Topical Scientific Workshop on
Risk Assessment for the Sediment Compartment
7-8 May 2013

"Thought starter" background document on

Effect assessment

This thought starter paper has been prepared by ECHA with the support of the Scientific Committee following a structured expert consultation process. Workshops participants were requested to respond to three sets of questions covering the main discussions areas:

- Problem definition and conceptual model for sediment Risk Assessment
  - Protection goals and ecological relevance
  - Risk characterisation and environmental impact assessment

- Exposure assessment
  - Environmental fate and transfer of chemicals from water to suspended matter and sediment
  - Behaviour processes, within sediment distribution, ageing, bioavailability estimations

- Effect assessment
  - Effect assessment for epi-benthonic organisms, relevant taxonomic groups and experimental tools
  - Effect assessment for benthonic organisms, relevant taxonomic groups and experimental tools

This document reflects the feedback obtained from the participants regarding the third area, effects assessment. Additional information has been obtained from the guidance documents, a review of available scientific literature and the input received from other experts in the field.

Disclaimer: This compilation has been prepared as a background document for facilitating the workshop discussions and does not represent a position of the European Chemicals Agency.

Index

Topical Scientific Workshop on
Risk Assessment for the Sediment Compartment
7-8 May 2013

1. Identification of relevant ecological communities and endpoints in the risk assessment for the sediment compartment

Distinguishing between epi-benthonic/benthonic and lotic/lentic communities could be largely assumed as an issue of exposure assessment and not necessarily effect assessment with the caveat of course that the exposure affects the effects not only in terms of the relevant exposure pathways but also in terms of bioavailability and uptake. For example, truly benthic organisms have a greater exposure to pore water and the actual sediments, which may result in a different contaminant exposure relative to epi-benthic organisms which are predominantly exposed to overlying water and suspended (or recently deposited) particles. In addition, the exercise published by Maltby et al 2005 (ETC 24:379-388) indicates that, for water-only exposure, there are no differences in HC5 for a number of insecticides for lentic and lotic species, and these include species or taxonomic groups that are suitable/recommended for the sediment compartment. For copper, final chronic values derived using SSDs (species sensitivity distributions) are within a factor of 2-3 (not significantly different) when using data sets comprising both pelagic and benthic species, or comprising sole benthic species life stages (Simpson et al 2011, Chemosphere 85:1487-1495). There are several reasons justifying the need to cover all these communities accounting for relevant differences. First, in most ecosystems these different communities are connected and interact in complex ways. In lakes, for example, pelagic community processes influence the flux of organic matter to benthic communities. The biochemical processes controlling bioavailability and exposure differ between benthic and pelagic communities. Second, many key species are "part" of different communities in different life stages, and could therefore have distinct contaminant exposure histories and transfer contaminant burdens from one community to another. For example, an organism that spends part of its life in the benthos (where it would have an exposure history unique to that environment) and then spends a part of its life in the pelagic zone (again, with unique exposure) could effectively transfer contaminants from one community to the next. Obviously, the relevance of this process depends on many factors; for example in historical contaminated sites where there are persistent and bioaccumulative substances (Dioxins, DDT, Mercury, PAH) the main ecological process is the transfer in the foodchain of the contaminants. Finally, services and values derived from communities differ greatly among ecosystems, which have a large influence on the societal relevance to sediment contamination. For example, contaminated sediments in the benthic zone of a deep, fishless oligotrophic lake are unlikely to have the same societal significance (e.g., human exposure) as contaminated sediments in a shallow estuary that is a nursery for seafood and shellfish. Clearly, the connections and uniqueness of these different communities must be distinguished in any sediment-effects assessment. The relevance of the contaminant exposure needs to be a consideration in evaluating the need for a sediment assessment. For instance, exposure of contaminants to benthic organisms in a lotic system is expected to be primarily associated to the dissolved phase and freshly deposited sediments, and therefore to actual emissions, while in a lentic system additional factors may be relevant, including actual and previous emissions, ageing processes, and distribution within the sediment. For generic assessments it is likely to be counter-productive to try and assess flowing and still waters separately (owing to the amount of studies required) and therefore the best approach is probably to include both in freshwater assessments, although it would be beneficial to include some from both environments. Both epibenthic and benthic communities should be looked at. It is even more important to make a distinction between life strategies and feeding strategies. In conclusion, both epi-benthonic and benthonic organisms should be represented in the effect assessment, as they are characterised by different living and feeding modes they should be represented in a base set. The Ecotoxicologically Relevant Concentration may differ between epi-benthos and infauna (certainly important in the risk assessment). Another approach could be to distinguish among differences in functional ecology (trait based approach like differences in feeding mode; filter feeder vs. deposit-feeder) rather than on epi-benthonic vs. in fauna.

Regarding freshwater and marine assessments, in principle they should be separate owing to likely differences in sensitivity and ecology. Marine and transitional waters (estuarine) assessments are probably not required to be separate but some account should be taken of the salinity tolerance of test organisms; combined marine/estuarine datasets should include studies from both groups or include organisms that live in both estuarine and marine environments. Beyond this, a range of routes of exposure are required (e.g. feeding on sediment particles, sediment ingestion, filter feeding, etc.) and this may mean that organisms from different micro-habitats are required (e.g. on and within the sediment). Because of differences in the sensitivity, bioavailability, and community structure, marine and freshwater species require independent evaluations; however, if evidence exists to indicate that there is no difference in sensitivity/tolerance, combining the datasets may be a consideration, although other factors, such as the ecological relevance of the database for addressing the assessed biological community, should be considered. An EFSA review is available for pesticides[1].

Thus, from a scientific perspective the effect assessment should distinguish between different communities. Nevertheless, considering the tests available in the regulatory dossiers (PPP, BP, REACH...), and the validated guidelines, rules for using, pooling or not, refining the available data should be investigated. A guidance document on these rules could be relevant. For organic chemicals with a narcotic mode of action, the species specific distributions available at the moment show no specific differences in sensitivity between marine and freshwater organisms. Therefore, data from the freshwater sediment compartment can be used to assess the marine sediment compartment. For chemicals with other modes of action specific evaluations are needed. Regarding the relevance of the database for assessing the over-all community, in the Risk Assessment for REACH and biocides, when setting the generic Assessment Factors (AF) a larger factor is proposed for the marine environment because of the overall higher diversity (higher number of taxonomic groups in general, although this is not necessarily the case for specific ecosystems/locations) in the marine environment as compared with freshwater environments, although the possibility for demonstrating equivalent sensitivities is also mentioned. In addition, there are issues that may be taken into account in a semi-generic way, for example the reduction of solubility in seawater for non-ionising substances (exposure) or differences in bioavailability between epi-benthic and benthic organisms and physiological adaptations such as osmoregulation (biology).

[1] http://www.efsa.europa.eu/en/supporting/pub/357e.htm

In conclusion, all relevant communities should be covered but there are several ways for getting this coverage:

- Conducting complementary assessments for different communities
- Focusing the assessment in the ecologically-driven (habitat, feeding strategy, etc.) exposure differences, if similar sensitivities among the relevant groups can be assumed (e.g. narcotic or other modes of action not leading to particular sensitivities for organisms only represented in a particular community). This can be further developed by exposure corrections (e.g., based on internal dose), but this approach requires setting the scientific basis for these corrections.
- Selecting the most sensitive groups by combining the potential for exposure and sensitivity (e.g., the risk lines approach)
- Applying extrapolations (e.g., freshwater to marine) based on corrections for exposure/bioaccesibility and a second consideration for the relevance of the dataset regarding the overall biological community.
- Considering the services and values derived from the communities. These differ greatly among ecosystems, which have large influence on the societal relevance to sediment contamination. For example, accounting for the different societal significance (e.g., human exposure) of sediment contamination in a shallow estuary that is a nursery for seafood and shellfish vs. a deep, fishless oligotrophic lake.

Elements for discussion

Do you agree with the conclusions above?
Should the assessment of different exposure pathways (e.g. porewater and particle ingestion) be a mandatory part of any sediment assessment?
When should a generic correction for phys-chem properties (e.g., salinity, pH, hardness) be considered?
How can the relative "similarity" in the toxic response among organisms from the different communities be demonstrated?
How can the relevance of the database to the biological community be assessed?
How should human health protection, in relation to consumption of fishery benthic products, be considered?

Effects assessment on strictly benthic organisms might be more relevant to exposure scenarios where contaminants sorb to and persist within the sediment compartment. Appropriate species can be selected to assess the worst case effects of the contaminant (i.e., via a combination of exposure routes - sediment contact, sediment ingestion and pore-water). In comparison, epi-benthic organisms may be exposed to lower concentrations in the pore water, with a greater contribution of exposure from the overlying water. Being more associated with the uppermost sediment layer, epi-benthic organisms are likely to be exposed to sediment with comparatively smaller and less dense particle sizes, where both quantity and quality of the organic carbon component may differ markedly. They may be exposed to, and feeding on, newly deposited sediments particles, with more potential for exposure to transient sediment contaminants and, perhaps, less exposure to accumulated contaminants forming in the deeper substrate layers. Conversely, however, bioturbation may result in contaminants being released from the deeper bed sediments and becoming more bioavailable to epi-benthic organisms. On the issue of relevance, many of the internationally accepted test methods advocate the use of artificial soil or sediment recipes as the solid matrix for benthic effects assessment, on the basis that results will be more standardised if sediment components are well controlled. This approach to standardisation may be at the expense of environmental realism, whereas the introduction of standard (reference sediments) may provide a more realistic alternative for risk assessment purposes. Standardised test methods have little consideration for the impact of sediment aging processes occurring in the environment. Temporal changes in toxicity due to the formation of non extractable residues (NERs) or their release over time are rarely accounted for in the effect-endpoints generated. Guidance on methods and approaches (when/how) to assess these temporal changes would improve realism in risk assessments for the sediment compartment. The sediment effects assessment should evaluate the impact to the sediment ecosystem/community structure, not to single benthic or epi-benthonic species. Effects assessments should be considered for benthic and epi-benthonic organisms for substances with expected high toxicity to sediment organisms, substances with different toxicity in the aquatic versus sediment compartment, or substances with high accumulation potential in sediment.

Both benthic and epi-benthic organisms are important in effect assessment for sediments and thus all benthic (e.g., oligochaetes, polychaetes, some bivalves) including benthic fish, and epibenthic species (e.g., amphipods at the sediment-water interface), including benthic fish, should be considered. They should be the object of protection as they often form the basis of food chains for many aquatic ecosystems. In addition they alter the structure of the sediment via sediment processing, burrowing and re-suspension. Relevance should certainly encompass the uniqueness of exposure pathways present at the sediment-water interface, and the physiological, morphological diversity of the species which occur in this habitat. All organisms that are in contact with sediments will be exposed to the contaminants within those sediments from both dissolved (porewater, burrow water, overlying water) and particulate sources (sediments, food sources – algae, detritus). The degree or exposure from each source contributes to the net exposure, but this is not always characterised by the definitions of what are benthic and epi-benthonic organisms. Some benthic species utilise sediments almost solely as their 'home', and the major contaminant exposure route is via the overlying waters. For other benthic species, both the dissolved phase and particulate phase (dietary, while feeding) are significant contaminant exposure routes. The dominance of an exposure route (water or particulate) may change, depending on the organism's behaviour and life stage. There exist both benthic and epi-benthic species that ingest large amounts of sediments (dietary exposure) while feeding, however as pointed out above the materials/particles ingested by different organisms may be very different, leading to very diverse dietary exposures. Also epi-benthic algae may experience a different exposure regime than rooted macrophytes that occupy a deeper sediment layer but are also exposed via the water column. For assessments, it is important to ensure that all the possible exposure routes are assessed, and in terms of use bioassays, it is important to use a range of organisms with an adequate range of (characterised) exposure routes. For regulatory assessment, it should be interesting to have a list on the different organisms and their relevant uptakes of substance such as water or ingestion of sediment... to link this information with the behaviour of the substance in water /sediment systems.

There is a need to address the typical 'variability' in sediment ecotoxicology results. Even for the most robust methods (and same laboratories), variability in single tests can be 20-30%, and is often greater for tests with longer durations. Non-contaminant factors (food availability, sediment type) often contribute to the variability.

Benthic and epi-benthonic organisms are an important ecological element, since they comprise a very high number of species with different ecological roles and they cover all trophic levels, from primary producers to consumers. Living in direct contact with the bottom sediments, they are an important link between detrital deposits and higher trophic levels. Moreover, their presence and assemblage can reflect eventual environmental changes occurring in the ecosystem, integrating the information provided by the chemical characterization. Because of considerable variation in sensitivity among species, community composition and the distribution and abundance of species are useful measures of ecological integrity. For these reasons these organisms have been included in many biomonitoring programs and are considered by the WFD (Water Framework Directive) as Biological Quality Elements for the assessment and classification of the ecological status. In particular, for what concerns toxic contamination, the WFD aims for the protection of the whole ecosystem, and has introduced a novel approach to assess ecosystem integrity, using results related to the whole-community response. Following this approach, benthic and epi-benthonic communities are being studied in response to gradients of contamination, producing the first effective toxicity indices. A key element for further discussion is how the experience of these integrated approaches may be used in the design and implementation of prospective risk assessments, including those related to generic marketing authorisation (e.g., REACH, biocides and pesticides).

Elements for discussion
Should benthic and epi-benthonic communities be covered together, independently, or sorted by exposure pathways linked to the ecology (e.g. trait-based approaches)? What are the exposure assessment requirements for each of the options above? How can the experience in assessing and monitoring contaminated sediment be used for improving (higher tier) prospective risk assessment?

2. Selection of taxonomic groups and relevant ecological functions.

Both ecosystem functions (decomposition, primary production, and nutrient cycling) and structure (covering at a minimum survival, growth and reproduction) should be considered. An integrated functional approach focusing on ecosystem services could be ideal, but function is a harder thing to test. In addition function can be redundant so you can remove species or groups of species that perform similar functions and still not be able to distinguish change. From a testing perspective, microcosm and field studies can be used to assess community function more effectively than single species toxicity tests, which can more effectively evaluate ecosystem structure. Endpoint selection should be driven by protection goals (e.g. population parameters for rare/endemic/commercially important species; community properties (e.g. taxon richness, diversity, trophic indices) for general protection.

Both the macro and meio-benthos in sediment should be considered. We generally don't know enough about keystone species in many meio and macro benthic communities to pinpoint these. There is no one species or genus that is the most important but the adage of testing from as wide a number of families, orders, classes, phylum as possible is needed. Ideally it should be important to test up to the community level, but this requires higher tier mesocosm and field studies.

The following groups should be considered as part of a sediment toxicity database:

- Micro organisms (including algae, bacteria) – growth, community composition/ abundance/function (decomposition, primary production, and nutrient cycling).
- Sediment rooting macrophytes (e.g. *Myriophyllum*, *Zostera*), growth/photosynthesis endpoints. Seagrass (*Posidonia oceanica*) meadows; it should be noted that this group is linked to specific habitat conditions and is also exposed via the water column.
- Invertebrates
    - Sediment ingesters and facultative suspension feeders. Feeding strategies: filter, deposit, detritus, scavengers, burrowing.
    - Benthic: Bivalves, oligochaetes, polychaetes, nematodes Epibenthic: Amphipods, gastropods, midge and mayfly larvae, cladocerans.
    - Additional species might be added as follows: a) Where specific toxicological modes of action are suspected, e.g. mollusc (FW, estuarine or marine species) for endocrine disruptors. b) Echinoderms (only present in the marine compartment) and may not be sufficiently protected using the traditional invertebrates given above.
    - Insecta (*Ephemeroptera, Plecoptera, Trichoptera*), *Mollusca*, Crustaceans in freshwater lotic ecosystems; *Diptera, e.g. Chironomidae, Oligochaeta, Mollusca*, Amphipods in freshwater lentic ecosystems
- Slow-moving fish inhabiting bottom waters; larval stages of organisms could also be particularly important since many live on or in the sediment (e.g. some fish and bivalve molluscs).
- Top carnivores

The endpoints depend on the objective, scope, and limitations of the assessment. In the first step the most important taxonomic groups, feeding strategies and micro habitats of organisms inhabiting the sediment should be summarised and sensitive endpoints and ecosystem functions be analysed. Then, the assessment should be based on tests performed with representative species. These species should differ in taxonomic group, feeding strategy and micro habitat (e.g. porewater). The test species and the test design should be selected in a pragmatic way taking into account sensitivity, practicability, reproducibility, etc. A practicable concept should be applied to perform a sediment assessment on the basis of limited available information. Substance properties and mode of action are also important parameters to consider when selecting appropriate test organisms. In addition to the endpoints measuring directly the adverse effects (e.g., reproduction or growth), the use of early-warning signals may be considered. Biomarkers and in particular genomic-based biomarkers and the new approaches related to the adverse outcome pathways (AOP) are promising tools but the ecological relevance of the observed response must always be considered.

For monitoring programmes if the objective is to detect an ecological change that is potentially attributable to contaminants, a large number of measures of community composition are available from which to select an endpoint. Depending on the ecosystem and the contaminants in question, certain sub-groups of organisms are likely to be the most sensitive. Thus, the selection of appropriate indicators should be based on a basic conceptual understanding of contaminant transport and fate, and especially of community dynamics. Key to detecting any ecological change is an understanding of baseline, or reference condition, which is the expected natural or pre-disturbance/contaminated condition. This understanding of baseline should especially include an estimate of natural variability in the ecological endpoint of interest, as well as the expected variability associated with sampling. Without this basic understanding of baseline conditions, inference is limited to merely variation in observed conditions, which may or may not be within the bounds of natural spatio-temporal variability. Change in an endpoint that is beyond the range of natural variation is likely to have real ecological consequences, and therefore be a meaningful measure of ecological effect due to sediment contamination. Assessments should not be limited to merely detecting ecological change. Even if a significant change in an ecological endpoint can be attributed to sediment contaminants, more thorough assessments are required to truly understand the potential consequences of this change to other components of the ecosystem, including humans. *De novo* in-depth ecological studies are not practicable, but additional assessments targeted to key ecosystem components will provide clues about the broader effects of changes in community composition. Included in these follow-up assessments are key measures of processes such as growth, reproduction, population dynamics, and primary production. Process-based measures are usually necessary to understanding the broader implications of contaminant-induced changes in community composition.

Within a biomonitoring framework, the most studied aquatic organisms are at present:

1) Macrobenthic invertebrates: benthic and epi-benthonic groups, with a very high number of species with different ecological roles. They cover all trophic levels among consumers (grazers, shredders, gatherers, filterers, and predators) and they can be found in all micro-habitat types in lotic and lentic ecosystems. Benthic invertebrates are exposed to contaminants in water, sediment and biofilm, providing a direct pathway to higher trophic levels. Different species are characterized by different resilience and resistance traits, providing a large spectrum of ecological adaptations to cope with environmental stress. The most important taxonomic groups are *Insecta* (*Ephemeroptera, Plecoptera, Trichoptera*), *Mollusca*, and Crustaceans in freshwater lotic ecosystems; Crustaceans *Gammarus* and *Diporiea spp*, *Diptera e.g. Chironomidae*, *Oligochaeta* and *Mollusca* in freshwater lentic ecosystems. Crustaceans are important indicators in marine waters. Riverine, lacustrine and marine communities are major Biological Quality Elements to be assessed within the EU-WFD for

evaluating the ecological status of aquatic systems.

2) Biofilms: among epi-benthonic organisms, biofilms are composed by green algae (diatoms and cyanobacteria) forming the autotrophic component and by bacteria, fungi and protozoa composing the heterotrophic part. Due to their omnipresence, their important role in primary production, in nutrient fluxes and trophic cascades as well as their sensitivity to organic and inorganic pollutants, biofilms have been recognized as proper indicators of integrated ecosystem health. Fluvial biofilm communities are one of the major biological quality elements to be assessed within the EU-WFD for evaluating the ecological status of aquatic systems. Multiple endpoints have been developed to assess both structure and function of macrobenthic and biofilm communities. Community abundance and composition are considered within most biomonitoring programs. Moreover, species sensitivity to different environmental stressors is often included as ecological weights of single taxa. Besides, functional endpoints are also considered, including biological (e.g., life cycle, respiration mode, reproduction, body size, etc.) and ecological (e.g., feeding habits, habitat preference, tolerance to stressors, etc.) traits of species. Most of the functional endpoints developed focus on functions directly linked with processes essential for the whole aquatic ecosystem, such as primary production, cycling of nutrients, flow of energy etc. Using a well-defined set of measurements these organisms may allow capturing both acute and chronic effects of a toxicant. The response of functional molecular biomarkers is expected to be quicker than community composition or growth, but the ecological relevance of these responses must be assessed. At site scale, micro-habitats characterized by fine sediment deposition, such as riverine pool habitats, may be preferred to assess the effects caused by toxic contamination of sediments.

A recent development in aquatic ecology is the characterization of the communities according to their functional composition. The advantage of using functional traits instead of taxonomic composition of communities is bound to the *a priori* predictable response of traits to individual stressors. For example, this approach was adopted to study the effects of toxic contamination on invertebrate communities in running waters. The trait approach has been used as framework for deriving species sensitivity to toxicants. For example, Archaimbault et al. (2010) developed a multimetric index based on the benthic macroinvertebrate community described in terms of 22 biological and ecological traits, considered as sensitive to sediment toxicity. Community composition at each site was thus described as relative abundance of trait categories. Based on sets of selected trait categories, a statistical procedure was used to allocate sites to toxic quality classes from the attributes of its benthic macroinvertebrate community. Similarly, the SPEAR (SPEcies At Risk) index, based on species traits, was shown to be highly sensitive to particular groups of toxicants, such as pesticides (Liess et al., 2008). The index is calculated as the proportion between sensitive (SPEAR) and less sensitive (SPEnotAR, "SPEcies not At Risk") species, on the basis of some traits which were considered sensitive to organic toxicants and on life-cycle traits responsible for recovery. The index is applicable across different biogeographical regions in Europe.

For regulatory assessments, the current Guidance recommends different taxonomic groups and feeding strategies but a logical format, such as the one for pelagic assessment (primary producers, primary consumers and secondary consumers), is missing for sediment assessment. A deeper understanding of keystone species and FW ecosystem structure with read across if possible to other compartments would be a good start. Wageningen is currently running an LRI project with new trophic levels included, such as plants and micro-organisms. Other species that could be considered and are frequently found in the literature are *Gammarus*, *Asellus*, insect larvae and molluscs. These organisms would dramatically increase the diversity and usefulness of current options. We should not get lost in endpoints with too high variability such as predator prey interactions and stick to reproductive, growth and biomass endpoints already used. The reality is that few test methodologies exist for addressing a broad range of taxonomic groups. Therefore, at this point we must extrapolate to these groups based on the available data. A tiered pragmatic approach, depending on the mechanism of action is advised.

Elements for discussion
What relevance criteria should be considered for laboratory toxicity test data in sediment risk assessment? And for selecting a representative, generic sediment toxicity database? What additional taxonomic groups are needed for toxicants with specific target species (e.g., herbicides)? Should microbial functions be included in the sediment risk assessment? Should non-invertebrate species (e.g. plants and fish) be included? What relevance criteria should be used when considering data from these groups? How should the invertebrate species be selected (e.g. taxonomy, habitat, feeding mechanism, behaviour, traits/SPEAR, a combination, etc.)? Is it possible to select sets of invertebrate species covering the most relevant groups/exposure pathways/ and ecological roles? How should oral exposure through the diet (secondary poisoning) in fish/amphibians be considered?

There are many experimental tools available depending on the objectives and limitations of the assessment. A review of those relevant for pesticide assessment has been published by EFSA[2].

Currently, the EPM is the most widely used tool. QSARs could also be used but need to have more data to be validated. It is also important to note that data on marine sediment organisms are scarce. Due to this small amount of available data on sediment organisms in regulatory dossiers, it will be worthwhile to investigate what is the best way to derive a $PNEC_{marine\_sediment}$ via $PNEC_{marine\_water}$ or $PNEC_{freshwater\_sediment}$.

Among the basic and simple tools, standardized sediment toxicity tests can be used as the first tool, supported by additional lines of evidence and higher tier studies when needed. The available tests include a broad number of different species, taxonomic groups with different feeding strategies and exposure routes. Insects and midge larvae, mainly *Chironomus sp* for which different testing

[2] http://www.efsa.europa.eu/en/supporting/pub/337e.htm

methods are available (OECD), and crustacean amphipods such as *Hyalella* (USEPA 2000) and the most common in Europe *Gammarus sp*, oligochaetes (*Lumbriculus variegatus* OECD), and nematodes (*Caenorhabditis elegans* ISO) were already present in the 2003 guidance document. These organisms would provide complementary data that is most suitable in a base set for sediment toxicity. Additional tests include the rooted macrophyte *Myriophyllum aquaticum* (OECD and ISO under development), ostracods (*Heterocypris incongruens* (ISO/DIS)), a sediment contact assay with early life stages of fish is also extensively considered for sediment assessment using different endpoints. Polychaetes, amphipods, molluscs such as bivalves are recognised test species for the estuarine and marine environment. Test methods are available for *Arenicola marina*, *Corophium volutator*, *Leptocheirus plumulosus*, and *Amphiascus tenuramis*, and tests with early life stages of sea urchins or bivalves that would be more representative of the sediment-water interface. Larval mollusc and echinoderm tests are often water only tests, the potential for exposure from suspended matter depends on the experimental conditions. Criteria for establishing the relevance of the test organisms and experimental exposure conditions for sediment RA need to be developed. Marine benthic microalgae have been also used in sediment-contact tests and should be considered. Some tests designed to measure water toxicity may also be useful in assessing sediments (e.g. embryo-larval or fish ELS tests) if the test organism satisfies relevance criteria, i.e., the sensitive life stage is in contact with sediment-associated contaminants. Regarding standardisation, the development of OECD test guidelines offering a proper coverage of species/organisms and endpoints is a priority.

The comparison of the relevant groups previously discussed with the current availability of standardised tests indicates an obvious need for more validated and standardized single species laboratory test methods. The validation process is essential in all cases, while the standardisation is particularly relevant in the regulatory context. Taking into account that the standardisation of a test method requires a significant investment, there is a clear need for identifying the current coverage and gaps, to be followed by a careful selection and prioritisation of key tests requiring further development and standardisation.

Experimental systems with greater environmental realism have been developed for some forms of benthic communities: mesocosms and microcosms; transplants and *in situ* caged and colonization studies; or standard benthic community analyses. The greater ecological realism of these approaches is a clear advantage but it is often associated with the specific conditions of the studied community, and this may create difficulties regarding the extrapolation of the results to other ecological conditions. This extrapolation may be solved in a weight of evidence approach leading to a better understanding of the ecotoxicological profile of the substance and the remaining uncertainties; in some cases there are also limitations regarding the statistical power of these approaches due to the reduced number of replicates. Whole-sediment toxicity tests, whole-sediment bioaccumulation tests, pore-water toxicity tests with field-collected sediments or with contaminants spiked into sediments are also available, but these tools should not be used alone for measurement and prediction, rather in a weight of evidence approach.

Collectively, all experimental tools available for sediment-contaminant assessments have limitations and should therefore not be relied upon alone. The most defensible approach to measure and predict effects in retrospective risk assessments includes a combination of carefully designed observational studies and experiments in a weight-of-evidence approach that are targeted to demonstrate key mechanisms and linkages.

Elements for discussion
Do we have sufficient tools for covering all relevant species/taxa and functions? Can these tools be applied in a tiered manner to evaluate missing information/residual uncertainty to develop a comprehensive testing program? What should be the priorities for further developments?

3. Accounting for inter-species sensitivity in the effect assessment

Inter-species sensitivity can be considered in Species Sensitivity Distributions (SSDs) and can also be considered based on mode-of-action. Inter-species sensitivity is the reason we see community-level ecological responses to contaminants. A species' sensitivity to a particular contaminant is influenced by its physiology, behaviour, life history, food preferences, and a host of other traits. Knowledge of these traits, in addition to the chemical properties of the contaminants, is needed to identify appropriate ecological endpoints and predict the ecological consequences of sediment contaminants. It may be possible to take into account inter-species sensitivity to some extent using approaches along the lines of critical body burden. Some of the differences may be due to the relative biovailability and therefore time to reach equilibrium of the substance in the organisms, which depends on both uptake and metabolism/elimination processes. For metals, homeostatic regulation and storage process may play a role. An alternative approach using biological traits showed that organisms' sensitivity to stress is a function of their biology, and can be predicted from species traits such as morphology, life history, physiology and feeding ecology. This approach showed that 4 species traits (skin respiration, insect/crustacean, life-cycle duration, gill respiration) explained 71% of the variability in sensitivity to toxicants within a group of 12 species exposed to 15 chemicals. This approach could revolutionize the SSD concept, showing which species within the community are most susceptible to specific toxicants.

For the sediment assessment, interspecies sensitivity is difficult to resolve using standard bioassays because of the limited suite of organisms for which we have sediment data. Field data and mesocosm experiments are more likely to yield information needed for characterization of interspecies variation. However, these approaches typically have lower statistical power than laboratory toxicity tests; and when several exposure pathways are relevant, the "apparent" inter-species sensitivity may be influenced by the role of each pathway under the specific experimental conditions. Thus, it is important to considered laboratory and field approaches as complementary tools. When quantitative understanding of interspecies variability for response to sediment contaminants is not available, the most conservative approach would be to use the most sensitive species as our criteria for effects assessment, adding the additional uncertainty using appropriate assessment factors.

In general, SSDs principles can be applied to sediment communities but the complexity of the exposure pathways would require careful screening of data to standardise or otherwise control the influence of bioavailability on any endpoint used to construct an SSD. This approach would be implemented as per aquatic SSDs although targeting species that are applicable to the sediment risk assessment. With regard to the minimum species requirements in order to apply the SSD concept, the London workshop (2001) formulated some recommendations for the aquatic compartment only (covering at least 8 taxonomic groups, containing at least 10 NOECs and preferably more than 15 for different species, etc.). Guidance on the minimum number of taxonomic groups needed to apply the statistical extrapolation technique for the sediment compartment was not part of the London workshop, and it can be even questioned if the same criteria developed for the aquatic compartment should be imposed on the sediment compartment. Applying the London workshop criteria to the sediment compartment would ignore: 1) the expected differences in species richness between sediment and water ecosystems, 2) the different exposure conditions and feeding behaviour of the organisms in the sediment (ingestion of sediment, body wall contact, exposure through pore water and overlying water). In addition, very few standardized toxicity test methods exist for benthic species, creating additional difficulties. Extrapolating the London Workshop guidance to sediments may therefore not be appropriate. The focus in establishing a SSD for sediment should instead be based on obtaining a reasonable cross section of the feeding behaviour of all benthic species. The use of fewer species than for the aquatic compartment (but representing different living & feeding conditions) can be covered by the complementary use of other lines of evidence (e.g. mesocoms).

Minimum requirements would depend on the end use of the data (i.e., allowing a substance to be used vs. clean up criteria). A broad range of taxa (different trophic groups, physiologies, feeding mechanisms, reproductive strategies) including plants, animals and microbial composition/function is needed. However, in practice for most chemicals the data is lacking, which makes the approach untenable. Using an SSD to calculate an HC5 would require at a minimum ~10 species, otherwise you would often be extrapolating a value beyond the range of your data. It is implicit to the HC5 concept that if the number of species is sufficient for ensuring intrapolation instead of extrapolation, the HC5 should be expected to be higher than the lowest NOEC or EC10; nevertheless this issue has created some discussions in the regulatory context. For most chemicals, there is not enough data to employ the SSD approach. Whenever you are estimating extreme values within the tails of a distribution like an HC5 you are only likely to get an accurate estimate if you have a lot of data. Additional generic issues associated to the SSD include pooling of data (within species, duration of test, endpoint, etc.), developing SSDs covering all species in a compartment or SSDs using only similar types of organisms, goodness of fit, choice of fit functions, confidence limits, among others. Also, the SSD approach is only protective of the community if the species within the SSD are representative of the community. With the limited suite of organisms for which data exists for a given chemical it is unlikely that those organisms are a good representation of the community you are trying to protect. Despite these limitations, the capacity of SSD approaches, when properly applied, for refining the effect assessment has been recognised for other compartments, also in the regulatory context. Thus, it should be important to consider which specific adaptations and developments are needed for setting guidance on its applicability in sediment risk assessment.

As a first step, emphasis is needed on choosing relevant species for inclusion in the SSD. A more conscious choice of test organisms is needed to represent the actual risk to benthic communities. Elements to be considered include aspects such as exposure potential (e.g., related to choice of habitat or feeding strategy), biotransformation capability (for organic compounds; inefficient biotransformers are generally more susceptible to chemical exposure), ecological importance (i.e., species that are crucial for certain ecosystem functions) and/or the expected mode-of-action of the chemical in question (if this is known) in cases where the chemical has a mode-of-action that targets particular taxonomical groups. These elements should be addressed when presenting the SSD outcomes, and the development of generic guidance would be very useful in the regulatory context.

Even more important is the metrics for constructing the SSD. As stated above, different organisms are exposed by different routes (and combinations of routes) and the distribution of the chemical among the different sub-compartments may be highly influenced by the experimental conditions. If not corrected for bioaccesibility/bioavailability, the SSD may reflect artificial distributions linked to conditions of each test, and even when corrected for bioaccesibility, the metrics used for the correction will affect the sensitivity of each species within the SSD. Therefore there is a need to either construct separate SSDs for sediments with different properties (low vs. high AVS, DOC, sand-silt) or construct SSDs using effects thresholds that account for variability in sediment properties. A possible approach is: 1) fist make SSD for a standard situation relating to relevant bioavailability parameters, then conduct tests to quantify bioavailability normalisation functions on all relevant types of species. 2) Make spot tests on species not used for building the normalisation approach. 3) Then apply the normalisation approach on all species and thereby transform the SSD towards different but relevant abiotic factors determining the bioavailabilities toward the different species. 4) Take account of uncertainty for bioavailability read across (from species with known bioavailability dependence to those only spot checked).

Elements for discussion
How should SSD principles be applied in sediment risk assessments? What are the key characteristics of benthic organisms that should be considered in the development of an SSD? What are the minimum data requirements for application of the SSD and can these requirements be practically met in a standard sediment assessment? Which metrics should be used? Which higher-tier supplementary tools and alternatives to SSD exist and how should they be used? How to conceptually employ bioavailability normalisation (metals, non-ionised organic substances (pH: 5-9), ionised organic substances.)

4. Minimum requirements and use of information from non-sediment dwelling organisms

The minimum requirements should (as noted above) cover a specified range of ecologies/ feeding strategies and taxonomic groups. The PNEC should be based on long-term data as any exposure of the sediment compartment is a long-term exposure.

In some current regulatory systems (e.g. REACH and biocides) a PNEC screening can be derived without using a single sediment toxicity data through the EPM, and a PNEC based on assessment factors can be derived from just one long-term test. The higher tier methods can be a mesocosm study or field-based considering that the most sensitive species is tested, and the analytical measurements are available.

Equilibrium partitioning theory (EqP) is often applied to predict sediment organism toxicity from pelagic organism data using the theory that exposures via water or dietary pathways are in equilibrium with freely dissolved pore water concentrations. Equilibrium partitioning between pore-water and sediment is often modelled using properties that predict binding potential, e.g. log Kow/Koc, to predict binding of non-polar organics to the organic carbon in sediment. The sorption behaviour for some substances is less predictable though (e.g. ionisable compounds). The method also assumes that the sensitivity of sediment-dwelling species is not significantly different to that of pelagic species as long as they are physiologically and ecologically similar. The assumptions have not been validated for a broad range of sediment-contaminants or for organisms at different trophic levels or with different feeding types. Uptake via ingestion of sediment particles may become more important for highly adsorbing chemicals and to account for these uncertainties, compounds with log Kow > 5 or correspondingly high adsorption/ binding behaviour in the case of metals and binding mechanisms not related to lipophilicity, require specific considerations for uptake via ingestion. Even for lipophilic compounds with very high log Kow, binding, bioavailability and bio-accessibility may vary widely according to the actual Kow value and other properties. The organism feeding behaviour should be also considered, at least between total sediment ingestion and more selective feeding approaches. In laboratory tests with artificial sediment, the partitioning process may be different than that expected for natural sediments, and this issue should be also considered.

The currently applied EPM approach under REACH and biocides was developed over a decade ago taking into account the very limited availability of data. The regulatory experience on the EPM suggests that a re-thinking is needed. Basically, for generic prospective risk assessments, the use of EPM in the chemical safety assessment produces the same RCR (risk characterisation ratios) for water and sediment for substances with Kow >5, while for higher Kow values, the difference is just related to a pragmatic generic decision assumed to be conservative. An additional issue frequently observed for substances with very low water solubility is that a PNEC aquatic cannot be derived, or that the information suggests that aquatic exposure is of very low relevance. Those are typical cases for substances particularly relevant for a sediment assessment, and the application of the EPM, even with an additional assessment factor, is highly questionable when exposure via water is assumed to be of low relevance.

A weight of evidence approach is needed to determine the minimum data requirements. An SSD with a lower number of species but with corroborating field data has less uncertainty than an SSD alone, the uncertainty reduction depends on the relevance, comprehensiveness, power and type of field data. At a minimum, an effects assessment should consider multiple taxonomic/functional groups and ecological niches for determining the appropriate method for PNEC derivation.

Specific approaches are required for metals. The equivalent approach to organic substances is the derivation of a screening PNEC using pelagic ecotoxicity data combined with Kd values. However, in the Ni case study sediment toxicity test organisms were tested in both water-only exposures and in sediment exposures. Using sediment-specific Kds, water-only toxicity data were used to estimate sediment toxicity. These estimates were both higher and lower than the toxicity measured in sediment exposures. This places great uncertainty on the use of pelagic data in combination with Kd values for estimating toxicity via sediment exposure. A weight of evidence approach combining data from sediment, pelagic and soil organisms has been suggested based on:

A) Evaluation of the benthic sediment ecotoxicity data, recognizing the importance of organic carbon and the Acid Volatile Sulphide pool to control the chronic toxicity of Me^{2+} towards sediment-dwelling organisms. The derivation of the freshwater HC5-50sediment (benthic SSD) can be been based on the organic carbon normalized dataset, using only low AVS sediments (e.g. the retained database includes 6 species-specific data points representing 62 NOEC values at various sediment chemistry (OC)).
B) Using the EqP approach, HC5-50sediment (EP) values can be derived for a range of EU scenarios, representative for the physico-chemical characteristics of EU surface waters (e.g. the EU scenario's defined in the aquatic effects section using the aquatic BLM). The scenario-specific HC5-50sediment (EP) values can be calculated from the scenario-specific aquatic HC5-50 values and the application of respectively, the EU median Kd suspended solids, the EU median Kd sediment and the scenario-specific Kd values as calculated from WHAM VI (Kd WHAM).
C) Considering sediments as "wet soils" allows for a comparison between the HC5-50 values, derived from sediment NOECs with OC normalization and the HC5-50 values derived from soil NOEC data and soil bioavailability models (pH, OC and CEC normalizations).

However, the applicability of this approach has not been checked. Some conceptual limitations to this approach are mentioned below:

- The anaerobic sediment conditions (AVS binding) are not yet considered in the above approach.
- Soil SSDs are often dominated by ecological processes with little relevance to aquatic/sediment communities and these factors are not covered by bioavailability corrections.
- Bioavailability relationships for metals in soil are governed by CEC and pH, whereas for sediments these involve sulphides, organic carbon, and Fe/Mn oxides.

If sediment RA is to be carried out in a tiered approach, the lower tier could be addressed with individual level toxicity testing as is the case for current RA procedures, though with the differences outlined in the answers for the above questions. Additional tiers may cover field and microcosm studies which address semi-realistic community structures and/or the new approach for population effect models. It should be noted however, that at each tier, the statistical power of the test to predict individual effects is reduced. The extrapolation to population level effects using mechanistic effect models is receiving significant attention. These types of models have in recent years become increasingly recognized in the scientific community for the potential in predicting risk of chemicals at the population level. Several new models with different degrees of complexity have been developed during the last few years, and guidelines for good modelling practice are now available. So far the models are not in routine use, but certain regulatory authorities seem to be slowly opening up for the possibility of including them as one out of several different tools for RA of plant protection products (PPPs). As the models cover all environmental compartments a generic discussion on this approach is out of the scope of this workshop; however it should be noted that some currently available models address sediment organisms.

In principle, information for terrestrial organisms might be used for screening in a similar way but this approach has been validated to an even smaller extent. A study performed by Dresden University in conjunction with Fraunhofer Institute found reasonable correlation between FW and sediment effects while the correlation with soil was not so clear. In a recent assessment, a volatile substance was found to be totally lost for the soil compartment while it was retained in the sediment compartment due to its poor water solubility. Thus it is not necessarily simple to read across between soil and sediment assessments due to differences in fate and physical properties. Nevertheless, an option to be considered is the use of this information in WoE for supporting the development of the testing strategy, e.g. high toxicity on soil microbial functions could trigger the need for testing microbial sediment functions, while indications that soil invertebrates are much more sensitive than terrestrial plants and soil microbial functions would trigger a testing strategy focusing on sediment invertebrates.

Elements for discussion
What is the value of an EPM based PNECsediment? For substances with low adsorption, is a new PEC/PNEC needed if the value is the same as that for the pelagic compartment? For adsorptive substances, is the additional factor of 10 justified? Should the effects on microbial functions and plants be always considered in sediment RA? Could data from non-sediment tests be used in a WoE for deciding the need for testing on these taxonomic groups? What are the criteria for selecting sediment invertebrates to be tested? What should be the basis for developing an Integrated Testing Strategy for sediment?

5. Addressing bioavailability and uncertainty in sediment effect assessment.

Bioavailability data can reduce uncertainty by providing more relevant info on exposure concentrations. This leads to a more realistic exposure assessment as compared to the conservative assumptions derived from bulk sediment chemistry alone. Bioavailability is determined by the chemical/physical properties of the compound and the environment-which would presumably be evaluated in the exposure assessment. In addition, bioavailability is a function of many biological/ecological conditions such as life stage and feeding habits, which should be evaluated in the effects assessment. Thus, bioavailability needs to be considered for both exposure and effects assessment, as does bioaccesibility. For the effects assessment, the biological responses associated with tolerance/detoxification and the relative costs of any tolerance (acclimation and adaptation) must also be considered. Wherever possible, it is preferable to include consideration of bioavailability for each receptor in the effects assessment. Bioavailability should best be dealt with on a local scale with site specific measurements. Acknowledging that this type of information is not always at hand a more generic bioavailability correction using realistic worst case conditions could be used as a first tier. It is likely that a generic or at best semi-generic assessment is all that will be possible at this time. As mentioned in Q1, the bioavailability of a substance between marine and FW compartments can change due to salting out. Such information could be included in a semi-generic assessment. It may be possible to get a rudimentary understanding of the difference between bioavailability of a substance to epi-benthic organisms in a lotic and a lenthic system or between benthic and epi-benthic organisms to help obtain a series of risk assessments without multiplying the number of studies. As mentioned earlier deposit-feeding organisms have developed mechanisms which make them efficient in extracting chemicals out of the sediment. Therefore some sediment bound chemicals may be more bioavailable to deposit-feeders as a group compared to benthic organisms with other feeding strategies and to an even higher degree when compared to pelagic organisms. Subsequent tiers can be evaluated by increasing levels of bioavailability normalization. Substances known to have complex interactions, such as biotic ligand interactions, should be evaluated on an individual basis.

Since bioavailability of sediment-associated contaminants are quantitatively and qualitatively different for different feeding groups, we need to apply different dosimetry for feeding groups. The use of freely dissolved concentrations clearly applies to the exposure assessment (PEC derivation) process, but is also vital for effects assessment. The uncertainty in long term chronic testing derived from e.g. sediment aging processes or biodegradation is an important factor that should be considered in toxicity testing. The use of passive sampling methodologies to determine equilibrium concentrations in the sediment and true exposure concentrations during the test is key, and will lead to reduced uncertainty in sediment toxicity testing.

Elements for discussion
Do you agree that bioavailability must be considered in both exposure and effect assessments? What should be the basic principles for developing a tiered approach for dealing with bioavailability in sediment effect assessments? Which metrics should be used in exposure/effect comparisons in low tier bioavailability approaches? And in higher tier approaches?

There is significant measurement uncertainty in many ecological endpoints, and steps should be taken to quantify this uncertainty during the data collection phase and propagate this uncertainty through the analysis phase. Assessments of ecological effect based on departure of some observed condition from an expected reference condition should express the natural temporal and spatial variability of that reference condition, and incorporate that variability in the assessment itself. Other assessment approaches can use bootstrapping or Monte Carlo simulations to generate distributions of parameter estimates or confidence intervals on predictions. Developing guidance for incorporating results of field-based data is of particularly high importance, given the debate over the use of such data in recent discussions of substances under the Water Framework Directive.

In addition to general tools accounting for uncertainty in any effect assessment, specific elements should be considered for sediments. Qualitative ways to address the uncertainty of sediment effects assessment would be to make some assessment (or critique) of the quality/relevance of tests selected to derive the PNEC, for example by answering the following questions:

- Was the species selection and test design likely to have assessed exposure to the sediment contaminant by all the major pathways?
- For each species used, was the sediment appropriate for the organism? e.g particle size distribution, organic carbon quality/quantity?
- Was the food incorporated into the sediment or untreated and could this have influenced the endpoint substantially?
- Was the sediment spiking method suitable and were interfering effects of any solvent (carrier) apparent/likely?
- What is known about species sensitivity for the contaminant or class of contaminants in question?
- Is there a mode of action that has not been properly addressed by the species and endpoints selected?
- Are other species likely to be more sensitive or at greater risk of exposure in the receiving environment, for example if bioturbation might influence bioavailability to epi-benthic and pelagic organisms.
- Are there factors related to the binding mechanism for the chemical that may suggest different species, inhabiting different sediment types (different particle size, mineral component, organic carbon types) may be at greater risk?

Elements for discussion
Do you agree with the questions above? Which elements should be added/removed/modified?

Annex. Questions for TOPIC 3: Effect assessment

1. Should the sediment compartment effect assessment distinguish between the epi-benthonic/benthonic, lotic/lentic, marine/estuarine/freshwater communities?

2. What is the relevance of the effect assessment for benthic and epi-benthonic organisms?

3. Which benthic and epi-benthonic taxonomic groups, feeding strategies, micro habitats, endpoints and ecosystem functions should be considered?

4. Which experimental tools are available for measuring and predicting the effects on benthic and epi-benthonic organisms?

5. How should inter-species sensitivity be considered?

6. Are the principles for Species Sensitivity Distribution approaches applicable to sediment communities? How can this approach be implemented in practice? Please mention applicability, minimum requirements and limitations.

7. What should be the minimum data requirements for establishing a Predicted No Effect Concentration? Which lower and higher tier methods can be used?

8. How can ecotoxicological information on pelagic and terrestrial organisms be used for screening and assessment purposes on sediment organisms?

9. Is a generic assessment of bioavailability under the exposure assessment sufficient or should bioavailability be part of the effect assessment and discussed independently for each ecological receptor?

10. Which ways are available for investigating and expressing the uncertainty of the effect assessment?