders

One of the objectives of the Sullied Sediments project was to develop a “better” assessment framework for sediments and dredged material, that is environmentally safe but does not increase costs for managers substantially. In Europe, most decisions on sediment or dredged material management are based on sediment quality criteria, which may or may not be derived from ecotoxicological data. From all the substances that could be present in sediments and adversely affect the biological community, the quality of dredged material is usually assessed on the basis of a relatively small number of chemicals which usually at least comprises 7 heavy metals and arsenic, as well as PAHs, PCBs, mineral oil and TBT with criteria for either single substances or sum values (Röper and Netzband 2011). In some countries, more substances are routinely measured for in-situ evaluation of sediments (Boden et al. 2020), but then, few regulations in Europe address environmental sediment management (e.g. removal of sediments for environmental reasons). Scientists have been arguing in favour of basing management decisions on at least 3 lines of evidence (LoE): exceedance of chemical quality criteria, indices reflecting the status of the biological community and ecotoxicological effect data (e.g. (Chapman and Wang 2001; Wolfram et al. 2012; Hall et al. 2017). However, all these different LoE have their limitations:

Chemical threshold values, even though most often used to guide decisions, are only available for relatively few substances among the many thousands that potentially end up in the environment, and may underestimate the hazardous potential in sediments. Routine analysis of chemicals in sediment does not inform on ageing effects that may reduce the adverse impact, or assess synergistic or antagonistic interactions of the contaminant cocktail in sediments.

Bioassays, on the other hand, react solely to bioavailable substances. They are performed in the laboratory by exposing test organisms (worms, shrimps, bacteria, algae, water fleas etc) to the elutriates of sediments or to sediments directly. Impairment of their physiological functions (e.g. photosynthesis, growth, reproduction, respiration) is assumed to reflect, what could potentially happen to the biological community in the environment. While these ecotoxicological data integrate the effect of contaminant mixtures, it is not easily possible to identify what substances are responsible for this measured effect – also because we are only aware of a tiny fraction of the pollutants that may be present. The transfer of laboratory data to the environment has some deficiencies: Sediments are usually tested only by a few bioassays, due to economic reasons but also because they are partly labour intensive and take time. But can these few test species represent the biological community? It is necessary to understand, that this is not the intention behind biotesting. With a test battery that comprises 10 to 20 biotests, we can say little more about the diverse biological community than with 4 or 5. Biotest data only give us an indication of the hazard that is present in the sediment and which can potentially to impact biota. They do not predict what actually happens in the environment. If adversely affected, biotests show that substances are present in the environment that can inhibit certain essential functions (e.g. photosynthesis) in representatives of important trophic levels (e.g. water fleas). In order to cover a wide spectrum of possible effects, bioassays that are sensitive to different substances, to different exposure pathways and belong to different trophic levels are combined into biotest batteries.

In order to assess in what status the biological community is, we have to look at the **diversity of organisms** in the sediment, but even this does not tell us everything what we need to know. The biological community may not have yet reacted to recently emitted contaminants, or it may have adapted to historic contaminants. A low diversity may be due to available toxicants, but it may also be the result of a recent oxygen deficiency or weather related saltwater intrusion.

Combining the 3 LoEs into a weight of evidence approach has been suggested as a method to reduce the likelihood of overlooking a hazard and increase the certainty of making a "correct" decision. At the same time, the pattern of the chemical, ecotoxicological and ecological results can inform on the kind of stress that affects the system, much like 3 individual pieces of a puzzle create a picture. This combined approach has been termed the Sediment Quality Triad

by Chapman as early as 1997 (Chapman et al. 1997). Most sediments are moderately contaminated or by specific substances. Depending on the kind of contaminants, the pollution history, emission pathways, environmental conditions etc, they may pose a risk which can not easily be deduced by chemical analysis which should then be complemented by bioassay data and information on the biological community. Only in the rare cases where sediments are either free of contaminants or very highly polluted, all 3 lines of evidence would be expected to be highly correlated, giving no or very high results. This correlations were checked for in this report. They were never more than moderate, which indicates that the sediments we looked at were not unpolluted (when none of the three "tools" would have been shown anything) nor were they very highly hazardous (when all tools would have indicated a high hazard). The probability to overlook a potential hazard is with contaminants that are in between these extremes. No correlation among ecotoxicological, chemical and community data with our data show, that they each provide complementary information to a weight of evidence approach.

In a more detailed view on the character of ecotoxicological data, these are usually based on a number of bioassay results compiled into a biotest battery. The individual assays provide complementary information to the overall ecotoxicological property. A biotest battery should comprise test organisms from several trophic levels with different exposure pathways, and which are sensitive to different substances. When testing a contaminated sediment, probably not all organisms will react to it, but only those that are sensitive to the present substances. If two biotests always react in the same way, there is no added value of having both in the battery and one is redundant and can be excluded. This is why we tested for redundancy.

The approach of applying different bioassays to sediments and then integrating the data on the assumption of complementary information is not common in European regulation. Often, the focus is laid on one (the worst) bioassay, not taking into account the test immanent variability of a living system.

To align our research with regulatory practice, the following information was collected within the project

- What is the status of assessment frameworks for sediment and dredged material in Europe? Are there countries that follow a weight of evidence approach?
- What is the justification for stakeholders to base decision making only on chemical data?
- Can we quantitatively assess their arguments?
- What improvements can be done?
- How can an improved ("better") sediment and dredged material management framework look like?

In the Sullied Sediments project, a number of activities have been carried out in order to provide the path towards a better assessment scheme. Objectives and outcomes are only summarized here. For details, the original documents should be addressed (See Sullied Sediments publications).

Sullied Sediments Database: By sampling sediments 6 times at 3 locations within 3 different watersheds between autumn 2017 and summer 2019, and testing ecotoxicological effects as well as analysing chemical contaminants, nutrients, grain size fractions and assessing the quality of the ecological community, a substantial database was compiled. This data base will provide the grounds for further analysis of the usefulness and meaningfulness of weight of evidence assessments in comparison with current frameworks.

In a workshop in 2018 on "sediment classification and management decision – *in situ* and *ex situ*", 26 participants from 8 countries came together to (1) compare existing regional or national regulations with regard to their components, decision-making and consequences for the catchment management, and (2) to exchange experiences and difficulties with the different frameworks. It became apparent that few national or regional decision frameworks in Europe take ecotoxicological information into account (Figure 1). If done at all, ecotoxicological data are usually only part of a second tier after chemical analysis has been performed. Few frameworks are truly integrated weight of evidence approaches with chemical, biological and ecotoxicological data as lines of evidence with the same weight (e.g. Italy). There is no apparent trend that biological effect-based decisions would be assigned a higher priority within frameworks. While Italy and the new framework in France (Philippe Bataillard, BRGM, personal communication),

assign the same or even a stronger weight to biological effect-based data as to chemical data, other approaches that used to base decisions on ecotoxicology test data have removed them from their classification (e.g. Netherlands) (see Figure 1).

Criticism of sediment managers with regard to ecotoxicological data was discussed and studied among Sullied Sediments partners and summarized in the **Sullied Sediments report "Investigating the Reliability of Bioassays in Ecotoxicology – Addressing Questions of Reproducibility, Uncertainty and Interpretation"**. It became clear during initial discussions, that biotest data were perceived as having a low reproducibility and reliability by stakeholders, even though they were regarded to be of added value in the overall assessment. A round robin test on natural sediment samples was then organized by the Hamburg Port Authority (HPA) to test the interlaboratory variability. The two biotests performed on sediment elutriates (Algae growth inhibition test, luminescence bacteria test test) showed high variability of test results between participating laboratories. It also became clear that sediment properties such as oxygen content, electric conductivity etc, that were measured by each of the participating labs, had changed differently prior to testing. Even though, instructions on how to handle samples and to perform bioassays were handed out, these seem to lack in detail and leave too much room for individual variation.

Intra-laboratory reproducibility was subsequently studied by HAW Hamburg with the same test systems and showed good reproducibility and high resolution of data (for intralaboratory reproducibility, see also the poster presented at the SETAC conference 2019, Annex 1). It was assumed that transport, storage, and small modifications in the test protocols could lead to high variability of the eventual biotest responses. To investigate the criticality of these different steps for the biotests result, an additional investigation was carried out by HAW Hamburg and IDN (Institut Dr. Nowak GmbH Co. KG). It became clear that the time period (in hours) between sampling and testing, as well as certain pre-treatment steps (e.g. whether centrifugation was performed and whether the elutriate was transported or not) had a strong influence on the test results.

A critical literature survey on the reproducibility of ecotoxicological data for sediments concluded that variability of biotest systems is, in principle, in the same order of magnitude as that of current chemical analytical data. In contrast to chemical methods, however, few interlaboratory comparisons on sediment testing have been carried out, which could largely improve performance of biotesting (Heise et al. 2020).

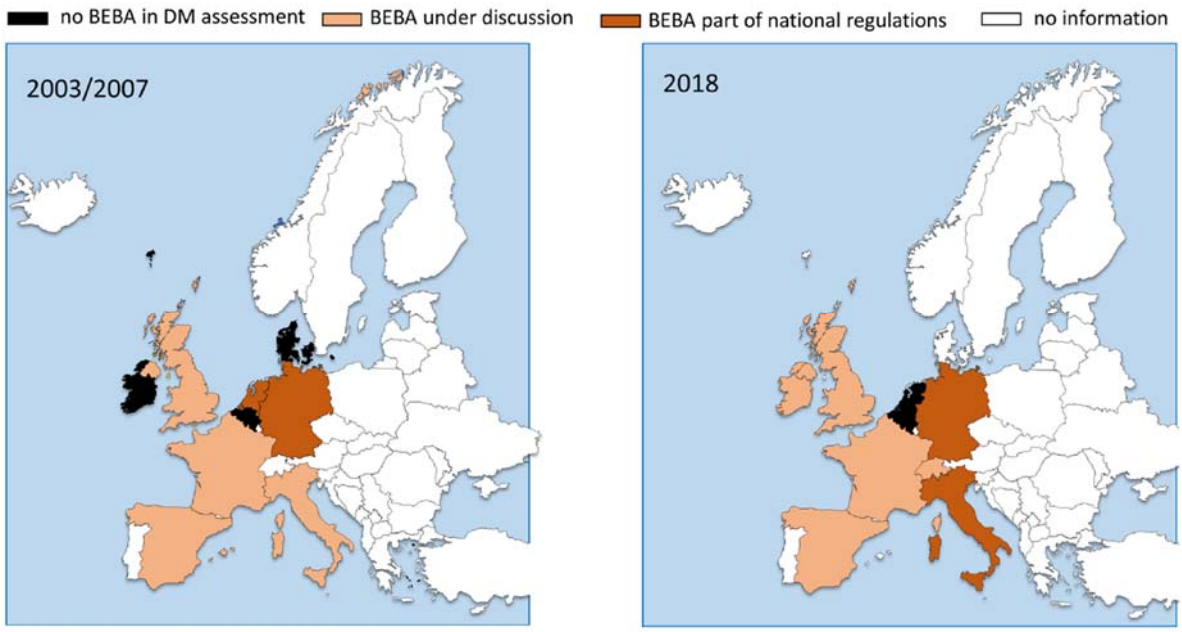


Figure 1: Status of the inclusion of biological effect-based assessments (BEBA) into national regulatory frameworks for dredged material (DM) in European states in 2003/2007 compared to 2018 ((based on den Besten et al. 2003; den Besten 2007 and the outcome of the SedNet & Sullied Sediments Workshop 2018) (from Heise et al. (2020))

In October 2020, a Sullied Sediments workshop on the "role of ecotoxicological data in sediment quality and dredged material assessment frameworks" was carried out. A controversy with regard to the reliability of ecotoxicological tests became clear and new methods with potentials for environmental assessment were presented and discussed.

As an important outcome of Sullied Sediments, the Hamburg Port Authority (one of the major stakeholders in the project), and the Hamburg University of Applied Sciences (HAW, the WP3 leader,), agreed on the following objective:

- The importance of single biotests should not be overstated in the decision-making frameworks of dredged material.
- A more integrative approach towards the interpretation of biotest batteries and their integration into decision making frameworks would be appreciated.

In this respect, this paper describes the development of a BEBA framework for the integrative assessment of sediment quality and for dredged material management.

Methods

A biological-effect based assessment ("BEBA") of sediments was to be developed on the basis of data from 3 sites within each of 3 different watersheds (Elbe estuary, Humber catchment and Schelde river district). Sediment samples from these 9 sites were taken 6 times per catchment between autumn 2017 and summer 2019. Sampling surveys at all catchments were carried out at roughly the same time (within 4 weeks at maximum).

The following data produced by the described methods were considered when developing the assessment scheme:

Biological indices:

The Belgian Sediment Index (BSI): The BSI is based on the composition of macroinvertebrate communities. Depending on the number of taxonomic groups per indicator group, the BSI is calculated corresponding to the diversity of macrozoobenthos with 7-10: good biological quality; 5-6: moderate biological quality; 3-4: poor biological quality; 0-2 poor biological quality.

The NemaSPEAR[%] Index uses the abundance of sensitive freshwater nematodes (or the lack thereof) as an indicator of pollution-related ecological quality. Two endpoints were included in the calculation: The NemaSPEAR index on species level (NemaSPEAR[%]) or on genus level (NemaSPEAR[%]$_{genus}$) for the ecological classification: >56=high; 30-56: good; 20-30: moderate; 10-20: poor; 0-10: bad ecological quality.

Bioassays:

AGI – Algae growth inhibition test: test species: *Rhaphidocelis subcapitata* (Green algae); test matrix: sediment elutriates; test duration: 72 h; endpoint: inhibition (%) of growth compared to the performance in an unpolluted control.

DIT – *Daphnia* immobility test: test species: *Daphnia magna* (Phyllopoda); test matrix: sediment elutriates; test duration: 48 h (acute); endpoint: % immobility compared to the performance in an unpolluted control.

LBT – Luminescence bacteria test: test species: *Aliivibrio fischeri* (bacteria); test matrix: methanol extracts (ex), elutriates (el); test duration: 0.5 h; endpoint: inhibition (%) of luminescence compared to the performance in an unpolluted control.

BCT – Bacteria contact test; test species: *Arthrobacter globiformis* (bacteria), test matrix: whole sediment; test duration: 2 h; endpoint: inhibition (%) of dehydrogenase activity compared to the performance in an unpolluted control.

Thamno – Shrimp test; test species *Thamnocephalus platyurus* (Phyllopoda); test matrix: sediment porewater; test duration: 24 h (acute); endpoint: acute mortality; Inhibition is transformed into effect units (E.U.) which correspond to the inverse value of the EC_{50} from a dilution series (multiplied by 100).

Ha – Amphipod test: test species: *Hyalella azteca* (Amphipoda); test matrix: whole sediment; test duration: 10 d (chronic); endpoint: inhibition (%) of survival and growth compared to the performance in an unpolluted control.

Lv – Lumbriculus test: test species: *Lumbriculus variegatus* (Oligochaeta); test matrix: whole sediment; test duration: 28d (chronic); endpoints: inhibitions (%) of reproduction rate (rr) and specific growth rate (gr) per day compared to the performance in an unpolluted control.

Ma – Chronic aquatic plants test: test species: *Myriophyllum aquaticum* (higher plants); test matrix: whole sediment; test duration: 10d (chronic); endpoint: inhibition (%) of growth compared to the performance in an unpolluted control.

Nema – Nematode test: test species: *Caenorhabditis elegans* (Nematoda); test matrix: whole sediment; test duration: 96h (chronic); endpoints: inhibition (%) of growth (gr) and reduction (r) compared to the performance in an unpolluted control.

Ostra – Ostacod test: test species: *Heterocypris incongruens* (Ostracoda); test matrix: whole sediment; test duration: 6d (subchronic); endpoints: inhibition (%) of survival and growth compared to the performance in an unpolluted control.

Chemical Contaminants

For this analysis, a selection of the most important chemical substances or substance groups that had been analysed in the project was chosen in order to identify strong signals:

TBT - as one of the most important organotin-compounds, which were formerly used as an anti-fouling agent on ships.

The **sum of HBCDD** – the sum of isomers of hexabromocyclododecane was used. HBCDD is used as flame retardant in polystyrene foam and in textiles

PAH EPA Sum – the sum of 16 EPA PAHs. PAHs can derive from natural or anthropogenic sources as a consequence of organic matter combustion.

PFOA and **PFOS**: Perfluorooctanoic acid (PFOA) and Perfluoroctanesulfonic acid (PFOS) are anthropogenic fluorosurfactants that have been listed as emergent contaminants and are of special concern because they are extremely persistant, bioaccumulative, toxic and possibly carcinogenic. Both have been used as the main component in fire fighting foams (AFFF) and PFOS in metal plating.

HCH-isomers – α- and β-HCH are isomers, that were produced during the production of the insecticide γ-HCH

p,p'-DDT, p,p'- DDD, p,p'-DDE – p,p'-DDT used to be widely applied as insecticide and is degraded in the environment to a number of metabolites, of which p,p'-DDD and p,p'DDE are the most important ones.

Chlorinated benzenes - Tri-, Tetra-benzenes and penta-chlorobenzene are intermediates in the synthesis of pesticides and other chemicals. They have also been used as components of dielectric fluids.

Hexachlorobenzene – has been widely applied as a fungicide and released as an industrial by-product.

Hexachlorobutadiene – is mainly an industrial by-product of chlorinated hydrocarbon synthesis (e.g. trichloroethylene). It has been listed as a persistent organic pollutant in the Stockholm Convention in 2015.

Diclofenac and triclosan – have been analysed in this project as PPCPs (pharmaceuticals and personal care products). Diclofenac is widely used as an anti-inflammatory drug. Triclosan is an antimicrobial agent that is present in consumer products such as toothpaste, soaps and detergents.

Total oil and total n-alkanes – These sum parameters indicate a potential oil pollution

2,3,7,8-TCDD – this 2,3,7,8-tetrachlorodibenzo-dioxin is the "Seveso" dioxin, considered to be the most toxic dioxin congener.

OCDD and OCDF - octachlorodibenzo-p-dioxin (OCDD) and octachlorodibenzofuran (OCDF) serve as markers or surrogates for the more toxic polychlorinated dibenzodioxins and dibenzofurans.

The metals **Ag, Cd, Cr, Cu, Hg, Ni, Pb,** and **Zn** as well as the metalloid As were used in this analyses due to their known toxicity. **Al** and **Rb** were included because they can serve as indicators of a grain size distribution and could thus point to a physical rather than chemical influence on the toxicity of bioassays.

For the development of the BEBA-framework, the following hypotheses were tested:

1) A single biotest can inform on the impairment of the biological community.

Using the Sullied Sediments data base, an analysis of correlation between single biotests and the biotic indices that are supposed to indicate the status of the biological community was carried out. A normality test (D'Agostino & Pearson test, software GraphPad prism) showed that only 7 out of the 15 biotest data compiled and the three community indices followed a Gaussian distribution. A non-parametric (Spearman) correlation was carried out after normalization of each parameter to a 100% scale from lowest (set to 0%) to highest (set to 100%) value.

2) Single bioassay responses are correlated with elevated concentrations of certain chemical contaminants.

Using the Sullied Sediments data base, an analysis of correlation between single biotests and the concentrations of the individual substances, normalized to a 100% scale from lowest (set to 0%) to highest (set to 100%) value, was carried out.

3) If there is no correlation between single bioassays and individual contaminants, the ecotoxicological response reflects the overall chemical contamination of the respective sediment.

In order to identify an "overall" chemical contamination, the individual concentrations of those substances that provide the basis for this report were normalized to a 100% scale from lowest (set to 0%) to highest (set to 100%). After that, they were summed up and by this, every substance was given the same weight. Sums were again normalized to a 0 to 100% scale. A correlation matrix with all biotest data was calculated.

4) The quality of the biological community is related to measured concentrations of chemical contaminants and to sediment parameters.

A correlation analysis was done with biotic indices (BSI; NemaSPEAR[%]), chemical data and sediment parameters (percentage dry matter, organic matter, exchangeable NH_4, Nitrate, Nitrite and Phosphate, grain size fraction <63 µm).

5) A biotest combination can be designed that reflects the diversity of the ecological community (NemaSpear or BL-index).

As different labs use different biotest batteries, it would be helpful to indicate, which tests give similar answers and are thus correlated to each other. Rather than using several of these tests in one battery and thus have redundant information, the biotests should be complementary. A correlation analysis was carried out in order to gain insight into the distribution of biotest results over the different sampling sites, seasons and watersheds.

With regard to the diversity indices: High correlations were not expected, as the biotic indices are not only influenced by chemical pollution, but also adaptations of the benthic communities to other environmental factors might occur. A description will be given to which extent the biotest combination correlates with the chemical stress classes.

Results and Discussion

A) Development of a BEBA system for integration into sediment quality assessment frameworks.

1) Significance of results from a single biotest for the impairment of the biological community as measured by the BSI and NemaSPEAR.

Using the community descriptors NemaSPEAR and Belgian Sediment Index, high numbers reflect the best quality, while in all bioassays, high numbers indicate high inhibition or toxicity. Thus, any correlation between biotest data and biotic indices should result in a negative correlation coefficient.

Figure 2 depicts r-values for the non-parametric Spearman correlation. Elevated r-values (> 0.4) are summarized in Table 1 given the respective p-values and samples sizes. Considering the small number of data points for the *Lumbriculus* tests and the accordingly high p-values, this correlation has a high uncertainty. The biotests with *Thamnocephalus* (Thamno), *Hylella azteca* (Ha), and *Caenorhabditis elegans* (Nema gr) show a negative moderate correlation with the Belgian Sediment Index (BSI). The *C. elegans* test shows a trend towards higher growth inhibition, if the NemaSPEAR (ecologic) (NS) indicates a lower quality, while the Ostracode test correlates moderately with the community diversity as indicated by the NemaSPEAR[%]$_{genus}$ (NS gen).

As the power of non-parametric tests is low, a parametric correlation was also carried out. It confirmed (underlined in grey, Table 1) the negative moderate correlations between the nematode test (Nema gr) and the BSI, and between the ostracode test (gi) and the NemaSPEAR[%]$_{genus}$ (NS gen). The positive moderate correlation between the nematode test (Nema gr) and the NemaSPEAR (ecologic) (NS) was also confirmed (data not shown).

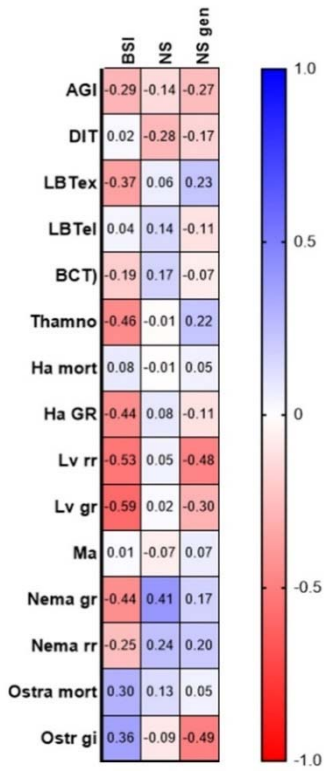


Figure 2: non-parametric Spearman correlation of normalized data from biotic indices and biotest data for all sediment samples

Table 1: Statistical descriptors of correlations of bioassays with biotic indices from Fig. 2 (bold: relationships with low p-Value and high sample size; grey cells: correlations confirmed by parametric correlation).

Correlated parameters	r-value	p-value	Sample size
Thamno / BSI	**-0.46**	**0.009**	**32**
Ha GR / BSI	**-0.44**	**0.024**	**26**
Lv rr / BSI	-0.53	0.145	9
Lv rr / NS gen	-0.48	0.195	9
Lv GR / BSI	-0,59	0.105	9
Nema gr / BSI	**-0.44**	**0.002**	**50**
Nema gr /NS	**+0.41**	**0.002**	**54**
Ostr gi / NS gen	**-0.49**	**0.002**	**36**

Conclusion:

The endpoint of all invertebrate tests, except the Ostracod test, reflect to some extent the quality of the macrozoobenthos community whereby the database of the *Hyalella azteca* and the *Lumbriculus* tests were too small to be reliable. Growth inhibition (gi) of ostracodes is moderately correlated to the quality of the nematode community as indicated by the NemaSPEAR. There is no correlation with either one of the biotic indices with any of the tests with microorganisms. There is neither a strong (>0.7) nor a very strong correlation (>0.9) of any of the biotests with the biotic-indices.

2) Are single bioassays able to reflect the chemical contamination of single compounds in sediments?

Table 2 depicts the coefficients when correlating the biotest data with the concentrations of the selected chemical analytes. Moderate correlations (r between |0.4| and |0.69|) are marked in light blue if positive and in light red if negative. Strong (r between |0.7| and |0.89|) and very strong (r > |0.9|) correlations are marked in dark blue (positive) and dark red (negative).

Strong correlations appear only twice in the analysis: The results of the shrimp-test (Thamno) are strongly correlated (r=0.717) with the concentration of 1,2,4,- Trichlorobenzene with an extremely small p-value (8.65 E-07). Thus, it is highly unlikely that this is a random correlation (n=36). The growth inhibition of *Lubriculus variegatus* (Lv gr) is strongly negatively correlated to the Sum of PAHs (r=-0.750) and the p-value of 0.025 indicates that the possibility that these results are correlated by chance is 2.5 %.

Some of the biotests are moderately correlated to chemicals, such as the

- AGI to γ-HCH and p,p'-DDD, to a number of chlorinated benzenes, to OCDD and OCDF and to Ag.
- DIT to the sum of HBCDD
- LBT ex to TBT, γ-HCH, DDE and DDD, chlorinated benzenes including HCB (Hexachlorobenzene), OCDD, Ag, Cd, Zn
- Thamno in its responses is similar to LBT ex with moderate correlations with γ-HCH, DDE, chlorinated benzenes, OCDD, Ag, Cd, Zn and Ni.
- Mortality of *Hyallela* is moderately correlated with sum of HBCDD, 1,2,3-Trichlorobenzene.
- Lv rr shows correlation with a large number of substances. Due to the small sample sizes of 9, however, this information is uncertain.
- *Myriophyllum aquaticum* shows a moderate correlation with diclofenac.
- With *C. elegans* ("Nema"), moderate correlations are indicated for a-HCH, two chlorinated benzenes, hexachlorobenzene and hexachlorobutadiene, OCDF, As and Rb. Inhibition of reproduction seems to be moderately correlated with diclofenac.
- The growth inhibition in the Ostracode test (Ostr gi) shows moderate correlation with p,p'-DDE and with total n-alkanes.
- The LBTel and the BCT show no correlation with any of the selected substances.

	AGI	DIT	LBTex	LBTel	BCT	Thamno	Ha mort	Ha GR	Lv rr	Lv gr	Ma	Nema gr	Nema rr	Ostra mort	Ostr gi
TBT	0,279	-0,312	0,498	0,157	0,111	0,356	-0,040	0,040	0,377	0,126	-0,314	0,362	0,175	-0,434	-0,057
HBCDD; Sum	0,001	0,609	0,119	-0,146	-0,005	0,276	0,639	0,240	0,429	0,250	0,082	-0,362	-0,327	0,346	-0,075
PAH EPA Sum	-0,090	-0,012	-0,088	0,124	-0,081	0,133	0,150	-0,199	-0,500	-0,750	-0,037	-0,524	-0,387	0,300	0,165
PFOA	0,024	0,162	0,148	0,003	-0,277	-0,237	-0,186	0,053	0,400	0,267	-0,027	0,144	-0,122	-0,124	-0,116
PFOS	0,394	-0,237	0,309	0,132	-0,137	0,249	-0,082	-0,243	0,333	0,017	-0,312	0,011	-0,176	-0,229	-0,021
aHCH	0,348	-0,351	0,310	0,125	0,107	0,224	-0,208	-0,079	0,287	0,426	-0,075	0,488	0,259	-0,413	-0,021
bHCH	0,396	-0,073	0,247	0,061	0,223	0,178	-0,142	-0,057	0,360	0,460	0,078	0,316	0,209	-0,215	-0,215
yHCH	0,418	-0,148	0,493	0,303	0,026	0,528	-0,125	-0,175	0,587	0,230	-0,412	0,063	-0,165	-0,456	0,026
p-p DDT	0,173	-0,300	0,345	0,048	-0,243	0,144	-0,106	-0,175	0,235	0,218	-0,219	0,153	-0,043	-0,303	0,142
p-p DDE	0,344	-0,164	0,547	0,274	-0,081	0,469	-0,076	-0,372	0,360	0,126	-0,405	0,287	-0,001	-0,229	0,423
p-p DDD	0,411	-0,404	0,472	0,204	-0,062	0,368	-0,301	-0,204	0,117	-0,067	-0,303	0,304	0,066	-0,558	0,090
1,2,3-Trichloro benzene	0,278	0,261	0,540	0,286	-0,035	0,653	0,434	-0,086	0,667	0,283	-0,334	-0,098	-0,347	-0,089	-0,023
1,3,5-Trichloro benzene	0,328	-0,155	0,419	0,179	0,122	0,439	0,138	0,115	0,678	0,644	-0,158	0,503	0,174	-0,346	-0,066
1,2,4-Trichloro benzene	0,406	0,156	0,605	0,281	-0,003	0,717	0,308	-0,093	0,633	0,217	-0,334	0,130	-0,140	-0,213	0,137
1,2,3,4-Tetrachloro benzene	0,427	-0,309	0,521	0,197	0,060	0,635	0,003	-0,118	0,322	0,070	-0,358	0,333	0,098	-0,318	0,174
1,2,3,5-Tetrachloro benzene	0,384	-0,260	0,560	0,188	0,034	0,508	-0,070	-0,123	0,492	0,220	-0,293	0,463	0,191	-0,368	0,158
1,2,4,5-Tetrachloro benzene	0,329	-0,385	0,275	0,193	0,039	0,394	0,186	-0,153	0,198	0,069	-0,267	0,304	0,029	-0,180	0,203
Pentachlorobenzene	0,460	-0,415	0,400	0,159	-0,042	0,331	-0,226	-0,148	0,035	-0,052	-0,277	0,317	0,044	-0,504	-0,011
Hexachlorobenzene	0,451	-0,311	0,492	0,139	0,004	0,319	-0,292	-0,161	0,383	0,267	-0,298	0,441	0,154	-0,553	-0,040
Hexachlorobutadiene	0,180	-0,289	0,395	0,091	0,011	0,255	0,057	-0,004	0,570	0,485	-0,162	0,527	0,243	-0,492	0,118
Diclofenac	-0,348	-0,227	-0,040	-0,170	-0,030	-0,157	0,197	-0,060	-0,200	-0,033	0,502	0,339	0,526	0,126	0,129
Triclosan	-0,245	-0,163	-0,076	0,074	-0,054	-0,187	0,128	-0,104	-0,383	-0,683	0,035	-0,105	-0,068	0,254	-0,277
Total Oil	0,115	0,123	0,187	0,071	-0,298	0,018	0,200	-0,458	0,183	-0,250	-0,327	-0,209	-0,331	0,152	0,132
Total n alkanes	0,138	-0,171	0,085	0,168	-0,220	0,070	0,000	-0,462	0,017	-0,317	-0,357	-0,127	-0,345	0,002	0,461
2378-TCDD	0,223	-0,313	0,288	0,034	0,140	0,283	-0,267	-0,312	0,517	0,617	-0,054	0,315	0,306	-0,229	0,081
OCDD	0,412	0,234	0,462	0,288	0,050	0,515	0,134	-0,192	0,577	0,151	-0,443	0,087	-0,015	0,099	0,208
OCDF	0,444	-0,277	0,375	0,150	0,138	0,295	-0,309	-0,125	0,517	0,417	-0,269	0,570	0,298	-0,352	0,131
Ag	0,469	0,062	0,440	0,249	-0,120	0,613	0,066	-0,115	0,669	0,318	-0,473	-0,040	-0,293	-0,262	0,021
Al	0,216	-0,155	0,034	0,390	0,067	0,154	0,062	-0,053	-0,293	-0,142	-0,086	0,262	-0,100	-0,302	-0,087
As	0,277	-0,411	0,168	0,162	0,018	0,173	-0,500	-0,127	-0,109	0,000	0,001	0,450	0,144	-0,474	0,042
Cd	0,270	0,015	0,684	0,251	-0,312	0,569	0,336	-0,379	0,467	0,050	-0,377	0,019	-0,167	-0,223	0,239
Cr	0,099	0,426	0,307	0,227	0,012	0,389	0,372	0,112	0,460	0,100	0,008	0,007	-0,188	0,063	-0,041
Cu	-0,052	-0,238	0,055	0,028	-0,214	0,189	-0,100	-0,404	-0,117	-0,393	0,059	-0,090	-0,114	-0,023	0,120
Hg	0,168	-0,452	0,275	0,078	-0,138	0,143	0,048	-0,261	0,052	-0,061	-0,067	0,383	0,189	-0,397	0,262
Ni	-0,072	-0,207	0,341	-0,070	-0,147	0,475	-0,068	-0,501	-0,452	-0,351	0,243	0,016	0,202	-0,051	0,109
Pb	0,102	-0,125	0,252	0,122	-0,294	0,280	0,138	-0,255	-0,167	-0,417	-0,435	-0,189	-0,263	-0,014	0,150
Rb	0,163	-0,133	0,172	0,191	0,144	0,192	-0,317	0,073	0,159	0,234	0,084	0,594	0,288	-0,331	0,065
Zn	0,330	0,085	0,473	0,221	-0,166	0,433	0,119	-0,196	0,633	0,217	-0,273	0,153	-0,135	-0,296	0,057

Table 2 Coefficients of the correlation of bioassays with selected contaminants, measured in the Sullied Sediments project.

Conclusion:

Bioassays show no to moderate correlation with the selected chemical substances. With the exception of the shrimp test responses (Thamno) which seem strongly correlated with the concentration of 1,2,4-trichlorobenzene, this points to most bioassay results not being determined by the concentrations of single chemical substances. If at all, the various test systems are sensitive to different substances. No single biotest responds to all substances.

3) Is there a correlation of the bioassays with the overall chemical contamination of the respective sediments?

Three bioassays show moderate correlations with the overall contamination of the sediments, calculated roughly as the sum of all concentrations of the selected substances after normalization of every parameter: LBTex (r=0.546; p=0.00001955), Thamno (r=0.507, p=0.002) and Lv rr (r=0.517, p=0.162), whereby again the small sample size of Lv rr renders the correlation uncertain (data not shown).

The LBTex is the only biotest which is performed on samples that have been extracted with an organic solvent (methanol). Methanol extraction is supposed to extract also those substances that could become bioaccessible. The difference in responses of this bioassay compared to the others may indicate, that the contaminants may not have been readily bioavailable in the other test systems.

Thamno, on the other side, is the only bioassay of those that are considered here, that is performed with porewater and assessed by a dilution series. While the porewater may contain contaminants in concentrated form (compared to e.g. elutriates), it may also be the dilution series which is of importance. This will be taken into account when setting up the classification system.

Conclusion:

Only two biotests reliably show a moderate correlation with the overall degree of contamination in the sediment, as estimated here by the sum of normalized concentrations of all contaminants: LBT ex and Thamno (and Lv but with a small sample size). Sensitivities of organisms towards specific substances will be different and, thus, classifying the sediment pollution as done here is a very rough approach. Nevertheless, the results point towards the significance of bioavailable/bioaccessible fractions of contaminants and towards the importance of taking dilution steps into account.

4) Are biotic indices correlated to chemical contaminants and sediment parameters?

An additional correlation analysis was done with biotic indices, chemical data and sediment characteristics (percentage dry matter, organic matter, exchangeable NH_4, Nitrate, Nitrite and Phosphate, grain size fraction <63 µm). Again, correlation coefficients would be expected to be negative, should the quality of the biological community be negatively impacted by specific parameters.

Only moderate (negative) correlations became apparent and only with BSI and chemical contaminants (see correlation matrix in Annex 2), especially with HCH isomers (alpha and beta), some chlorinated hydrocarbons including HCB and hexachlorobutadiene, OCDF, As but also Rb, which is an indicator of fine grained sediment, and, accordingly, with the grain size fraction <63 µm. As the grain size fraction itself is positively correlated with most of those chemicals, that show correlation coefficients with the BSI between -0.4 and -0.69, the question remains whether the BSI is strongly influenced by the fine sediment structure or whether it is an effect of the contaminants that adsorb to the small sediment fraction. Even though no correlation with ammonium was detected, low oxygen content and elevated NH_4^+-concentrations cannot be excluded as potential stress factors..

The NemaSPEAR index shows no correlation with any of the chemical contaminants and not with the sediment parameters, which might be due to the overall poor ecological quality at all stations.

Conclusion:

The correlation analysis of biotic indices with chemical data and sediment parameters show that the NemaSPEAR index shows no correlation with contaminant concentration or sediment parameters, while the BSI is moderately correlated to a number of substances and to the fine grain size fraction (<63 µm), to which the contaminants will specifically adsorb. These sediments are however, likely to become anoxic, so low oxygen contents may have been one stress factor. A correlation with the organic content and with ammonium concentration, however, is not apparent.

5) Design of a biotest combination to be used as one line of evidence in decision making

The biotest combination should consider different trophic levels and fulfil the following requirements:

(1) No redundant information. As few biotests as possible and as many as necessary should be compiled into a battery. In order to gain as many signals as possible with regard to potential pollution of the sediment, the sensitivities and responses of those biotests that are going to form the battery should be most divers and complement each other.

(2) Sensitivity towards pollution. The different test systems show different response ranges. As sampling sites differed substantially with regard to the different chemical stressors, bioassays for a biotest battery should have a wide

response range, thus show that they are sensitive to the contaminants in the sediments and allow differentiation between samples.

(3) Consideration of test-inherent variability. Biotests are carried out with living organisms and thus show a variability that can vary between test systems, due to the genetic diversity of the test organisms, the matrix, the testing procedure. Thus, for example, the lowest inhibition value that reflects a toxic signal, will differ between assays. This should be considered when translating inhibition into potential hazard.

The results from integrating the biotest responses of the designed biotest battery should then be assessed with regard to environmental relevance, and to practicability and cost efficiency (Chapter B).

5.1. Redundancies / reduction of variables.

If biotest responses are correlated to each other, they tend to deliver the same information. They could thus replace each other but should not be integrated into the same battery.

In order to identify correlations, and thus redundancies, between the applied biotests, a correlation matrix was produced (Figure 3). After analysis of the p-values according to the method of Holm-Sidák with an alpha of 0.05, only the following correlations were significant: The moderate correlation (r=0.64) of the Thamno test with the LBTex results, the low correlation (r=0.39) of BCT with LBT el, and the moderate correlation of the two endpoints for the nematode test with *C. elegans*: Growth and reproduction (r=0.64). A drawback in this analysis is the small number of data. At maximum, 54 samples were examined with some comparisons of data only being based on 6 to 9 simultaneously gathered biotest data (e.g. Ostracode versus Myriophyllum-test).

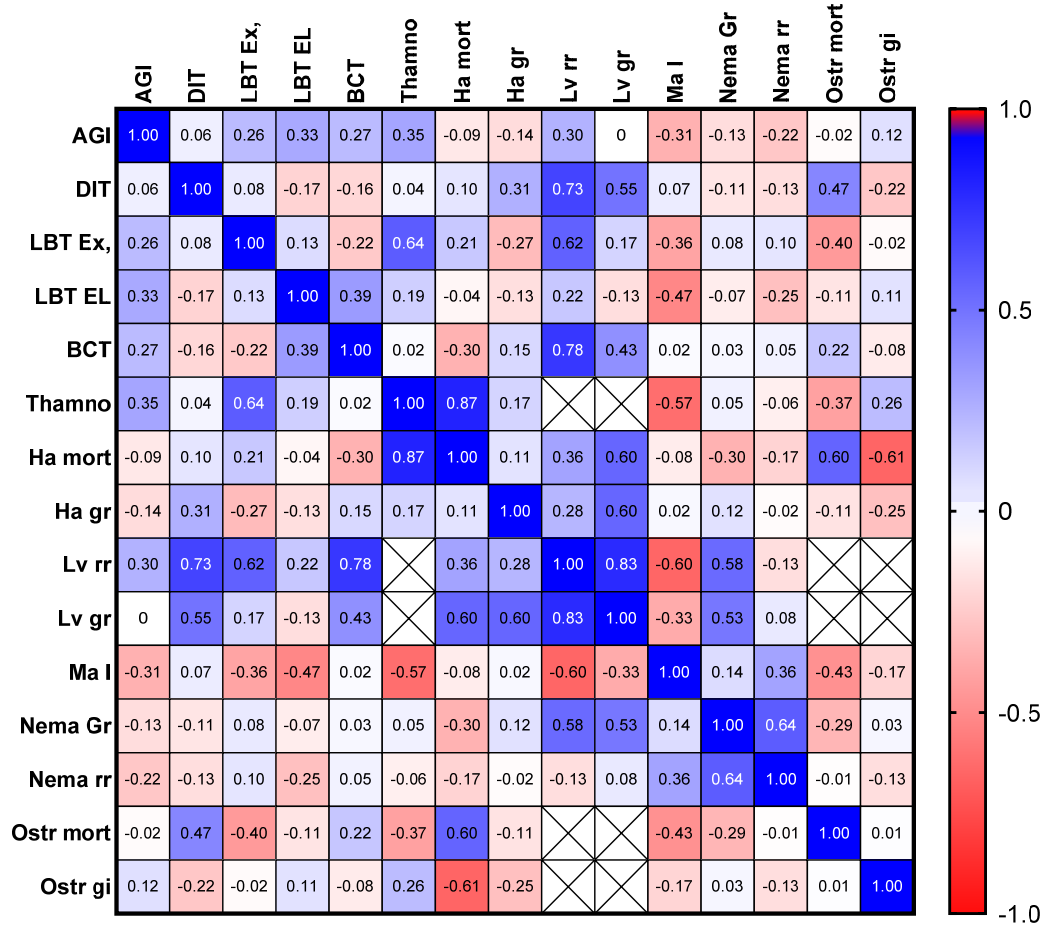


Figure 3 Correlation matrix of all biotests applied in the Sullied Sediment project, depicting the correlation coefficients.

CONCLUSION

No test system can be convincingly omitted from the biotest battery on the basis of its strong correlation with any other of the applied test systems.

5.2. Sensitivity towards pollution

Figure 4 depicts the response ranges of biotest results for all measured sediments. While some bioassays showed responses from stimulation to more than 80 % inhibition, others hardly varied (e.g. LBTex with a response range of 40% points, or Ha mort with 30% points) and therewith did not add a lot of information. More than 75% of all responses from LBTex, Nema gr, Ha mort and Ha Gr were below 20% inhibition. These biotests were not further considered as informative on the basis of these data.

Additionally, those biotests were disregarded for a future biotest battery design, for which too few data (less than 50%) had been provided during the project: Ma, Lv rr and Lv gr. While all 54 samples were measured with the miniaturized and routinely measured assays (AGI, BCT, LBT ex, LBT el, Nema Gr, Nema rr), some otherwise routinely measured assays could not be performed on every sample due to methodological issues (DIT: n=42; Ostr mort: n=32; Ostr gi: n=36; Thamno: n=36).

Some biotests show a jump function: For most samples, inhibition in these assays was very low, but in some cases, the organisms were highly impacted. This is apparent from DIT, Thamno and Ostr gi. As these are ecologically relevant organisms, their strong responses will be considered in the further discussion.

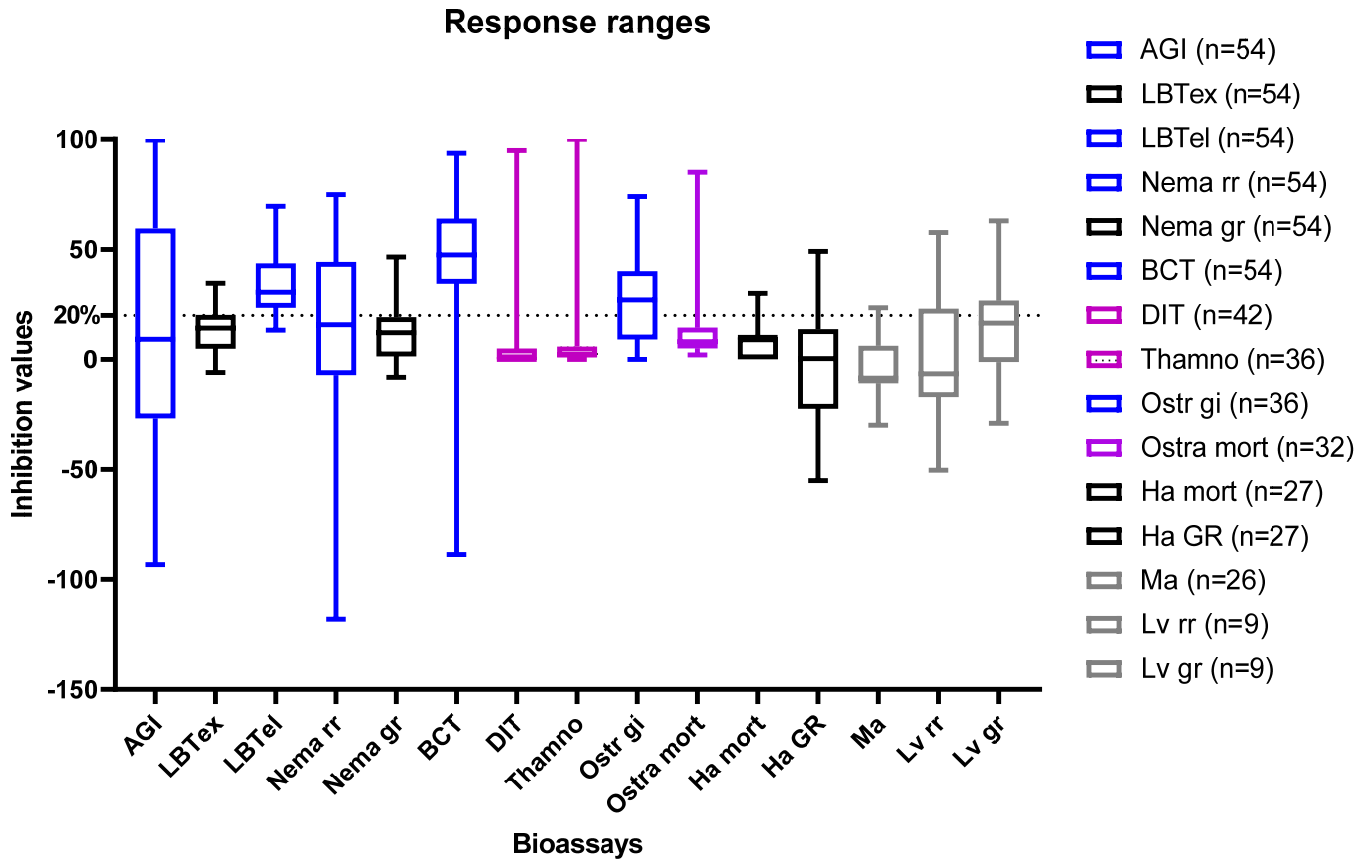


Figure 4: Response ranges of applied biotests in the highest concentrations of sediments or elutriates. The number of measured samples with the respective biotest is given in the legend (grey: number of measured samples <50 %; black: low response range: 75 % of all data below 20 % and no high responses; purple: biotests with apparent jump function: generally low inhibition but strongly impacted by a few; blue: biotests considered as major components of a biotest battery.

5.3 Consideration of test-inherent variability and subsequent comparison of toxicity categories for the respective sediment samples

Test inherent variability of biotests can derive from a number of different sources: Sediment as a matrix is heterogenous by nature and insufficient mixing or small sample sizes can lead to high deviation of replicates; miniaturization and the handling of small volumes and quantities can add to variability; the test organisms, especially if not lab-cultured and genetically identical, will respond differently to stress. For classification purposes, it will be helpful to know these variabilities for biotests. An identified toxicity threshold should differentiate between test-inherent variability and a reliable inhibition signal, thus avoiding false positive results. Additionally, a large variability of biotest results limits the number of toxic categories, which can be differentiated for the respective bioassay (Ahlf and Heise 2007; Heise and Ahlf 2008; Keiter et al. 2009; Höss et al. 2012)

For some bioassay, response ranges have previously been evaluated and translated into toxic categories. These are shown in Table 3. For the other tests, these interpretations have not yet been carried out or still need to be communicated.

Table 3 Toxicity categories (inhibition in % of an uninhibited control), preliminary information

Toxic category		AGI	LBTex	LBTel	BCT	Nema rr	Ostr gi
1	No or low impact	<25	<40	<25	<40	<40	<10
2	Some impact	25-50	40-70	25-50	40-60	40-60	10-40
3	Severe impact	>50	>70	>50	>60	>60	>40
References		Ahlf & Heise, 2005				Preliminary assessment	Preliminary assessment

5.4 The integrative evaluation of biotest combinations / derivation of Hazard Classes (HC)

In Table 4, the first 5 columns show those bioassays, that were selected on the basis of their wide response ranges, and which had been applied to at least 50 % of sediment samples (see discussion above). A wide response range is necessary to enable differentiation between sediment samples with regard to toxicity. All biotest data in Table 4 are coloured to indicate the respective toxic category (according to Table 3). An integrative assessment is then performed following a weight of evidence approach as depicted in Figure 5. As has been pointed out in the Sullied Sediments-report on reproducibility of biotest data, prepared by HPA, HAW and IDN, decisions and assessments should not be based on a single biotest response.

Sediment samples have been coded in the project with regard to the watershed (UK – Humber watershed; BE – Schelde river district; DE – Elbe estuary), the number of the respective sampling survey (“1” for first sampling survey in autumn 2017, to “6” the last sampling survey in summer 2019), and the site of sampling (e.g.: 1 – upstream; 2 – close to WWTP; 3 – downstream).

All sediment samples within the project are herewith assigned to hazard classes (Table 4), based on the following rules:

Class 1 (no hazard): all biotests show no or low impact, in maximum 1 test shows some impact

Class 2 (potential hazard): 2 to 3 biotests show some impact, while the others do not show any.

Only one test shows severe impact while all others show no or low impact

Class 3 (moderate hazard with high certainty): more than 3 tests show some impact, none shows severe impact.

1 or 2 biotests show some impact and 1 test shows severe impact

2 biotests show severe impact while the others show no impact

Class 4 (high hazard with high certainty): at least 3 tests show some impact, and 1 test shows severe impact.

two tests show severe impact and at least one other test shows some impact.

More than 2 tests show severe impact.

Toxicity category						HC
1	1	1	1	1		1
1	1	1	1	2		1
1	1	1	2	2		2
1	1	2	2	2		2
1	2	2	2	2		3
2	2	2	2	2		3
2	2	2	2	3		3
2	2	2	3	3		4
2	2	3	3	3		4
2	3	3	3	3		4
3	3	3	3	3		4
3	3	3	3	1		4
3	3	3	1	1		4
3	3	1	1	1		3
3	1	1	1	1		2
3	1	1	1	2		3
3	1	1	2	2		3
3	1	2	2	2		4
3	3	1	2	2		4
3	3	1	1	2		4
3	3	3	1	2		4

Figure 5: conditions for hazard classes

As pointed out above, additional attention should be paid to the information of those three bioassays, which follow a jump function (data in Annex 3):

Ostr mort: In the samples UK_1.3 and UK_2.3, more than 75% of mortality in the ostracode test had been measured, while responses were low for all other samples. These strong responses were reflected by the high growth inhibitions in the same test (ostr gi). They confirm and do not change the assignment to the respective hazard class.

Thamno: Only one result for the Thamnocephalus test was high: a toxicity of 100% of the BE_1.2 sample. This sample has been categorized as class 4, so this result would have no implications for the hazard assessment.

DIT: Daphnia were strongly inhibited when exposed to the elutriates of 4 sediments: BE_2.1 (HC 3); BE_2.2 (HC3), BE_4.1 (HC2) and BE_4.2 (HC3), shifting the respective samples (except BE_4.1) into the highest Hazard class (see Table 4; "HC with DIT"). Due to this additional information, which improves the reliability of the assessment, it is recommended to include the Daphnia test in the test battery. The resulting classes are depicted in Table 4.

Conclusion:

From the available data, an integrative assessment of ecotoxicological sediment quality could be designed with 6 bioassays, which were chosen for their added value in the hazard assessment (no redundancy) and their response ranges (→ sensitivity to pollution, differentiation between toxicities of sediments). Only those bioassays were considered for which at least 50 % of data were available, as this indicated the practicability of the test procedure. Hazard classes were assigned following a weight of evidence approach and considering the test-inherent variability and responsiveness of the biotests.

Table 4: Biotest data (%) with resulting hazard class (HC) for all samples. Colour coding in biotest columns corresponds to toxicity categories of Table 3. (n.d.: no data)

Sample ID	AGI	Ostr gi	LBT el	Nema rr	BCT	HC	DIT	HC with DIT
BE_1.1	35,34	9,00	63,13	23,50	63,50	4	nd	4
BE_1.2	98,55	33,00	35,46	9,76	85,92	4	nd	4
BE_1.3	94,44	50,00	28,22	2,68	67,13	4	nd	4
BE_2.1	30,02	16,00	31,09	44,06	33,23	3	75,00	4
BE_2.2	87,84	0,00	23,72	-15,38	49,68	3	70,00	4
BE_2.3	36,37	37,00	16,85	15,41	33,48	2	0,00	2
BE_3.1	34,38	24,00	30,14	-12,26	32,29	2	5,00	2
BE_3.2	-18,72	39,00	44,07	30,79	44,32	2	25,00	2
BE_3.3	8,65	39,00	52,18	0,60	42,56	3	0,00	3
BE_4.1	-25,10	30,00	27,71	-27,88	36,65	2	95,00	3
BE_4.2	97,27	33,00	28,30	-40,85	47,07	3	95,00	4
BE_4.3	32,87	34,00	29,84	-72,78	38,20	2	0,00	2
BE_5.1	-36,88	nd	22,50	42,42	48,27	2	5,00	2
BE_5.2	-69,55	nd	32,72	18,34	47,70	2	45,00	2
BE_5.3	30,46	nd	65,39	33,86	46,47	3	0,00	3
BE_6.1	-24,37	nd	32,85	12,15	67,08	3	0,00	3
BE_6.2	-65,40	nd	30,83	4,73	43,57	2	0,00	2
BE_6.3	27,68	nd	66,98	5,20	49,26	3	0,00	3
DE_1.1	91,49	23,00	15,71	33,19	-5,86	3	nd	3
DE_1.2	75,76	11,00	63,26	53,44	79,39	4	nd	4
DE_1.3	99,67	9,00	37,17	52,88	75,13	4	nd	4
DE_2.1	86,96	40,00	53,29	3,33	34,01	4	0,00	4
DE_2.2	-2,76	1,00	38,20	-1,59	60,47	3	0,00	3
DE_2.3	22,58	8,00	34,41	41,55	66,88	3	0,00	3
DE_3.1	89,57	43,00	37,97	-6,81	58,69	4	0,00	4
DE_3.2	82,46	34,00	57,20	22,62	40,96	4	0,00	4
DE_3.3	10,33	3,00	34,16	44,85	65,45	3	0,00	3
DE_4.1	28,56	48,00	48,21	-48,71	58,60	3	0,00	3
DE_4.2	57,30	40,00	17,85	53,02	47,25	4	5,00	4
DE_4.3	-31,82	12,00	17,99	57,23	53,75	2	0,00	2
DE_5.1	35,19	nd	25,15	51,22	10,00	2	0,00	2
DE_5.2	-55,34	nd	22,60	74,84	44,05	3	0,00	3
DE_5.3	-44,26	nd	16,61	70,18	44,75	3	0,00	3
DE_6.1	39,66	nd	46,31	4,27	14,11	2	0,00	2
DE_6.2	-5,46	nd	19,01	27,58	48,59	1	0,00	1
DE_6.3	-15,25	nd	26,85	52,29	69,19	3	0,00	3
UK_1.1	8,35	17,00	37,14	20,07	87,69	3	nd	3
UK_1.2	9,54	0,00	21,78	-8,20	82,26	2	nd	2
UK_1.3	98,83	58,00	43,55	27,05	93,68	4	nd	4
UK_2.1	8,84	11,00	30,98	-13,86	-11,11	2	nd	2
UK_2.2	-13,48	0,00	22,70	4,40	-88,68	1	nd	1
UK_2.3	-14,46	74,00	13,21	-33,10	22,05	2	nd	2
UK_3.1	-22,64	14,00	23,63	-15,86	34,52	1	0,00	1
UK_3.2	-2,38	7,00	29,47	-5,94	44,52	2	5,00	2
UK_3.3	-9,47	44,00	30,04	69,80	52,82	4	0,00	4
UK_4.1	95,86	41,00	52,94	-100,14	73,07	4	0,00	4
UK_4.2	66,05	3,00	69,55	-112,06	74,84	4	0,00	4
UK_4.3	-55,80	53,00	67,95	-118,08	48,76	4	0,00	4
UK_5.1	-55,40	nd	18,96	59,32	0,87	1	0,00	1
UK_5.2	-69,82	nd	22,29	52,75	-54,11	1	0,00	1
UK_5.3	-58,48	nd	27,72	8,86	8,54	1	0,00	1
UK_6.1	-85,88	nd	32,88	67,11	60,35	4	0,00	4
UK_6.2	-93,22	nd	27,81	16,13	46,07	2	5,00	2
UK_6.3	-68,48	nd	24,79	22,31	54,42	1	0,00	1

B) Application of resulting hazard classification to Sullied Sediment samples:

Applying the derived hazard classification, from 54 samples that were taken between autumn 2017 and Summer 2019, 7 are considered to be not hazardous, 15 of potential hazard, 14 of moderate hazard with high certainty, and 18 of severe hazard with high certainty (Table 5).

Table 5: Distribution of Hazard Classes per watershed.(bold: all samples of one survey fall into the same category)

Watershed	Non-hazardous	Potentially hazardous	Moderately hazardous	Severely hazardous
DE	DE_6.2	DE_4.3 DE_5.1 DE_6.1	DE_1.1 DE_2.2 DE_2.3 DE_3.3 DE_4.1 DE_5.2 DE_5.3 DE_6.3	DE_1.2 DE_1.3 DE_2.1 DE_3.1 DE_3.2 DE_4.2
UK	UK_2.2 UK_3.1 **UK_5.1** **UK_5.2** **UK_5.3** UK_6.3	UK_1.2 UK_2.1 UK_2.3 UK_3.2 UK_6.2	UK_1.1	UK_1.3 UK_3.3 **UK_4.1** **UK_4.2** **UK_4.3** UK_6.1
BE		BE_2.3 BE_3.1 BE_3.2 BE_4.3 BE_5.1 BE_5.2 BE_6.2	BE_3.3 BE_4.1 BE_5.3 BE_6.1 BE_6.3	**BE_1.1** **BE_1.2** **BE_1.3** BE_2.1 BE_2.2 BE_4.2

From Table 4 and Table 5 the following observations can be drawn:

- There is no site in any watershed, which always indicates a high or low hazard.
- UK sampling sites showed lower overall hazards, with all samples from sampling survey 5 showing the lowest category while the previous sampling 4 resulted in highest hazard classes.
- Belgian and German samples both showed a trend of becoming less hazardous with time: While most sediments of the earlier surveys (1, and part of 2 for Belgium; 1 to 3, and part of 4 for Germany) fall into category 3 and 4, the later sampled sediments were categorized into hazard classes 1 to 3.
- In contrast to a principal component analysis using the chemical-physical data, which revealed partly strong differences between other sampling stations (see Annex 4), but similar to the diversity indices for the biological community (see Annex 5), no apparent pattern of HC classes can be observed – neither with regard to sampling station or catchment, nor with regard to seasons – only a trend with time for Belgian and German samples.
- Thus, ecotoxicological tests can deliver an additional line of evidence for sediment classification in form of potential toxic effects of sediments, adding value to the assessment on the basis of chemical contamination and of biological quality. This will be discussed in the following using the example of the Elbe Estuary.

Example: Discussion of sediment hazard classes in the Elbe estuary.

The 3 sampling sites in Germany can be shortly characterized as follows:

- Site 1, “Stover Strand”, river-km 592, is the most upstream sampling site in the Elbe Estuary and strongly influenced by historically contaminated sediments that are transported downstream from sub-catchments which had been polluted by mining and industry up to the 1980s. This contaminant transport is dominated by high water discharge (Heise et al. 2007; Heise et al. 2008).

- Site 2, “Köhlbrand”, at approx. river-km 626, is located in area of Hamburg Port. It still receives part of the contaminant load reaching the estuary from upstream, even though concentrations are diluted due to extensive mixing processes due to tidal cycles. This station is close to the discharge of the waste-water treatment plant (WWTP) “Köhlbrandhöft/Dradenau”. The WWTP treats the waste-water of 2.3 million inhabitants from a catchment area of 300 km^2. Here, sediment can be potentially exposed to contaminants typical for effluents from urban WWTPs, such as pharmaceuticals, estrogenic compounds and personal care products (Reemtsma et al., 2013).
- Site 3, “Wedel”, (approx. river-km 640) is located about 15 km further downstream of site 2 and close to the Mühlenberger Loch, the biggest freshwater Wadden area in Europe, designated a Flora Fauna Habitat (FFH) and EU-Bird protection area. The site receives marine sediments that are transported upstream from the North Sea with the flood, but is expected to still carry the imprint of the historical pollution of the catchment.

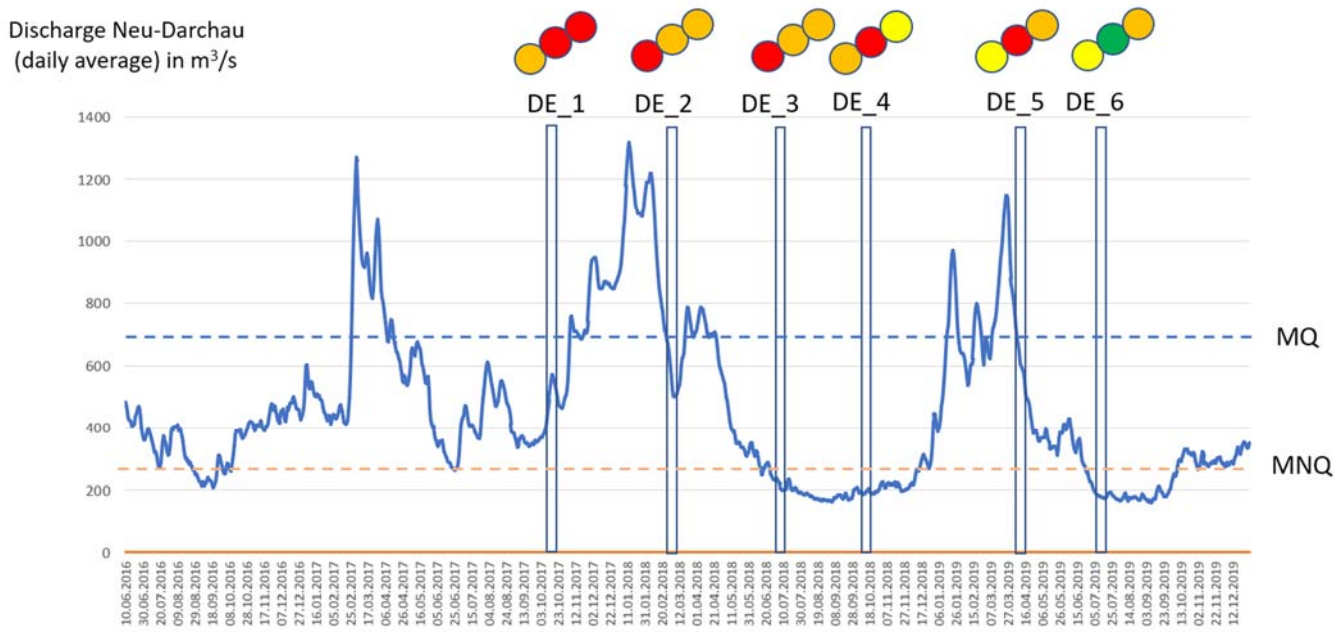


Figure 6: Water discharge at Neu-Darchau, upstream of Site 1 (Stover Strand), from Summer 2016 to the end of 2019 (Source: Wassergütemessnetz Hamburg). Vertical bars indicate the sampling dates with the resulting hazard classes for the three sites in circles (from left to right) in the order upstream to downstream (Stover Strand, Köhlbrand, Wedel). Horizontal dashed lines indicate the average water discharge (MQ) and the average low water discharge (NMQ) (Source: Undine) at Neu-Darchau.

Figure 6 shows the discharge at Neu-Darchau, the water level gauge close to Stover Strand, and the hazard classes of the 3 sites per sampling survey. These point towards a slight reduction of potential toxic impact over the course of the project. Also visible is the trend of decreasing water level in 2019 compared to 2018. From July to November (2018) or September (2019), the water discharge of the river sank below the average low water discharge. The peak of water discharge in March 2019 was well below the peak of 2018. As pointed out above, contaminant transport is flood dominated in the Elbe River due to the still large historical legacies in upstream sediments. The highest contaminant load during the course of the project will have entered the estuary with the elevated discharge in 2018. The following low water levels will not have resulted in increased contaminant transport from the river catchment, but they facilitate upstream transport of marine material toward Hamburg. This may explain the non-hazardous classification of Site 2 (Köhlbrand) during Sampling Campaign 6.

C) Quantitative Comparison of the derived Integrated Hazard Classification with Classification based on Worst Case Results

How big is the difference between an integrative assessment and an assessment based on worst cases? In the following, the hazard classes (with Daphnia-Test results "DIT") that were derived from the Sullied Sediment data (Table 4) are compared with the results when choosing the worst single test result.

The detailed comparison for all cases is shown in the Annex 3 (column: "HC based on 1 test only"). The shift in hazard assessment is demonstrated by Table 6. The integrative assessment allows a higher resolution the classification into four hazard classes (Figure 5) vs. only three categories in the worst-case approach in which the strongest response of the individual biotests in a battery would determine the toxic category of the sediment.

Table 6: Comparison of the number of Sullied Sediment samples (n=54) assigned to hazard classes based on the integrated assessment derived in this report (upper rows) and based on a theoretical worst case (lower rows).

Hazard class	1	2	3	4
Number of assigned samples based on integrative assessment	7	15	14	18

Toxic category	No or low impact	Moderate impact	Severe impact
Number of assigned samples based on worst outcome of single tests	1	19	34

With the integrative assessment, less samples are classified as severely impacted. Single strong responses in one test, that are not backed up by at least elevated inhibitions in others, will not dominate the classification. An assignment to hazard class 3 or 4 by the integrative assessment also reflects the high certainty, by which sediment is classified as hazardous.

Thus, the strength of the classification of sediments not only depends on the sensitivity of individual bioassays but also on the reliability of the final conclusion.

Conclusions

The analysis of the ecotoxicological data of the Sullied Sediments project confirms the hypothesis, that for moderately contaminated sediments, bioassay results are not strongly correlated with either ecological quality or chemical contaminant concentrations. They provide one line of evidence to an overall assessment and provide information that is complementary to chemical data and to the diversity of the biological community in sediments.

The biotest battery that was derived from the Sullied Sediments database consists of 5 bioassay for routine measurements. They were selected on the basis of ease to use, response range and sensitivity to different contaminants. Biotest responses were assigned to 3 toxic categories that had previously been reported or used. Following a weight of evidence approach, 4 hazard classes were derived from the combination of toxic categories of those 5 tests and applied to the sediments from the project. Broadening this battery and adding further test systems only confirmed the previous results.

A comparison between a worst case approach, by which the sediments were classified according to the worst response on the most sensitive bioassay, and the use of the integrated biotest battery resulted in a less severe classification. This may point to an underprotective assessment with more false negative evaluations. This assessment scheme, however, allows for a more resolved prioritization of classes by taking both, the extent of measured toxicity and the number of affected test species, into account. A sediment which caused a strong reaction in 3 or 4 of the bioassays is considered to present a hazard with a higher probability than a sediment, which to which only one test species reacted.

As shown in the application of the classification towards the German sites, interpretation of biotest data may require additional knowledge of the environmental conditions and is often not easily explained or theoretically derived. Hydrological conditions, the time that contaminants have been adsorbed to sediments, and co-occurrence of chemicals or of other sediment parameters may influence the toxic effect of a sample. Consequently, while a BEBA-approach is a useful tool, it should be seen in combination of chemical, environmental and biological data in order to inform on the ecological risk that is posed by sediments with the overall goal of safer decision making.

Literature

Ahlf W, Heise S (2007) Quality assurance of ecotoxicological sediment analysis. In: Förstner U, Salomons W (eds) Sediment Dynamics and Pollutant Mobility in Rivers - Interdisciplinary Approach. Springer, Berlin, pp 380-391

Boden A, Lauryssen K, Teuchies J et al. (2020) Onderzoek van waterbodem en oevers. Code van goede praktijk OVAM. 68

Chapman PM, Anderson B, Carr S et al. (1997) General Guidelines for using the Sediment Quality Triad. Marine Pollution Bulletin 34:368-372

Chapman PM, Wang F (2001) Assessing sediment contamination in estuaries. Environmental Toxicology and Chemistry 20:3-22

Hall AT, Belanger SE, Guiney PD et al. (2017) New approach to weight-of-evidence assessment of ecotoxicological effects in regulatory decision-making. Integrated Environmental Assessment and Management 13:573-579

Heise S, Ahlf W (2008) Potenziale und Grenzen von Fuzzy-Ansätzen in der ökotoxikologischen Bewertung. Tagungsband der BfG zu "Verfahren der ökotoxikologischen (Risiko-)Bewertung in der Umweltsicherung

Heise S, Babut M, Casado C et al. (2020) Ecotoxicological testing of sediments and dredged material: an overlooked opportunity? Journal of Soils and Sediments

Heise S, Calmano W, Ahlf W et al. (2007) Environmental challenges for the Hamburg stretch of the River Elbe and its catchment with regard to the Water Framework Directive. In: Leal W, Hellsten S, Krahn D, Ulvi T (eds) Baltic River Basin Management Handbook. Strategies for Sustainable River Basin Management. pp 47-61

Heise S, Krüger F, Baborowski M et al. (2008) Bewertung der Risiken durch Feststoffgebundene Schadstoffe im Elbeeinzugsgebiet Im Auftrag der Flussgebietsgemeinschaft Elbe und Hamburg Port Authority. Hamburg. 349

Höss S, Ahlf W, Bergtold M et al. (2012) Interlaboratory comparison of a standardized toxicity test using the nematode Caenorhabditis elegans (ISO 10872). Environmental Toxicology and Chemistry 31:1525-1535

Keiter S, Braunbeck T, Heise S et al. (2009) A fuzzy logic-classification of sediments based on data from in vitro biotests. Journal of Soils and Sediments 9:168-179

Röper H, Netzband A (2011) Assessment criteria for dredged material with special focus on the North Sea Region. Hamburg Port Authority. https://sednet.org/download/Dredged_Material_Criteria_North_Sea_0611.pdf. Accessed 21.12.2019

Wolfram G, Höss S, Orendt C et al. (2012) Assessing the impact of chemical pollution on benthic invertebrates from three different European rivers using a weight-of-evidence approach. Science of The Total Environment 438:498-509

Annexes

Annex 1

Poster on intralaboratory variability of microbiological biotests, presented at the Annual Conference of the Society of Environmental Toxicology and Chemistry, 2019, Helsinki

Microbiological Biotests for Sediment Risk Assessment - a Question of Reproducibility and Reliability

Susanne Heise[1], Sonja Faetsch[1], Samantha Warmuth[1] Nicole Wittig[1]
[1]Hamburg University of Applied Sciences (HAW), Work Group Applied Aquatic Toxicology, Hamburg, Germany; contact: susanne.heise@haw-hamburg.de

Background

- Sediment and dredged material (DM) managers that rely on ecotoxicological data as one line of evidence are concerned with reproducibility, precision and accuracy of biotest results.
- If biotest results that could potentially cause high financial investments are considered not plausible and, thus, are not trusted, ecotoxicity tests may eventually be removed from the decision making process.
- This study was initiated in response to observations of decision makers, that ecotox-testing of sediments gave different results when performed in different laboratories.
- Diverging results may be caused by
  - (1) non-reproducibility of biotests when applied to natural sediments in general.
  - (2) variations in the standard operating procedures of different laboratories
  - (3) influence of transport and storage conditions
  - (4) inhomogeneity of sediment samples
  - (5) non-suitability of the applied classification framework for sediment/DM assessment.
- Of all these potential causes, the first one, a general non-reproducibility of sediment biotesting, is certainly the most destructive one as it directly questions the applicability of ecotoxicological tests as management tool for natural sediment.

Photo 1: S. Warmuth with multicorer at Hohendeicher Hafen

Objective

Investigate the reproducibility of testing natural sediments applying 2 regulatory assays and a bacterial sediment contact test.

Methods

Sediment sampling	Locations	Analysis
- Sampling of sediment cores with multisampler (Eijkelkamp), ca. 80 cm length, 3-5 times every 2 weeks from the same location - The core was divided horizontally into slices of 1.5 to 3 cm thickness.	**Alte Süderelbe (spring 2017)** - Old sidearm of the Elbe river, not tidally influenced - Samples taken after wading into the shallow water (distance from bank: ca 8 m). **Hohendeicher Hafen (autumn 2018)** - Small harbor upstream of Hamburg - Samples were taken from a wooden swimming pier without boat traffic.	**X-ray fluorescence (XRF):** Vertical element profiles were determined to compare stratification of sediment cores. **Biotesting:** • Algae growth inhibition (AGI) assay with *Raphidocelis subcapitata* with eluates (DIN EN ISO 8692) • Luminescence bacteria test (LBT) with *Allivibrio fischeri* with eluates (DIN EN ISO 11348-2) • Bacterial sediment contact test (BCT) with *Arthrobacter globiformis* (ISO 18187:2014, draft)

Results and Discussion

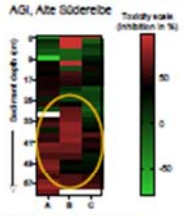


Fig. 1: AGI data of 3 sediment cores (A-C) from Alte Süderelbe

AGI (Location: Alte Süderelbe, Fig. 1)
- >30 cm depth: strong algae growth inhibition in all 3 cores from the station „Alte Süderelbe" (orange circle)
- The upper 30 cm showed less inhibition – also in all 3 cores.
- Close to the surface (0-15 cm), AGI varied between „very low inhibition" (A) and „very strong inhibition" (B, C).
  - All cores, taken 2 weeks apart, showed a similar profile between 80 and 25 cm, regarding the inhibition of algae growth by sediment eluates.
  - The strong differences in the upper 25 cm are most probably due to different sampling conditions: While Core A was taken through a hole in an otherwise closed ice cover with least disturbance (Photo 2), B and C were taken while wading into the water, resuspending the soft sediment.

Photo 2: Ice covered sampling site at Alte Süderelbe

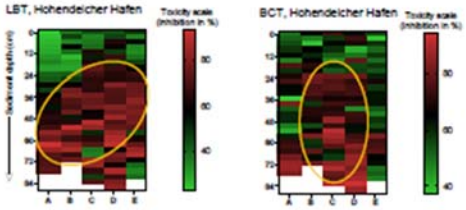


Fig. 2: LBT and BCT data of 5 sediment cores (A-E) from Hohendeicher Hafen

LBT (Location: Hohendeicher Hafen)
- high inhibition in samples below 15 - 30 cm in all cores (orange circle)
- little inhibition in the upper layer (Fig. 2)

BCT (Location: Hohendeicher Hafen):
- >60 cm: increased toxicity; 60 – 24 cm: toxicity high in cores B - D (circle)
- Upper layer (0 - 25 cm) less toxic in all cores
- 3 of 5 cores showed similar toxicity stratification (Fig. 2).
  - Eluate testing (LBT): same trend for all sediment cores
  - Bacterial contact test (BCT): more variable results than eluate testing, probably due to complexity and small scale variation of the matrix
  - Surface sediments were reproducibly non toxic in LBT and BCT.
  - Deeper sediments reproducibly showed increased toxicity.

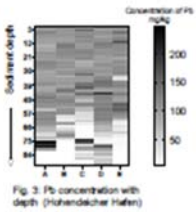


Fig. 3: Pb concentration with depth (Hohendeicher Hafen)

Sediment cores from Hohendeicher Hafen were similar in their stratification but not identical, as shown by Pb-depth profiles (Fig. 3).

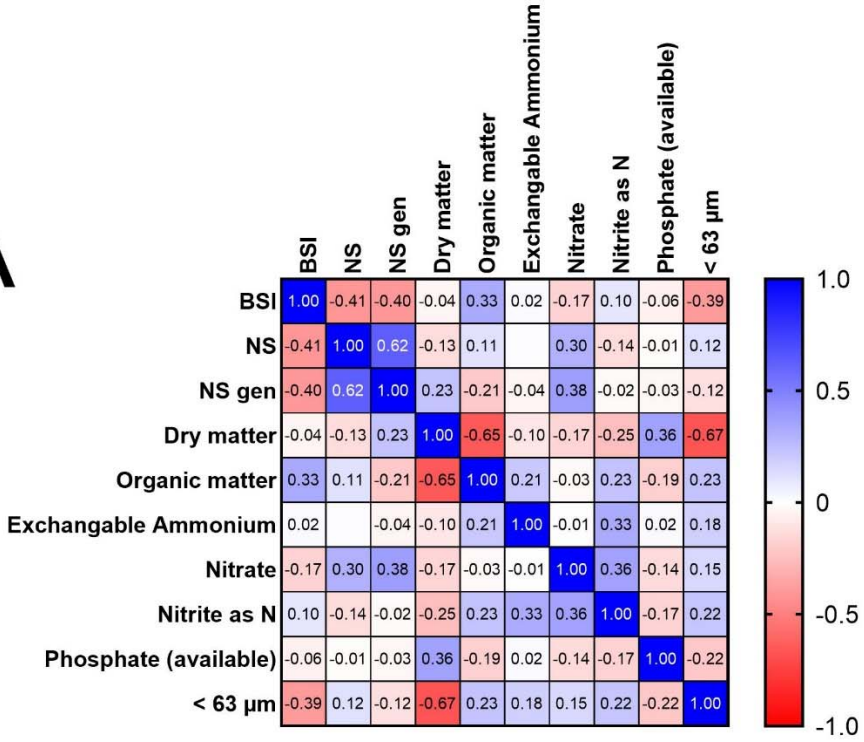
Conclusion & Outlook:

- AGI and LBT data reflected similarity of sediment cores, taken from the same location at different times.
- The direct contact test showed higher variability than eluate tests, which may be due to the complex and inhomogeneous matrix (shown by Pb profiles, Fig. 3).
- Concluding from the level of comparability of data, the tested assays are able to indicate toxicities in natural sediment samples in a reproducible way.
- Within the Sullied Sediments-project, we will next look at test-immanent uncertainties and at the way, regulatory frameworks interpret ecotox data.

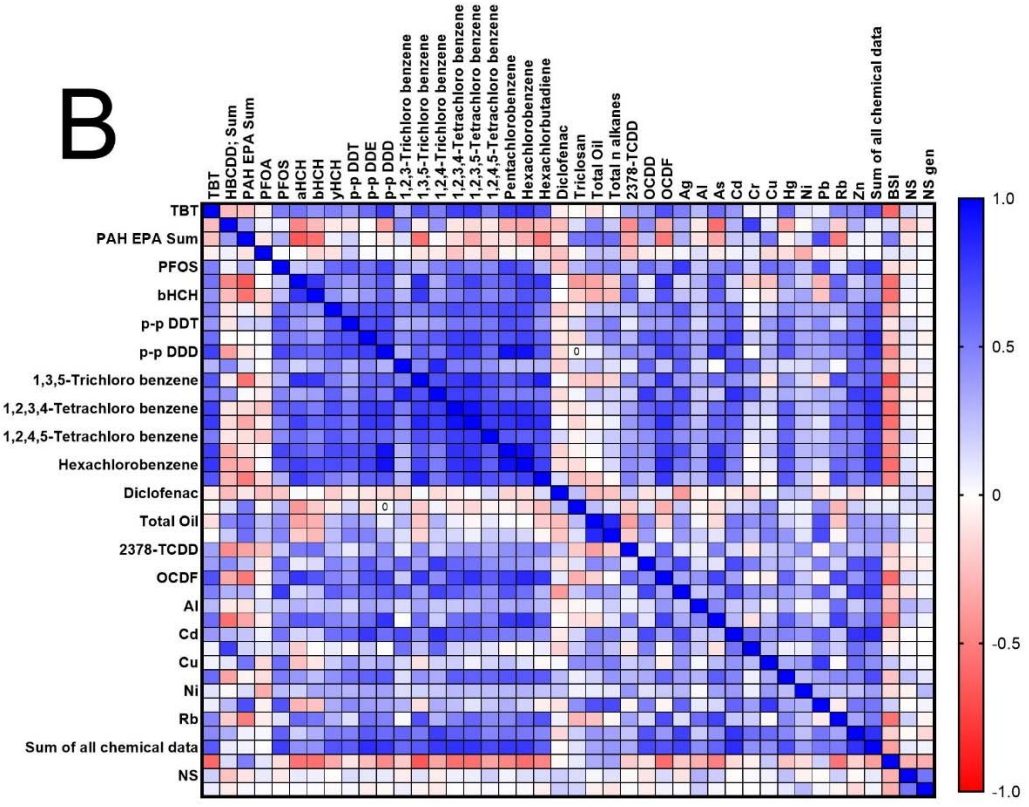
Annex 2

Analysis of correlation of biological community data with (A) nutrients and geochemical sediment properties; (B), with chemical data.

A



B



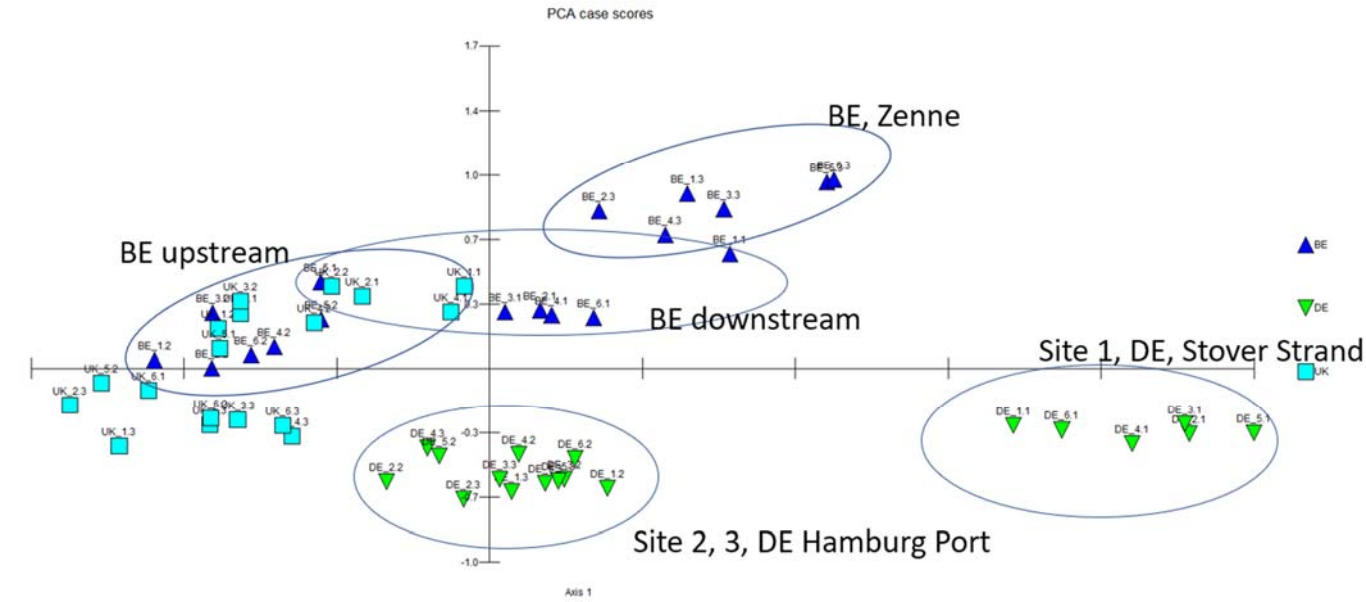
Annex 3:

Biotest data (%) with resulting hazard class (HC) for all samples including the DIT, Thamno, Ostr. Mortality and Nematode-test with the endpoint growth. Colour coding reflects toxicity categories.

Sample ID	AGI	Ostr gi	LBT el	Nema rr	BCT	HC	DIT	HC with DIT	HC based on 1 test only	Thamno	Ostr mort	Nema Gr
BE_1.1	35,34	9,00	63,13	23,50	63,50	4		4	4	38,78	30,00	15,76
BE_1.2	98,55	33,00	35,46	9,76	85,92	4		4	4	100,00	12,00	-2,87
BE_1.3	94,44	50,00	28,22	2,68	67,13	4		4	4	34,55	10,00	1,45
BE_2.1	30,02	16,00	31,09	44,06	33,23	3	75,00	4	4	10,15	7,00	18,54
BE_2.2	87,84	0,00	23,72	-15,38	49,68	3	70,00	4	4	0,01	8,00	-7,09
BE_2.3	36,37	37,00	16,85	15,41	33,48	2	0,00	2	3	7,93	5,00	6,11
BE_3.1	34,38	24,00	30,14	-12,26	32,29	2	5,00	2	3	1,73	8,00	5,81
BE_3.2	-18,72	39,00	44,07	30,79	44,32	2	25,00	2	3	1,90	15,00	22,20
BE_3.3	8,65	39,00	52,18	0,60	42,56	3	0,00	3	4	11,60	3,00	7,46
BE_4.1	-25,10	30,00	27,71	-27,88	36,65	2	95,00	3	4	3,15	25,00	11,97
BE_4.2	97,27	33,00	28,30	-40,85	47,07	3	95,00	4	4	5,51	7,00	4,04
BE_4.3	32,87	34,00	29,84	-72,78	38,20	2	0,00	2	3	5,95	7,00	2,77
BE_5.1	-36,88		22,50	42,42	48,27	2	5,00	2	3			14,64
BE_5.2	-69,55		32,72	18,34	47,70	2	45,00	2	3			6,33
BE_5.3	30,46		65,39	33,86	46,47	3	0,00	3	4			11,47
BE_6.1	-24,37		32,85	12,15	67,08	3	0,00	3	4			12,14
BE_6.2	-65,40		30,83	4,73	43,57	2	0,00	2	3			8,04
BE_6.3	27,68		66,98	5,20	49,26	3	0,00	3	4			12,23
DE_1.1	91,49	23,00	15,71	33,19	-5,86	3		3	4	5,30	5,00	17,12
DE_1.2	75,76	11,00	63,26	53,44	79,39	4		4	4	1,00	8,00	21,31
DE_1.3	99,67	9,00	37,17	52,88	75,13	4		4	4	1,00	15,00	15,70
DE_2.1	86,98	40,00	53,29	3,33	34,01	4	0,00	4	4	13,79	3,00	19,13
DE_2.2	-2,76	1,00	38,20	-1,59	60,47	3	0,00	3	4	1,90		7,09
DE_2.3	22,58	8,00	34,41	41,55	66,88	3	0,00	3	4	3,51	2,00	19,03
DE_3.1	89,57	43,00	37,97	-6,81	58,69	4	0,00	4	4	6,09	3,00	16,33
DE_3.2	82,46	34,00	57,20	22,62	40,96	4	0,00	4	4	1,46	2,00	19,04
DE_3.3	10,33	3,00	34,16	44,85	65,45	3	0,00	3	4	0,01		21,59
DE_4.1	28,56	48,00	48,21	-48,71	58,60	3	0,00	3	4	2,61	5,00	16,69
DE_4.2	57,30	40,00	17,85	53,02	47,25	4	5,00	4	4	2,65		35,89
DE_4.3	-31,82	12,00	17,99	57,23	53,75	2	0,00	2	3	1,78	5,00	24,93
DE_5.1	35,19		25,15	51,22	10,00	2	0,00	2	3			25,11
DE_5.2	-55,34		22,60	74,84	44,05	3	0,00	3	4			17,21
DE_5.3	-44,26		16,61	70,18	44,75	3	0,00	3	4			27,22
DE_6.1	39,66		46,31	4,27	14,11	2	0,00	2	3			18,75
DE_6.2	-5,46		19,01	27,58	48,59	1	0,00	1	3			24,05
DE_6.3	-15,25		26,85	52,29	69,19	3	0,00	3	4			28,17
UK_1.1	8,35	17,00	37,14	20,07	87,69	3		3	4	1,00	25,00	5,16
UK_1.2	9,54	0,00	21,78	-8,20	82,26	2		2	4	1,00	13,00	0,43
UK_1.3	98,83	58,00	43,55	27,05	93,68	4		4	4	1,00	75,00	-2,16
UK_2.1	8,84	11,00	30,98	-13,86	-11,11	2		2	3	0,01	23,00	-0,12
UK_2.2	-13,48	0,00	22,70	4,40	-88,68	1		1	1	0,01	8,00	12,37
UK_2.3	-14,46	74,00	13,21	-33,10	22,05	2		2	4	0,01	85,00	-1,85
UK_3.1	-22,64	14,00	23,63	-15,86	34,52	1	0,00	1	3	0,01	8,00	1,06
UK_3.2	-2,38	7,00	29,47	-5,94	44,52	2	5,00	2	3	2,54	3,00	-4,23
UK_3.3	-9,47	44,00	30,04	69,80	52,82	4	0,00	4	4	0,01	12,00	46,46
UK_4.1	95,86	41,00	52,94	-100,14	73,07	4	0,00	4	4	2,28	12,00	-4,34
UK_4.2	66,03	3,00	69,55	-112,06	74,84	4	0,00	4	4	2,83		-5,42
UK_4.3	-55,80	53,00	67,95	-118,08	48,76	4	0,00	4	4	2,00	5,00	-6,03
UK_5.1	-55,40		18,96	59,32	0,87	1	0,00	1	3			-1,29
UK_5.2	-69,82		22,29	52,75	-54,11	1	0,00	1	3			-8,24
UK_5.3	-58,48		27,72	8,86	8,54	1	0,00	1	3			4,40
UK_6.1	-85,88		32,88	67,11	60,35	4	0,00	4	4			27,59
UK_6.2	-93,22		27,81	16,13	46,07	2	5,00	2	3			19,70
UK_6.3	-68,48		24,79	22,31	54,42	1	0,00	1	3			19,24

Annex 4:

PCA of chemico-physical data of all sites (PC1 versus PC2), showing the clustering according to sites.



Annex 5:

Comparison of relative chemical contamination, quality of the biological community (Belgian sediment index, NemaSpeAR, ecological status and on genus level (NSgen) and Hazard class. Colours reflect the severeness of impact from green: low to red: high (for chemical contamination this is done on a relative scale).

SuSe Sample ID	Sum of all chemical data	Belgian Sediment Index	NemaSpeAR, ecological status	NSgen	HC with DIT
BE_1.1	68	1	21	41	4
BE_1.2	17	6	21	51	4
BE_1.3	55	5	23	25	4
BE_2.1	40	5	26	37	4
BE_2.2	20		19	29	4
BE_2.3	45		16	27	2
BE_3.1	35	7	17	18	2
BE_3.2	28	7	21	11	2
BE_3.3	57	8	14	25	3
BE_4.1	38	8	16	26	3
BE_4.2	24	6	13	25	4
BE_4.3	55	4	18	38	2
BE_5.1	34	8	19	26	2
BE_5.2	31	8	11	27	2
BE_5.3	82	8	25	38	3
BE_6.1	45	8	26	27	3
BE_6.2	25	7	25	24	2
BE_6.3	67	5	38	30	3
DE_1.1	69	2	15	16	3
DE_1.2	37	6	29	24	4
DE_1.3	31		14	20	4
DE_2.1	90	7	29	39	4
DE_2.2	27	2	33	52	3
DE_2.3	25	1	45	51	3
DE_3.1	86	6	12	10	4
DE_3.2	36	6	15	11	4
DE_3.3	28	1	27	31	3
DE_4.1	88	7	8	12	3
DE_4.2	31	1	15	25	4
DE_4.3	28		11	13	2
DE_5.1	100	6	25	30	2
DE_5.2	27	1	28	35	3
DE_5.3	35	0	37	49	3
DE_6.1	74	6	33	40	2
DE_6.2	39	1	51	60	1
DE_6.3	36	0	33	33	3
UK_1.1	44	7	24	31	3
UK_1.2	22	7	9	27	2
UK_1.3	12	9	4	9	4
UK_2.1	28	9	21	30	2
UK_2.2	34	9	13	35	1
UK_2.3	12	10	21	21	2
UK_3.1	25	7	19	25	1
UK_3.2	22	9	19	25	2
UK_3.3	19	10	35	14	4
UK_4.1	35	9	32	24	4
UK_4.2	27	5	21	19	4
UK_4.3	17	10	18	17	4
UK_5.1	29	9	18	30	1
UK_5.2	20	8	4	29	1
UK_5.3	21	9	12	39	1
UK_6.1	19	9	32	52	4
UK_6.2	24	3	31	42	2
UK_6.3	25	9	34	41	1

‘Sullied Sediments’ has been co-funded by the European Regional Development Fund through the Interreg VB North Sea Region Programme with a grant of 2.043.413 euros with equivalent match funding from the partners involved. The project partnership includes public, private and third sector organisations based in the United Kingdom, Germany, Belgium and the Netherlands

.

Partners

The Sullied Sediments project partnership comprises 13 project beneficiaries:

Canal and River Trust (UK)

East Riding of Yorkshire Council (UK)

Ecossa (Germany)

Hamburg Port Authority (Germany)

Hamburg University of Applied Sciences (Germany)

Institut Dr Nowak (Germany)

Openbare Vlaamse Afvalstoffenmaatschappij (Belgium)

Radboud University (The Netherlands)

Socotec UK Ltd (UK)

University of Antwerp (Belgium)

University of Hull (UK)

University of Leeds (UK)

Vlaamse Milieumaatschappij (Belgium)

The partnership also receives expert advice from 12 strategic partners who form our Advisory Group:

East and North Yorkshire Waterways Partnership (UK)

Elbe Habitat Foundation (Germany)

Environment Agency (UK)

Federal Institute of Hydrology (Germany)

Foundation for Applied Water Research (Europe)

Hamburg Ministry of the Environment and Energy (Germany)

Northumbrian Water (UK)

River Hull Board (UK)

Sediment European Network Steering Group (European)

Thames Water (UK)

Vlakwa (water research consultancy) (Belgium)

Yorkshire Water (UK)

Contact us

For more information about this publications, please contact:

 Visit our webspace: www.northsearegion.eu/sullied-sediments

 Follow us on Twitter: @SulliedSediment

 Join our blog: www.sulliedsediments.wordpress.com/

 Join our blog: www.sulliedsediments.wordpress.com/